ANN.BULL.RILP(1991)No.25, 91-98

EFFECTS OF PRECEDING VOWELS ON THE VOICING DISTINCTION FOR STOP CONSONANTS AMONG HEARING IMPAIRED SUBJECTS

Akiko HAYASHI*, Satoshi IMAIZUMI, Takehiko HARADA**,
Hideaki SEKI*** and Hiroshi HOSOI****

1. Introduction

In order to provide suitable hearing-aids and evaluation, the effects of the ear's temporal resolution on the speech perception of hearing impaired subjects have been investigated. Concerning this issue, we have already reported the following results[1,2,3]. 1) For some sensori-neural hearing-impaired listeners, the ear's temporal resolution tends to be poorer than for normal hearing listeners. 2) For normal-hearing listeners, the lower the presentation level, the poorer the temporal resolution. 3) The deterioration in temporal resolution may affect the voicing distinction of stop consonants.

Results 1) and 2) were obtained from measurements of the ear's temporal windows with normal and hearing-impaired subjects. The shape of the temporal window was measured using a temporal masking technique with non-speech stimuli according to the method proposed by Moore et al[4,5]. The results showed that the width of the temporal window was broader for some hearing-impaired subjects than for the normal hearing subjects. This means that hearing-impaired patients need a longer silent interval to recover from forward and backward masking than normal hearing listeners. For normal subjects, as the stimulus level decreased the width of the ear's temporal window increased. The temporal window shape indicated that the forward masking was more prolonged than the backward masking.

Result 3) was obtained from the measurement of the VOT phoneme boundary of bilabial plosives for CV(/pa-ba/) and VCV (/apa-aba/) stimuli. For the hearing-impaired subjects having a broad temporal window, a longer VOT value was needed to identify a stimulus as a unvoiced consonant. For both the normal and the hearing impaired subjects, the VOT phoneme boundaries were significantly longer for the VCV stimuli than for the CV stimuli. The difference in the phoneme boundary between the CV and the VCV contexts tended to be larger for the hearing-impaired subjects than for the normal subjects. It was assumed that the partial masking by the preceding vowel may have affected the voicing distinction for the following plosive consonant.

---

*The Research Institute of the Education for Exceptional Children, Tokyo Gakugei University; ** Department of Otolaryngology, Faculty of Medicine, University of Tokyo; *** Department of Computer Science, Chiba Institute of Technology; **** Department of Otolaryngology, School of Medicine, Kinki University

If this assumption is true, the following hypotheses seems valid for the voicing distinction for stop consonants in VCV sequences. H1) The lower the level of the preceding vowel, the higher the identification rate for an unvoiced stop should be. H2) The longer the silent interval between a preceding vowel and a following stop consonant burst, the higher the identification rate for an unvoiced stop should be. H3) For hearing-impaired subjects having poorer temporal resolution, a longer silent interval between a preceding vowel and a consonant burst should be required to identify unvoiced stop consonants. H4) At a lower presentation level, normal subjects should need a longer silent interval to identify unvoiced stops, as their temporal resolution becomes poorer.

In this study, to verify H1)-H4), the voicing distinction for stop consonants in VCV sequences, with changing levels in the preceding vowels and in the silent interval between preceding vowels and consonant bursts, is examined.

2. Method

2.1 Subjects

Thirteen normal hearing subject (ages:21-35 years) and 3 hearing-impaired subjects (4 ear's) participated in this study. Fig.1 shows the audiograms and the ages of the hearing-impaired subjects. In Fig.1, KL and KR indicate the left and right ear of Subject K.

2.2 Stimuli

In this experiment, the stimuli were VCV (/apa-aba/) sequences with a VOT of 30ms. It was confirmed in our preliminary study[3)] that a bilateral plosive with a VOT of 30ms was perceived as /pa/ in a CV context but as /aba/ in a VCV context (with no silent interval between the V and CV) by normal and hearing-impaired subjects.

The stimuli were generated using a Klatt type formant synthesizer[6)] with synthetic parameters basically similar to those used by Kuhl[7)]. Fig.2 illustrates the stimulus configuration (a) and the contour of each formant transition (b). Following the release of the burst (at 0ms), the VOT was conducted by a cutback of the first formant and an excitation of the higher formants with a noise source simulating aspiration instead of the periodic source during the cutback. The amplitude of this noise source fell linearly until the VOT as shown in Fig.2(a). The preceding and following vowels were 300ms in duration with rise/fall times of 10ms. The first VC transition contour was set symmetrically to the following CV transition as shown in Fig.2(b).

The variable parameters were the stimulus level differences between the first vowel and the second vowel (da) and the silent interval between the first vowel and plosive consonant (dt). For the da changing condition, the silent interval(dt) was fixed at 0ms.

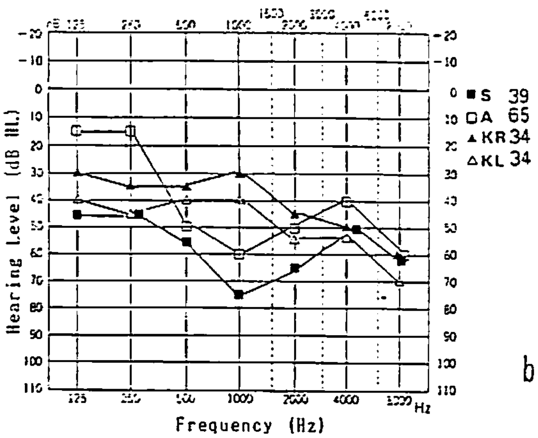
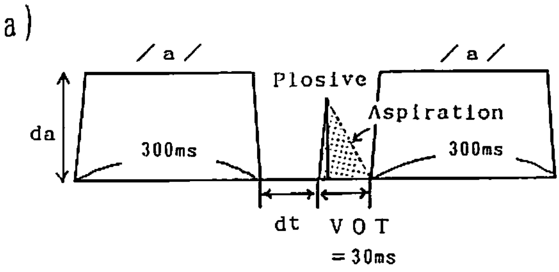


Fig.1 Audiograms and ages of the hearing-impaired subjects.



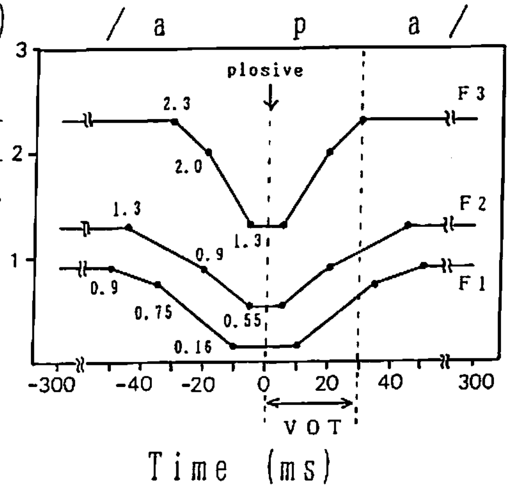


Fig.2 a) Stimulus configuration.

b) Contour of each formant transition.

2.3 Procedure

The phoneme boundary between /apa/ and /aba/ was measured separately for various levels of the first vowel (da condition), and various silent intervals (dt condition). For the da condition. The da was changed in 5dB SPL steps. For the dt condition, the dt was changed in 5ms steps for the normal subjects and in 10ms steps for the hearing-impaired subjects. All stimuli were randomly presented ten times each. The subjects identified the consonant in each stimulus as either /b/ or /p/. The 50% response level da (in dB) and dt (in ms) were used as the phoneme boundary.

For two experimental conditions (da and dt conditions), the level of the second vowel was set at the most comfortable level individually for the hearing-impaired, but was set at 50dBSPL and 80dBSPL for the normal-hearing subjects.

3. Result

For the da condition, Fig. 3(a) shows the percent of /b/ responses at the two presentation levels for a normal subject. The presentation level was the level of the following vowel, and was set at 50 and 80dBSPL. When the stimulus level difference (preceding vowel level minus following vowel level) increased in other words, when the preceding vowel's level decreased, the percent of /b/ responses decreased for both presentation levels, although the VOT phoneme boundary was different.

Fig. 3(b) shows the percent of /b/ responses at one level for the hearing-impaired subjects. Also as shown in this figure, the percent of /b/ responses decreased when the preceding vowel level decreased.

For the dt condition, Fig.4 (a) shows the percent of /b/ responses for a normal subject. The percent of /b/ responses decreased when the silent interval, dt, increased. The dt phoneme boundary shifted from 20ms to 42ms when the presentation level was changed from 80dBSPL to 50dBSPL.

Fig.4 (b) shows the percent of /b/ responses for the hearing-impaired subjects. The percent of /b/ responses decreased when the silent interval, dt, increased. The dt phoneme boundaries tended to be longer than those of the normal subjects.

Fig.5 shows the da phoneme boundary distribution and Fig.6 the dt phoneme boundary distribution for the normal and the hearing-impaired subjects.

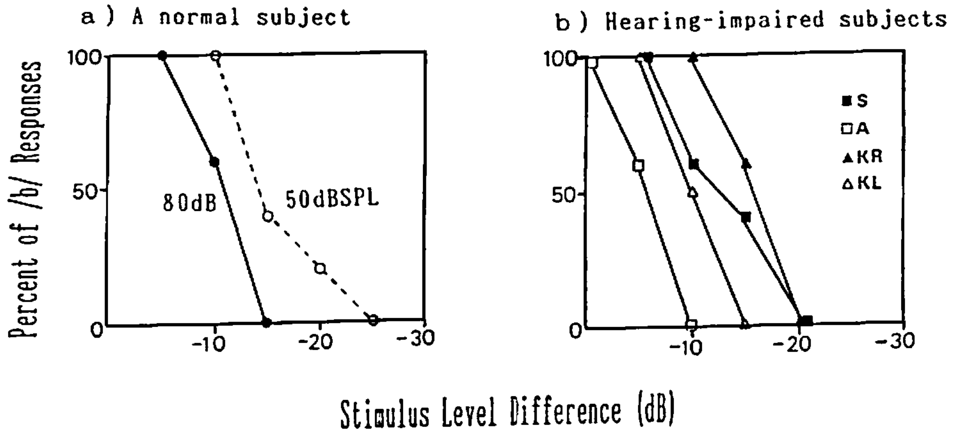


Fig.3 The percent of /b/ responses for the da condition.
a) A normal subjects at two presentation levels.
b) Hearing-impaired subjects.

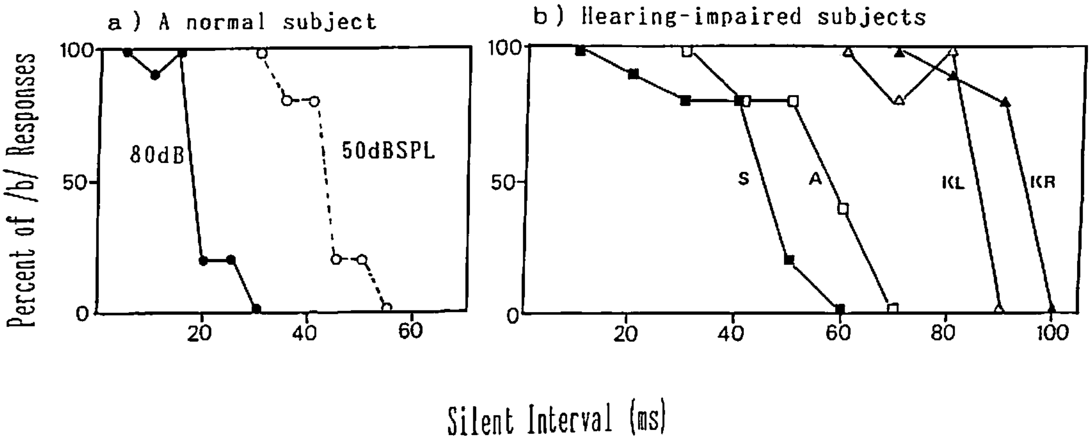


Fig.4 The percent of /b/ responses for the dt condition.
a) A normal subjects at two presentation levels.
b) Hearing-impaired subjects.

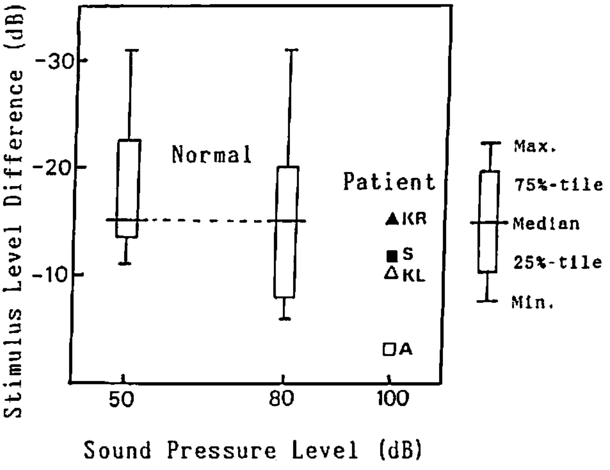


Fig.5 The da phoneme boundary distribution.

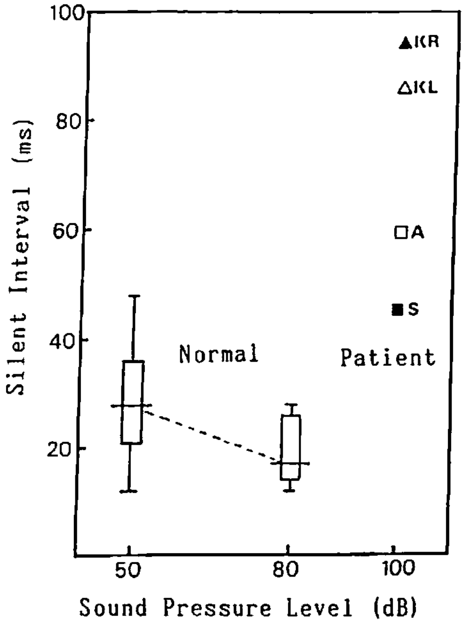


Fig.6 The dt phoneme boundary distribution.

The results shown in Figs.5 and 6 indicate that the voicing distinction for stop consonants in VCV sequences was affected by the level of the preceding vowel and the silent interval between the preceding vowel and the consonant burst. The identification rate for the voiced stop /b/ decreased when the preceding vowel's level decreased and the silent interval became longer.

For the normal subjects, the difference in the dt phoneme boundary between the two presentation levels was significant (t=-2.34, f=24, p<0.05), while the difference in the da phoneme boundary was not significant.

4.Discussion

The results shown in Figs.3, and 4 indicate that the voicing distinction for stop consonants in VCV sequences was affected by the preceding vowel level and the length of the silent interval between the preceding vowel and the consonant burst. The identification rate for the unvoiced stop /p/ increased when the preceding vowel's level decreased and/or the silent interval was longer. Accordingly, H1) and H2) were confirmed.

It seems that these results can be explained using our model of voicing perception. An outline of our temporal masking model for VOT perception is illustrated in Fig.7(a) for the da condition and in Fig.7(b) for the dt condition.

For the da condition, when the level of the preceding vowel decreases from da1 to da2, the falling time of the excitation pattern, or loudness, for da2 becomes shorter than for da1. Therefore, although the acoustical VOTa is same, the psychoacoustical VOTp2 is longer than the VOTp1. Consequently, it can be expected that the identification rate for unvoiced stops increases for lower das.

Similarly, VOTp changes by with changes in the silent interval between the vowel and the following consonant burst. As shown in Fig.7 (b), when the silent interval is lengthened from dt1 to dt2, VOTp2 becomes longer than VOTp1. Accordingly, the identification rate for unvoiced stops becomes higher when the silent interval is longer.

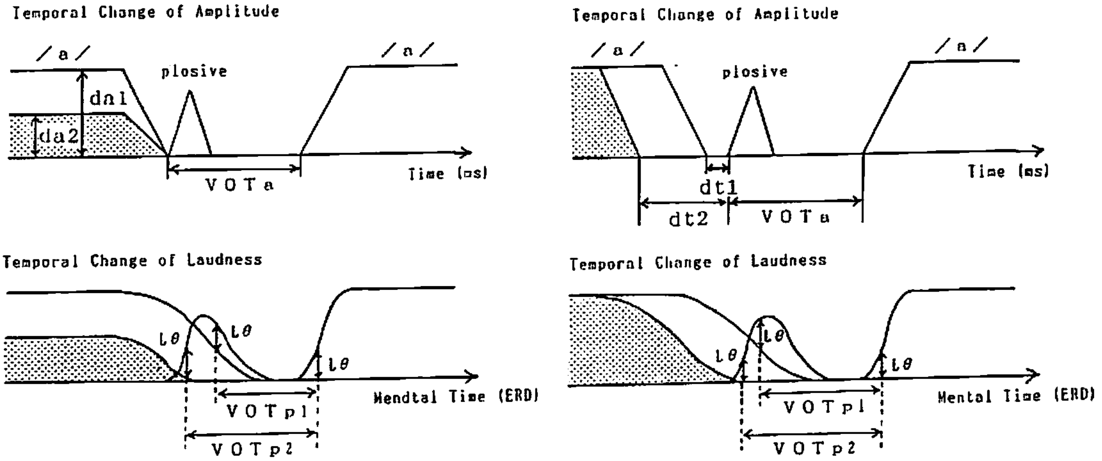
As the temporal resolution becomes poorer, the falling time of the loudness for the preceding vowel is lengthened, and thus, the length of VOTp is shortened. In addition, according to our preliminary study, temporal resolution deteriorate for some of the senso-ri-neural hearing-impaired and also for normal subjects when the presentation level is significantly low. Consequently, subjects having poorer temporal resolution require a larger level difference between a preceding and following vowel (da), or a longer silent interval to identify an unvoiced stop consonant.

As shown in Fig.6, the dt phoneme boundaries were significantly longer for the hearing-impaired than for the normal subjects. Also, for the normal subjects, the dt phoneme boundary was significantly longer at the low presentation level than at the high level. From these results, H3) and H4) were confirmed.

The other possible explanation for the present results is the effect of the auditory filter shape or auditory filter ringing. If the auditory filter ringing caused by the preceding vowel is processed as a buzz bar in the ear, this would increase the percentage of /b/ responses when the silent interval is decreased or the preceding vowel level is increased.

Subjects with an inner ear impairment tend to have a wider auditory filter, and thus this ringing must be expected to be shorter than in normal subjects. Subjects with retrocochlear hearing loss may have normal auditory filters if they have no impairment to their inner ear, and thus at least this ringing should be the same as in normal subjects. Therefore, although ringing is a possible factor in our present results, it seems incapable of completely explaining why hearing impaired subjects need longer silent intervals to identify unvoiced stops.

Ringing effects should be examined further to obtain a cleaner explanation of our present results.



a) The da condition

b) The dt condition

Fig.7 Temporal masking models of VOT perception.

5. Conclusion

The voicing distinction for bilabial stop consonants was examined in VCV sequences changing the level of preceding vowels or the silent interval's length between a preceding vowel and a following consonant. The following results were obtained.

1) The identification rate for the voiced stop /b/ decreased when the preceding vowel level decreased when the condition that VOT was fixed at 30ms.

2) The identification rate for the voiced stop /b/ decreased when the silent interval's length between a preceding vowel and the consonant burst increased under when the VOT was fixed at 30ms.

3) Sensori-neural hearing-impaired subjects needed a longer silent interval compared to normal subjects to identify /p/.

4) Normal subjects needed a longer silent interval to identify /p/ when the presentation level was lower.

These results suggest that the voicing distinction of for stop consonants in VCV sequences is affected by forward masking from the preceding vowel.

Acknowledgement

This work was supported in part by a Grant-in Aid for Scientific Research, the Ministry of Education, Science and Culture, Japan.

References

1) Yamada,A., Imaizumi,S., Harada,T., Mikami,Y., Decuchi,T. andHosoi,H.:Effects of temporal factors on the speech perception of the hearing-impaired. -- A preliminary report --, Ann.Bull.RILP., 21, 131-140, 1987.
2) Hayashi,A., Imaizumi,S., Harada,T. and Hosoi,H.: Effects of stimulus duration and inter-stimulus interaction on vowel intelligibility for normal and hearing-impaired subjects, Ann.Bull.RILP., 23, 163-172, 1989.
3) Hayashi,A., Imaizumi,S., Harada,T., Seki,H. and Hosoi,H. : Relationships between the ear's temporal window and VOT perception for normal and hearing-impaired listeners, Ann.Bull.RILP., 24, 145-156, 1990.
4) Moore,B.C.J., Glasberg,B.R., Plack,C.J., and Baswas,A.K.:The shape of the ear's temporal window, J.Acoust.Soc.Am., 83(3), 1102-1106, 1988.
5) Plack, C.J. and Moore, B.C.J.: Temporal window shape as afunction of frequency and level, J.Acoust.Soc.Am., 87(5), 2178-2187, 1990.
6) Klatt,D.M.:Software for a cascade/parallel formant synthesizer, J.Acoust.Soc.Am., 67(3), 971-995, 1980.
7) Kuhl,P.K. and Miller,J.D.:Speech perception by the chinchilla:Identification function for synthetic VOT stimuli., J.Acoust.Soc.Am., 63, 905-917, 1978.