Child development of adaptive phonetic gestures observed in vowel devoicing

Satoshi Imaizumi[*], Kiyoko Fuwa[+], Yoko Shimura[+]
[*]*Dept. of Speech & Cognitive Sciences, University of Tokyo, Tokyo, 113-0033 Japan*
[+]*Faculty of Education, Saitama University, Saitama, 338-0825 Japan*

Abstract

The developmental properties of vowel devoicing were investigated for 72 children of 4 and 5 years of age, and 37 adults in two dialects of Japanese. One was the Osaka dialect, with a low devoicing rate, and the other the Tokyo dialect, with a high devoicing rate. In the Tokyo dialect, the devoicing rate of children significantly increased and reached an adult-like level by the age of 5 years, whereas it remained low irrespective of age in Osaka. The vowel devoicing of 5 year-old children exhibited the same characteristics as that of the adults of their respective dialect. These results suggest that children growing up with the Tokyo dialect acquire the articulatory gestures which do not inhibit vowel devoicing by the age of 5 years, whereas children growing up with the Osaka dialect acquire those which inhibit the devoicing of vowels by the same age.

1. INTRODUCTION

It is well known that high vowels between voiceless consonants are often devoiced in many languages such as Korean, English, French, and German, as well as in many Japanese [1-6]. There has been a lively controversy, however, whether vowel devoicing should be described as allophonic variation [2, 3] or as a lower-level coarticulatory effect [4, 6]. This controversy results from the fact that the probability of such devoicing is dependent on phonological [2, 3], phonetic [4, 5], pragmatic [6], and dialectal conditions [1].

Based on acoustic analyses of listener-adaptive characteristics of vowel devoicing in Japanese dialogues, Imaizumi *et al.* [6] suggested the hypothesis that vowel devoicing is a consequence of the adaptive organization of the articulatory gestures which reduce or permit gestural overlap and hiding [7]. Fig. 1 schematically represents the possible laryngeal adjustments related to vowel devoicing [6], where Ab_1 and Ab_2 represent the devoicing gestures of the laryngeal abductors, like the posterior cricoarytenoid, for the unvoiced consonants C_1 and C_2, and A_d represents the voicing gesture of the adductors, like the interarytenoid, for the vowel V in a C_1VC_2 context. Two different strategies can be suggesting as possible controlling vowel devoicing. Strategy 1 (S_1) simply involves a lengthening/shortening of the interval between the successive devoicing gestures, as represented by the arrow S_1 in Fig. 1. S_1 is dependent on the articulation speed. Strategy 2 (S_2) involves adjusting the size of the gestures and changing the relative size of the voicing and the devoicing gestures, as represented by the arrow S_2. This strategy could be independent of the articulation speed.

If this hypothesis is valid, the organization of articulatory gestures should be dialect-dependent, since the probability of occurrence of vowel devoicing is dialect-dependent. When children become members of a particular dialect community, they have to acquire this dialect-dependent organization of articulatory gestures, particularly in terms of size and timing. Children in Osaka have to learn the organization of articulatory gestures to inhibit devoicing, while those in Tokyo do not have to inhibit devoicing. When and how do children acquire such adaptive articulatory organizations? This is an important issue in development of motor control for speech in children.

2. METHOD

2.1. Materials

Two birds and two other animals were used to elicit utterances meeting the following three phonological conditions for inducing vowel devoicing: P1) they contained one of high vowels /i/ and /u/; P2) the vowel was positioned between voiceless consonants; and P3) the syllable with the high vowel had a lack of accent.

Table I shows the stimulus words with their surface pitch patterns in terms of an HL description, where H and L indicate a high or low surface pitch level, respectively. The underlined vowel was the target of our analysis. The words had common accent types whether spoken in the Osaka or Tokyo dialect, except for /kitsune/ which is LHH in Tokyo but LLH in Osaka. The second /tsu/ in /kitsutsuki/ was the only accented mora analyzed here. Most of the analyzed moras were marked L, and thus not accented. The only exception was the /ku/ which is marked H in /hakutʃoo/, and was not accented since it was followed by an H mora. Among the analyzed moras, the accented second mora of /kitsutsuki/ is not usually devoiced in either dialect according to phonological condition P3. On the other hand, the first mora of /kitsutsuki/ and of /kitsune/, the third of /kitsutsuki/, the /ku/ of /hakutʃoo/ and the /hi/ of /hitsudʒi/, all of which are unaccented, are usually devoiced with a high probability only in the Tokyo dialect.

The four-mora word /kitsutsuki/, consisting of high vowels and voiceless consonants, was of particular interest. Based on phonological condition P3, the vowel /u/ of the second mora /tsu/ should be voiced, since it carried accent. The first and third moras, which were adjacent to this accented mora, should be devoiced, particularly in the Tokyo dialect. The last mora should be voiced following phonological condition P4 -- that is, devoicing in successive moras should be avoided in order to avoid a series of consonant clusters at the surface level which is not a favored structure in Japanese. This prediction, however, does not agree with phonological condition P5 – that is, word-final high vowels following a voiceless consonant tend to be often devoiced in the Tokyo dialect.

2.2. Subjects

Kindergarten children of 4 and 5 years of age growing up within the Osaka dialect (O4 and O5) and children within the Tokyo dialect (T4 and T5) participated as speakers in the test. Children whose parents were from the prescribed dialect group were

selected based on reports made by kindergarten teachers who recorded their utterances of the names of the four familiar animals described above. Children who could not name any of the four names were excluded from further analyses. Although ten 3-year-old children also participated, only four of them could name all of the test words. Therefore, 3-year-old children were excluded from the analyses. Children older than 3 years of age, 2 4-year-old and 2 5-year-old children in Osaka and 4 4-year-old and 3 5-year-old children in Tokyo, were also excluded since they named an animal in an unexpected way such as /tori/ (a bird), /kitsuki/ (deletion of the 3rd mora), or /kitsukitsu/ (mora metathesis) for /kitsutsuki/. Errors in accentuation such as LLH for LHH or in consonant articulation such as /tʃu/ for /tsu/ were carefully checked but not excluded from further analysis, since those errors are commonly found in children's speech. Thus, the children analyzed were 25 4-year-old and 18 5-year-old children in Osaka and 15 4-year-old and 14 5-year-old children in Tokyo. As controls, 19 adults speaking Osaka dialect (OA) and 18 speaking Tokyo dialect (TA) also participated.

2.3. Data Analysis

The analysis method and parameters adopted were the same for the children's speech as for the adult speakers. Using an acoustic analysis system which generated a wide-band sound spectrogram [6], the length of the unvoiced segment, U_m and that of the voiced segment, V_m,, and their sum, $L_m = U_m + V_m$, could be measured for each target mora, M_m.. To monitor the articulation speed for an assessment of its possible influence on the devoicing rate, the word length, T_u, was also measured for each utterance *u*.

M_m was classified as "Devoiced" if $V_m = 0$. The devoicing rate, *D*, was calculated as the percent of the number of moras classified as "Devoiced" against the total number of moras analyzed for each word and each speaker group.

Test word	Corresponding English	Tokyo dialect	Osaka dialect
/kitsutsuki/	woodpecker	LHLL	LHLL
/hakutʃoo/	swan	LHHH	LHHH
/kitsune/	fox	LHH	LLH
/hitsudʒi/	sheep	LHH	LHH

Table I. Test words and their accent types.

Dialect	Speaker Group	No. Subject	Mean Age Year	Mean Age Month
Osaka	O4	25	4	4.4
	O5	18	5	5.9
	OA	19	22	2.2
Tokyo	T4	15	4	2.7
	T5	14	5	5.3
	TA	18	21	7.1

Table II. Profile of the participants.

Speaker Group	Utterance Length (ms)	Devoicing Rate (%) M1	M2	M3	M4
O4	1037	22	0	14	9
O5	885	28	0	12	6
OA	635	36	15	40	0
T4	935	42	0	15	30
T5	807	62	0	54	15
TA	791	88	0	73	13

Table III. Devoicing rate and utterance length of the four moras, M1-M4, of the word /kitsutsuki/.

3. RESULTS

3.1. Devoicing Rate

With regards to the devoicing rates within the dialect groups for the first mora of /kitsutsuki/, as shown in Table III, the chi-square test revealed that those of O4, O5 and OA did not significantly differ from each other ($p>0.43$), whereas those of TA and T5 were significantly higher than that of T4 ($p=0.038$). The devoicing rate tended to increase with age in both dialects, although significant differences were found only in the Tokyo dialect. With regards to the devoicing rates between the same age groups across dialects, those of O4 and T4 were low and showed no significant difference ($p=0.22$), whereas those of T5 and TA were significantly higher than those of O5 and OA, respectively ($p<0.01$), but showed no significant difference between themselves ($p=0.11$).

A similar tendency, with some differences, was observed for the other tested moras which are supposed to be devoiced in the Tokyo dialect. The devoicing rate tended to increase with age within each dialectal group. OA had a significantly higher devoicing rate than O4, but not O5, for the third mora /tsu/ of /kitsutsuki/ ($p=0.04$) and for /ku/ in /hakutʃoo/. For the other moras, there was no significant age-related difference within the Osaka group.

On the other hand, T5 and TA had significantly higher devoicing rates than T4 ($p<0.04$) for all the moras tested. The devoicing rates of T5 and TA were significantly higher than those of O5 and OA, respectively ($p<0.01$), but showed no significant difference between themselves ($p=0.28$). For /ki/ in /kitsune/ and /ku/ in /hakutʃoo/, the devoicing rate of T4 was significantly higher that that of O4 ($p=0.01$). On the other hand, the devoicing rates for the second and the fourth moras of /kitsutsuki/, which are not supposed to be devoiced, were low irrespective of age and dialect.

3.2. Segment Length

The devoiced segment length, U_m , and the voiced segment length, V_m , as shown in Fig. 2, varied depending on age and dialect. For most of the target moras, the devoiced segment was longer for the Tokyo speakers than for the Osaka speakers irrespective of age, whereas the voiced segment was significantly shorter for the older speakers than for the younger ones in both dialect groups and was significantly shorter in the Tokyo groups than in the Osaka groups -- particularly so in T5 and TA.

The older speakers had a significantly shorter mora length than the younger speakers, without significant effects of dialect for most of the devoicable moras tested. This was also true for utterance length.

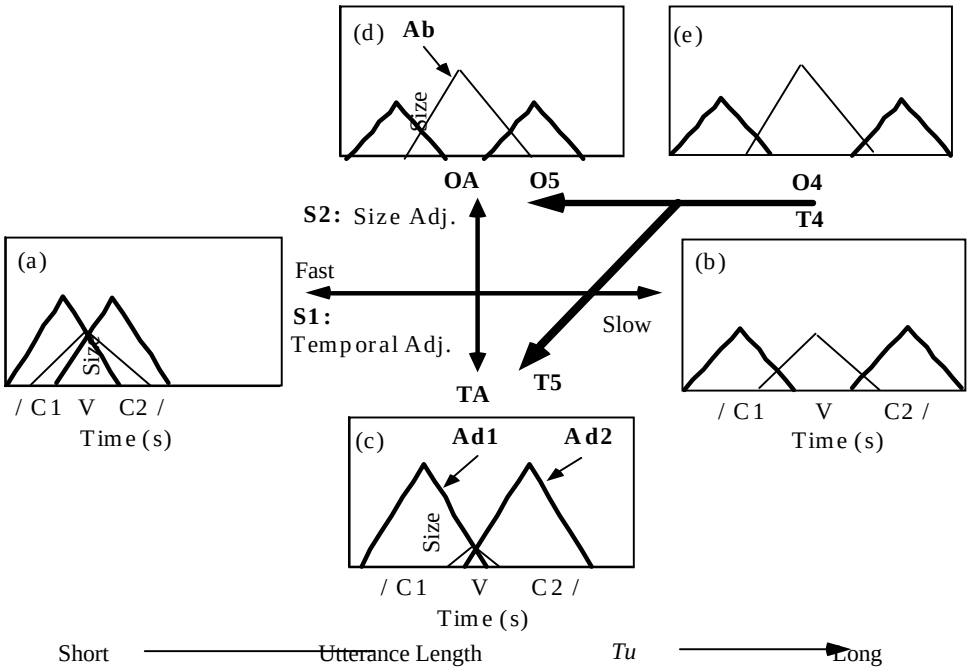
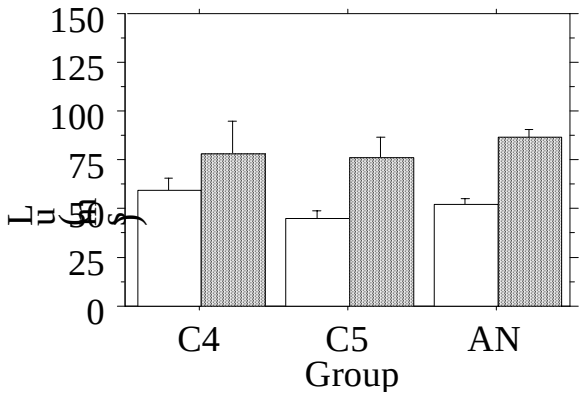
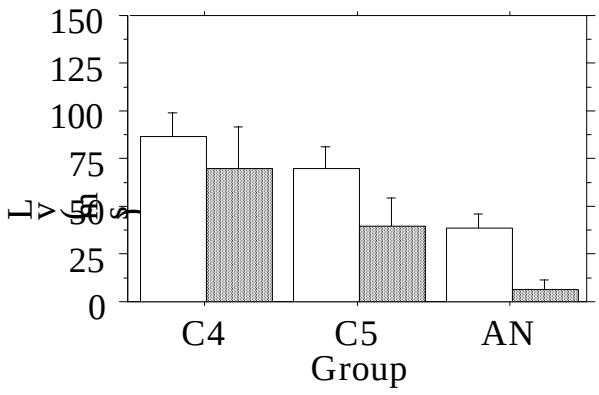


Fig. 1. Schematic representation of two different control strategies for vowel devoicing. Ab_1 and Ab_2 are the devoicing gestures for the unvoiced consonants C_1 and C_2, while A_d is the voicing gesture for the vowel V in the C_1VC_2 context. There are two component strategies to control vowel devoicing: Strategy 1 (S_1) simply involves lengthening or shortening the interval between successive devoicing gestures affecting utterance length -- compare (a) vs. (b). Strategy 2 (S_2) involves adjusting the size of the articulatory gestures independent of utterance length -- compare (c) vs. (d). 4-year-old children acquire articulatory gestures at a slow articulation speed with large voicing gestures as represented by panel (e). Articulation speed increases with age, and children exhibit dialect-specific strategies by the age of 5 years. Osaka children adopt articulatory gestures with increased voicing gestures, resulting in a low devoicing rate, as represented by panel (d). On the other hand, Tokyo children reduce their voicing gestures resulting in a high devoicing rate.



• • • • (a) *Lu* of the first mora /ki/ in /kitsutsuki/



• • • • • (b) *Lv* of the first mora /ki/ in /kitsutsuki/

Fig. 2. The average length of the unvoiced segment and of the voiced segment for the first mora of /kitsutsuki/. C4 includes O4 and T4, C5 to O5 and T5, and AN to OA and TA. Dark bars are for the Tokyo speakers while white bars for the Osaka speaker.

4. DISCUSSION

The results can be summarized as follows:

(1) Although the children tended to produce more errors than the adults both in accentuation and consonant articulation, no tokens were assessed to have consonant clusters -- such as /kts/, /tsk/,or /hts/ -- which are not permitted in Japanese.

(2) TA and T5 had significantly higher devoicing rates than the other speakers for the devoicable moras tested, although there were some age-related increases in the devoicing rate in both dialect groups. There were no significant differences among the speaker groups in the devoicing rate for the tested moras which should not be devoiced.

(3) The devoiced segments were longer for the Tokyo speakers than the Osaka speakers, whereas the voiced segments were significantly shorter for the older speakers than the younger ones in both dialect, and were significantly shorter in the Tokyo groups than in the Osaka groups, particularly for T5 and TA.

These results suggest that children growing up as speakers of the Tokyo dialect acquire articulatory gestures which permit vowel devoicing, whereas children growing up with the Osaka dialect acquire articulatory gestures which inhibit vowel devoicing until the age of 5 years. Only adults (TA) and 5-year-old children (T5) of the Tokyo dialect showed a significantly higher devoicing rate than that of 4-year-old Tokyo children (T4) and all of the Osaka speakers (O4, O5, and OA) tested. Although mora length and utterance length decreased significantly with age in Osaka, the devoicing rate remained significantly lower than in T5 and TA, irrespective of age. Therefore, the significant differences found in the devoicing rates of 5-year-old children and adult speakers between the two dialects can not be accounted for by differences in utterance length or articulation speed.

The 4-year-old children of both dialects in our study showed the longest utterance/mora length, suggesting the slowest articulation speed, as well as the lowest devoicing rate among the speakers tested. This result suggests that children acquire articulatory gestures at a slow speed of articulation, with clear voicing gestures for the vowel, as represented by the panel (e) in Fig. 1. By the age of 5 years, children exhibit different characteristics. In the Osaka dialect, they decrease mora length by reducing vowel length, while keeping their devoicing rate low. Therefore, they seem to acquire articulatory gestures at a fast articulation speed without significantly decreasing the voicing gesture for vowels, as represented by the panel (d) in Fig. 1. On the other hand, in the Tokyo dialect, children decrease mora length by shortening the voiced segment while lengthening the devoiced segments which increases their devoicing rate. Therefore, they seem to acquire articulatory gestures at a fast articulation speed, with reduced voicing gestures for vowels and enhanced devocing gestures for consonants, as represented by the panel (c) in Fig. 1. Lengthening the devoiced segment of a devoiced mora is important for maintaining the mora-based timing structure of Japanese. It should be noted that the 5-year-old children in Tokyo have already acquired vowel devoicing, which is not the same as vowel deletion. Vowel deletion generates undesirable consonant clusters and would destroy the Japanese mora-based timing. The vowel devoicing of T5 and TA did not result in such phonetic structures.

Articulatory development can be considered to vary depending on aspects of the target speech. Speech development may very well involve gaining greater independence among gestures that use articulators of increasingly closer anatomical proximity [8]. In this sense, vowel devoicing is a difficult task for children to acquire, because it involves the coordination of the devoicing and voicing gestures, both of which are mainly performed by a single articulator, the larynx. Furthermore, it is probably not easy for children to devoice vowel nuclei while maintaining mora-based timing without generating any undesirable consonant clusters.

5. CONCLUSIONS

Based on the hypothesis that vowel devoicing is a consequence of the organization of articulatory gestures adapted to phonological, phonetic and dialect-specific requirements, we tested when and how children acquire such an adaptive articulatory organization by analyzing vowel devoicing in two contrasting dialects of Japanese, those of Osaka and Tokyo. Tokyo adults and 5-year-old children showed a significantly higher devoicing rate than 4-year-old Tokyo children and all of the Osaka speaker groups tested. Although the length of utterance, moras and voiced segments decreased significantly with age in both dialects, the devoicing rate of Osaka speakers remained significantly lower than the adult and 5-year-old speakers of the Tokyo dialect, irrespective of age. The present results suggest that dialect-specific adaptive strategies to coordinate voicing and devoicing gestures as required to match the adult vowel devoicing pattern is acquired by the age of 5 years.

ACKNOWLEDGEMENTS

This research was partially supported by a Grant-in-Aid for Scientific Research on the Priority Area of "Development of Mind," Ministry of Education, Science and Culture, Japan.

REFERENCES

[1] Sugito, M. 1969. Vowel devoicing in Tokyo and Osaka. The Study of Sounds, Phonetic Soc. Jpn. 14, 1-4.

[2] Hirose, H. (1971). The activity of the adductor laryngeal muscles in respect to vowel devoicing in Japanese. Phonetica 23, 156–170.

[3] Yoshioka, H., Lo_fqvist, A., and Hirose, H. 1982. "Laryngeal adjustments in Japanese voiceless sound production. J. Phonetics 10, 1-10.

[4] Beckman, M., and Shoji, A. 1984. Spectral and perceptual evidence for CV coarticulation in devoiced /si/ and /syu/ in Japanese. Phonetica, 41, 61–71.

[5] Kuriyagawa, F., and Sawashima, M. 1989. Word accent, devoicing and duration of vowels in Japanese. Annual Bulletin (Research Institute of Logopedics and Phoniatrics, University of Tokyo, Japan) 23, 85–108.

[6] Imaizumi, S., Hayashi, A., and Deguchi, T. 1995. Listener adaptive characteristics of vowel devoicing in Japanese dialogue. Journal of the Acoustic Society of America, 98, 768-778.

[7] Browman, C. P., and Goldstein, L. 1986. Towards an articulatory phonology. in Phonological Yearbook 3, 219-252.

[8]Nittrouer, S. 1993. The emergence of mature gestural patterns is not uniform: Evidence from an acoustical study," Journal of Speech and Hearing Research, 36, 959-972.