INDEPENDENCE OF CONSONANTAL VOICING AND VOCOID F0 PERTURBATION IN ENGLISH AND JAPANESE

Shunichi Ishihara

Japan Centre (Asian Studies), and
Phonetics Laboratory, Department of Linguistics (Arts), The Australian National University

ABSTRACT

Data collected from Japanese and English showed that both phonetically fully voiced and (partially) devoiced allophones of /d/ have very similar perturbatory effect on the F0 of the following vowel. It is considered, therefore, that the phonetic voicing of /d/ (periodicity during the closure) is not clearly correlated with lower levels of F0 on the following vowel. Although the F0 perturbation may be caused by some aspects in the production of the preceding stop which is not necessarily manifested in actual vocal cord vibration, this result indicates that there is still a possibility that people may deliberately control the F0 of the following vowel as an additional cue to the phonological difference between voiceless and voiced stop consonants.

1. INTRODUCTION

The basic phonetic correlate of the phonemic [±voice] contrast in a stop consonant is the presence vs . the absence of vocal cord vibration during the closure of the stop. However, it has been noted in some languages that the phonemic [±voice] contrast in a stop consonant is phonetically realised on the basis of differences in VOT (Lisker and Abramson 1964). There are two types of acoustic cues distinguishing phonemic voiced and voiceless stops. In English, therefore, word-initial /voiced stops/ may be produced either with shortly-delayed phonation (short-lag VOT) or advanced phonation (VOT lead).

The first aim of this study is to see whether there is a difference or not between these two allophones of voiced stops, in terms of the perturbatory effect on the F0 of the following vowel. The second aim of this study is to reconsider the characteristics of consonantally induced F0 perturbation in the light of Tatham's (1988) allophonic classification on the basis of the results obtained for the first aim.

Following Wang and Fillmore (1961) and Ladefoged (1965, 1967), Tatham (1988) classified allophones into three types; uncontrollable intrinsic events , controllable intrinsic events , and extrinsic events . Intrinsic events can be divided into uncontrollable intrinsic ev ents ; those that must occur because there are limitations (i.e. physiological, acoustic, and neurological) on the use of available components and organs, and controllable intrinsic events ; those that can be suspended under extrinsic resistance.

The data included in this study were part of a large study on consonantally induced F0 perturbation in Japanese, Thai, and English (Ishihara 1996). The method of data collection, therefore, reflects this fact. The data obtained from the experiments on Japanese and English are presented in the present study.

2. EXPERIMENT: JAPANESE

10 native speakers of Japanese (the Tokyo pitch-accent system) who were recruited from the student and staff populations at the Australian National University (ANU), participated in this experiment. There were 5 males and 5 females between 21 and 45 years of age. All of them were from Tokyo or the prefectures adjacent to Tokyo, and finished their schooling there. A reading list containing 4 disyllabic words shown in table 1 was prepared in standard Kanji orthography. The first CV segments of these words contrast both in terms of the pitch of V; low and high pitch, and in terms of phonemic voicing type of C; [+voice]/[-voice], and these CV segments are followed by /k/. The informants were asked to read the list twice in the frame in the recording studio of the phonetics laboratory at the ANU. The material was recorded on high-quality normal position tapes using a Nakamichi tape recorder and a Nakamichi CM300 cardioid microphone.

Stimuli	taki "waterfall" L H	tako "kite" H L
	daku "to hold" L H	daki "steering wheel" H L
Frame	ima ___ to itte kuda sai [H L L L L L L LL] now part. say honorific "Please say ___ now"	

Table 1: Four test words and a carrier sentence. L and H stand for low pitch and high pitch, respectively.

The carrier sentence in which each test word was uttered was digitised with Macrecorder at a 22 kHz effective rate (sampling frequency = 22254), then edited and analysed with Signalyse version 2.08. F0 values of the first vowel of each test word were measured to see the perturbatory effect of preceding stop consonant and the duration of periodicity absence before the release of /d/—if it is (partially) devoiced—to see the perturbatory effect of devoiced /d/s. F0 values were pitch-synchronously measured with a mouse directly from digitised audio waveforms displayed on a Macintosh computer screen in a time-domain representation by calculating each period of successive vocal fold cycle. Duration of periodicity absence was also manually measured with a mouse directly from digitised audio waveforms.

2.1. Results and Discussion: Japanese

Two types of VOT; short-lag VOT and VOT lead, were recognised in /d/ from the Japanese data. Figure 1 contains a speech waveform of /da/ (L) segment of /daku/. In this production of /d/, the periodicity of /d/ dies out after 30 ms into its hold phase (short-lag VOT). Therefore, there is an absence of periodicity before the release of /d/ for 45 ms presumably because of the well-known aerodynamic conditions (cf. transglottal pressure equalisation). Table 2 shows the distribution of these two types of VOT. In the case of short-lag VOT, the duration of periodicity absence before the release of /d/ is presented in ms.

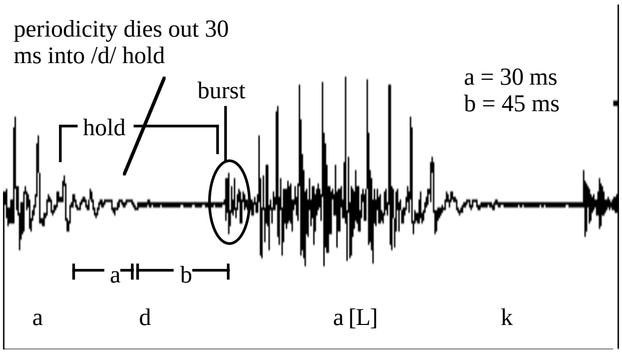

Figure 1: /da/ (L) segment of /daku/ (LH) preceded by the offset of /a/ and followed by the burst of /k/ (The first token of Speaker JM3).

	Low Pitch		High Pitch	
Speaker	1	2	1	2
JM1	-	-	-	-
JM2	73	63	30	-
JM3	45	54	30	-
JM4	10	-	31	-
JM5	17	-	18	22
JF1	8	14	15	14
JF2	-	-	-	16
JF3	-	-	-	-
JF4	28	18	27	19
JF5	-	-	-	-
Mean	33		22	

Table 2: The distribution of VOT lead and short-lag VOT in Japanese."-" means that /d/ was fully voiced in production and numerals represent the duration of periodicity absence before the release of /d/ (in ms). 1 stands for first token, and 2 for second token.

Judging from table 2, there are no clear restrictions relating to the occurrences of short-lag VOT and VOT lead for /d/. VOT lead constantly appeared in all of the /da/ segments articulated by JM1, JF3, and JF5, while short-lag VOT constantly appeared in all of the /da/ segments uttered by JF1 and JF4. Moreover, for JM2, JM3, JM4, JM5, and JF2, both short-lag VOT and VOT lead were observed. It seems, therefore, that the selection of short-lag VOT and VOT lead is idiosyncratic, or they are in free variation with each other, rather than one that is governed by phonological or phonetic rules.

Figures 2 and 3 show the F0 perturbation shapes of /ta/ (H) and /da/ (H) segments uttered by JM2 and JM5, respectively. The F0 values are graphically presented, plotting the raw F0 values as a function of glottal periods. The raw F0 values were t-tested (unpaired one tail) between the vowels following /t/ and /d/ at each period. In the t-tests, 0.05 was set as a level of significance.

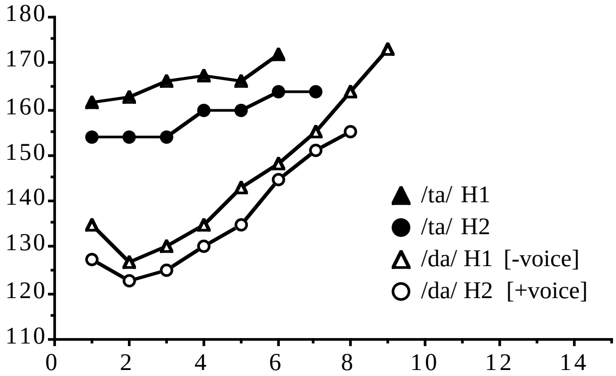

Figure 2: F0 perturbation shapes of /ta/ (H) and /da/ (H) (speaker JM2). X-axis is glottal period, and Y-axis is F0 (Hz). [+voice] stands for a fully voiced /d/, and [-voice] for a not fully voiced /d/.

In both figures, the F0 values of the vowels followed by phonemic /d/ are significantly lower than the ones followed by /t/. In figure 2, although /d/ was devoiced in the first token, the F0 perturbation shapes of the first and second tokens of /da/ (H) are very similar, and the F0 values of /da/ (H) are still significantly lower ($p \leq 0.0204$) than those of /ta/ (H) from the first period to the sixth period which is the last comparable period. In figure 3, although /d/ was devoiced for both tokens of /da/ (H), the difference in F0 values caused by /t/ and /d/ is still significant ($p \leq 0.0122$) from the first period to the fourth period.

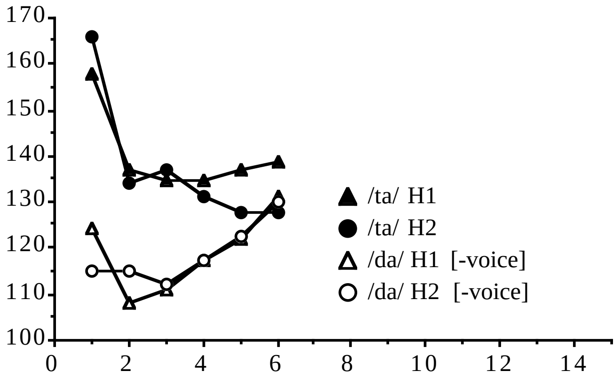

Figure 3: F0 perturbation shapes of /ta/ (H) and /da/ (H) (speaker JM5). X-axis is glottal period, and Y-axis is F0 (Hz). [-voice] stands for a not fully voiced /d/.

Table 3 shows the duration of the F0 perturbation caused by phonemic /t/ and /d/ for the individual speakers. In the case of JM1, for example, the difference in F0 values /ta/ (L) and /da/ (L) is statistically significant at the first period, but the difference is not significant any more after the first period. That is, the perturbatory effect of /t/ and /d/ disappeared after the first period, and it is 9 ms after the vowel onset. In the case of the high pitch vowels, the difference in F0 values between /ta/ (H) and /da/ (H) is statistically significant until the sixth period. That is, the perturbatory effect of /t/ and /d/ lasted until the sixth period in the high pitch vowels, and it is 41 ms after the vowel onset.

Judging from tables 2 and 3, it is clear that the two primarily acoustic allophones of /d/, represented by short-lag VOT and VOT lead, have an almost identical effect on the F0 of the following vowel. Although /d/ was devoiced 10 out of 20 /da/ (H) tokens (See table 2), the F0 values of the high pitch vowels following phonemic /d/ were significantly lower than those following phonemic /t/ for all of the subjects, except for JM3. This indicates that both fully voiced and devoiced (or partially voiced) /d/ functioned similarly and lowered the F0 of the following vowel. Therefore, it is understood that the low F0 value after a voiced stop does not correlate with the voicing of /d/ in Japanese.

	Low Pitch Vowel		High Pitch Vowel	
Speaker	Period	Duration	Period	Duration
JM1	1	9	6	41
JM2	0	0	6	40
JM3	0	0	0	0
JM4	0	0	6	45
JM5	4	40	4	31
JF1	0	0	14	48
JF2	0	0	6	25
JF3	0	0	1	4
JF4	0	0	8	27
JF5	5	25	10	36
Mean	1	7.4	6.1	29.7

Table 3: Duration of consonantally induced F0 perturbation in Japanese. The numeral in the period columns represent Nth period and the perturbatory effect of the prevocalic consonants lasted until that period. The numeral in the duration columns is the corresponding duration (in ms) of the perturbatory effect.

3. EXPERIMENT: ENGLISH

The participants were 8 native speakers of Australian English, recruited from students and staff at the ANU. They consisted of 4 males and 4 females between 22 and 32 years of age. All of them were born and finished their schooling in Australia (New South Wales or Australian Capital Territory). A reading list containing 2 disyllablic words shown in table 4 was prepared. In the experiment of Japanese, the perturbatory effect was investigated at two different frequency registers. In order to make the English study comparable with the Japanese study, the reading list was read in two different frames shown in table 4. When these words are read in the imperative frame, their pitches are realised as HL, yet when read in the interrogative frame, they are realised as LH. That is, when these words are read in these frames, the first CV segments of the test words make a contrast both in terms of phonemic voicing type of C; [+voice]/[-voice] and the frequency of V; low and high frequencies, and these CV segments are followed by /k/. The elicitation procedures and mensural procedures are the same as Japanese.

Stimuli	Æak« “tucker”	Æaki “duckie”
Frame 1	Say ___ again. (Imperative) L L L	
Frame 2	Say ___ again. (Interrogative) L H H	

Table 4: Two test words and two carrier sentences. L and H stand for low pitch and high pitch, respectively.

3.1. Results and Discussion: English

Similar to Japanese, two types of VOT; short-lag VOT and VOT lead, were observed for /d/ in the English data. Table 5 shows the distribution of these two types of VOT in all of the /da/ segments. The organisation of table 5 is the same as that of table 2.

	Low Pitch		High Pitch	
Speaker	1	2	1	2
EM1	77	46	46	49
EM2	-	16	19	-
EM3	89	89	105	113
EM4	43	40	71	63
EF1	38	-	23	-
EF2	72	49	56	57
EF3	24	23	32	54
EF4	24	-	-	-
Mean	48		57	

Table 5: The distribution of VOT lead and short-lag VOT in English.” -” means that /d/ was fully voiced in production and numerals represent the duration of periodicity absence before the release of /d/ (in ms). 1 stands for first token, and 2 for second token.

Judging from the English results shown in table 5, similar to Japanese, there are no clear restrictions relating to the occurrences of short-lag VOT and VOT lead for /d/. It seems that the selection of short-lag VOT and VOT lead is idiosyncratically decided, or they are in free variation with each other because EM1, EM3, EM4, EF2, and EF3 consistently have short-lag VOT, while EM2, EF1, and EF4 have both short-lag VOT and VOT lead.

Figure 4 shows the F0 perturbation shapes of /ta/ (H) and /da/ (H) uttered by EM1. In figure 4, although /d/ was devoiced for both tokens, the difference in F0 values caused by /t/ and devoiced /d/ is still statistically significant (p ≤ 0.0368) from the first period to the thirteenth period which is the last comparable period.

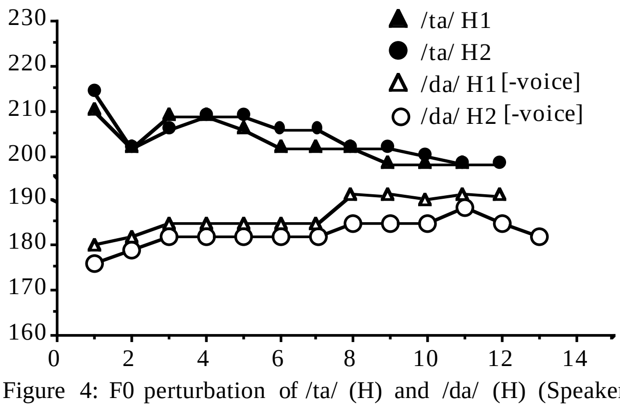

Figure 4: F0 perturbation of /ta/ (H) and /da/ (H) (Speaker EM1).

Figure 5 shows the F0 perturbation shapes of /ta/ (L) and /da/ (L) segments uttered by EM2. In figure 5, although /d/ is fully voiced in the first token of /da/ (L) and partially devoiced in the second token, the perturbation shapes are very similar to each other, and the F0 values of the vowels following /d/ are significantly lower ($p \leq 0.0195$) than those following /t/ from the second period to the seventh period.

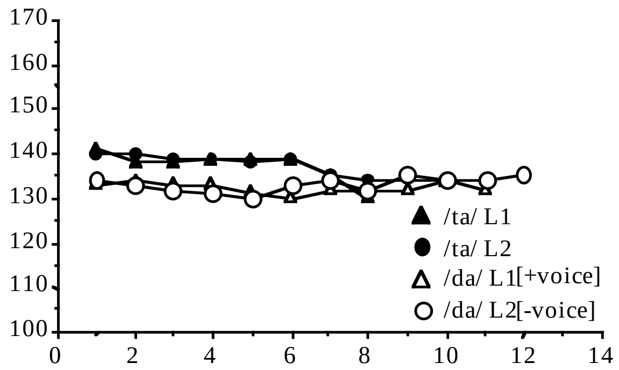

Figure 5: F0 perturbation of /ta/ (L) and /da/ (L) (Speaker EM2).

Table 6 shows the duration of the F0 perturbation caused by phonemic /t/ and /d/ for the individual speakers. The layout of table 6 is the same as that of table 3.

	Low Pitch Vowel		High Pitch Vowel	
Speaker	Period	Duration	Period	Duration
EM1	8	51	12	61
EM2	7	50	11	71
EM3	0	0	9	71
EM4	3	23	7	43
EF1	4	19	12	50
EF2	11	44	14	41
EF3	0	0	9	32
EF4	2	8	6	27
Mean	4.3	24.3	10	49.5

Table 6: Duration of consonantally induced F0 perturbation in Japanese. The numeral in the period columns represent Nth period and the perturbatory effect of the prevocalic consonants lasted until that period. The numeral in the duration columns is the corresponding duration (in ms) of the perturbatory effect.

Judging from tables 5 and 6, similar to the Japanese data, it is clear that the two primarily acoustic allophones of /d/, represented by short-lag VOT and VOT lead, have an almost indistinguishable effect on the F0 of the following vowel. Although /d/ was devoiced 12 out of 16 /da/ (H) tokens, the F0 values of the high pitch vowels following phonemic /d/ were significantly lower than those following phonemic /t/ for all of the subjects. This indicates that both fully voiced and devoiced (or partially voiced) /d/ behave similarly in both occurring together with.

EM 3 has the longest absence of periodicity (105 and 113 ms) of all before the release of /d/ of /da/ (H), and the duration of the perturbatory effect of /t/ and /d/ is 71 ms between /ta/ (H) and /da/ (H). EM4 has a shorter absence of periodicity (71 and 63 ms) before the release of /d/ of /da/ (H) than EM3, and the duration of the perturbatory effect of /t/ and /d / is 43 ms between /ta/ (H) and /da/ (H), and the duration is shorter than that of EM3. On the other hand, EF1 has a shorter absence of periodicity (46 and 49 ms) before the release of /d/ of /da/ (H) than both EM3 and EM4, however, the duration of the perturbatory effect of /t/ and /d/ is 61 ms between /ta/ (H) and /da/ (H), and the duration is longer than that of EM4, but shorter than that of EM3. This indicates, therefore, that the duration of the perturbatory effect of /t/ and /d/ does not correlate with the duration of the absence of periodicity. Therefore, it is understood on the whole that the F0 value after a voiced stop is not correlated with the voicing of /d/ in English.

4. CONCLUSION

It has been shown that both phonetically fully voiced and devoiced (or partially devoiced) allophones of /d/ have a indistinguishable perturbatory effect on the following vowel. This indicates that the presence or absence of periodicity during the closure is not correlated with the F0 perturbation contour of the following vowel. A similar finding is reported in the Chinese dialect of Wenzhou (Rose; personal communication). The F0 perturbation may be caused by some aspects in the production of the preceding stop which are not necessarily manifested in actual vocal cord vibration, such as vocal cord tension. The absence of periodicity during /d/ hold is actually caused by transglottal pressure equalisation not by the lack of vocal cord tension. That is to say, the tension of the vocal cords still exists even if there is no periodicty during the closure. If this is true, the F0 perturbation may still be considered as an intrinsic effect. However, this result still implies that there is a possibility that speakers may deliberately control the F0 of the following vowel (i.e. for the sake of perceptual significance) as an additional cue to the phonemic difference between voiceless and voiced stops in their language. If this is true, although the F0 perturbatory effect of stop consonants has been considered to be an intrinsic event, it may be considered to be an extrinsic event.

5. REFERENCES

1. Ishihara, S. Consonantally induced F0 perturbation in Japanese, Thai, and English. Unpublished MA thesis. The Australian National University, 1996.

2. Ladgefoged, P. “The nature of natural phonetic theories,” Georgetown University Monographs and Linguistics 18: 1965.

3. Ladgefoged, P. “Linguistic phonetics,” UCLA Working Papers in Phonetics, 6: 1967.

4. Lisker, L. and Abramson, A.S. “A cross-language study of voicing in initial stop: acoustical measurements,” Word 20: 384—422, 1964.

5. Tatham, A.A. “Classifying allophones,” Language and Speech 14: 140—145, 1988.

6. Wang, W. S-Y. and Fillmore, C.J. “Intrinsic cues and consonant perception,” Journal of Speech and Hearing Research, 4: 130, 1961.