

# Dialectal Characteristics of Osaka and Tokyo Japanese: Analyses of Phonologically Identical Words

Kanae Amino, Takayuki Arai

Department of Electrical and Electronics Engineering, Sophia University, Tokyo, Japan

amino-k@sophia.ac.jp, arai@sophia.ac.jp

## Abstract

This study investigates the characteristics of the two major dialects of Japanese: Osaka and Tokyo dialects. We recorded the utterances of the speakers of both dialects, and analysed the differences that appear in the accentuation of the words at the phonetic-acoustic level. The Japanese words that are phonologically identical in both dialects were used as the analysis target. The results showed that the pitch patterns contained the dialect-dependent features of Osaka Japanese. Furthermore, these patterns could not be fully mimicked by speakers of Tokyo Japanese. These results show that there is a phonetics-phonology gap in the dialectal differences, and that we may exploit this gap for forensic purposes.

**Index Terms:** pitch patterns, dialectal characteristics, phonetics-phonology gap, forensic phonetics

## 1. Introduction

Speech sounds convey various kinds of information including linguistic phonological information, emotional information, and information about the speaker. Speaker information can be classified into two categories: inherited and learned [1]. The former derives from the speaker's physiological and anatomical characteristics, and the latter derives from the knowledge that the speaker has learned in the process of language acquisition.

One aspect of the learned speaker individualities is the social background of the speaker. It includes social and regional dialects, habits in the articulatory gestures, speech rate and speaking style, and so forth [2-4]. They can be bundled together as "speaking habits" in a broad sense, and may carry linguistic information as well. For example, dialectal features carry linguistic meaning, and articulatory habits such as lisp appear on particular phonemes. On the other hand, speech rate do not necessarily accompany linguistic information.

Research on learned speaker individualities, especially the dialectal characteristics of speech, is important when it comes to social and applied linguistics, linguistic morphology, and second language education. It also contributes to the research on language change and variation, and forensic phonetics. In forensics, criminals often attempt to hide their identities and disguise their voices by intentionally mimicking another dialect than their own or pretending to be some else, and it is difficult for even the trained investigators to tell if there is a voice disguise in the recorded speech materials [5]. Dialectal characteristics can also be used in criminal profiling. These characteristics may appear on the acoustic features such as fundamental frequency (F0), articulatory gestures, formant structures of the vowels, and so on [6].

In this study, we will show the dialectal characteristics of Osaka (spoken in Kansai district including Osaka, Kyoto, Hyogo, and so on) and Tokyo (spoken in Kanto district including Tokyo, Chiba, Saitama and Kanagawa) Japanese by

the analyses of phonologically identical words. Especially, we focus on the pitch patterns. The following section describes the Japanese word accent and the accentuation systems of the two dialects.

## 2. Accentuation in Osaka and Tokyo Japanese

In Japanese, regional dialects appear in the lexicon, syntactic and morphological structures, phoneme inventory, and other phonological properties [7-8]. One of the most varying properties as for the regional dialects is the accentuation. Japanese is a pitch-accent language, and each mora in a word is associated with a specific pitch, high or low. In most dialects the pitch patterns distinguish words.

Examples of the word accents in two major dialects: Osaka and Tokyo Japanese are shown in (1a) and (1b), respectively. Tokyo Japanese has two-way distinction while Osaka Japanese has three-way distinction for the identical segment sequence /haçi/. In Tokyo Japanese, not only the accent patterns but also the location of the accent is important. The accent locations, or the pitch falls, are indicated by "\*" in (1b), although some words like the last one in (1b) do not have the pitch falls.

/haçi/	"bridge"	(high-low)	(1a)
/haçi/	"chopsticks"	(low-high)	
/haçi/	"edge"	(high-high)	
/haçi*/	"bridge"	(low-high)	(1b)
/hâçi/	"chopsticks"	(high-low)	
/haçi/	"edge"	(low-high)	

In these two dialects, some words have phonologically identical accent patterns as shown in (2). The accent pattern in the second word, high-low-low-low, is called "Type 1 accent," as the first mora is accented [9].

/murasaki/	"purple"	(low-high-low-low)	(2)
/akat̪an/	"baby"	(high-low-low-low)	

## 3. Experiment

### 3.1. Speech materials

In this experiment, we used the words with Type 1 accent that are identical at the phonological level in both dialects and compared the acoustic-phonetic differences between them. In addition, we recorded the utterances where the speakers of Tokyo Japanese disguised themselves and mimicked the utterances of Osaka Japanese. These speech data were also analysed and we found that the speakers of Tokyo dialect could not copy the dialectal characteristics of Osaka Japanese.

Nine speakers participated in the experiment. A list of the speakers is shown in Table 1. The selection of the speakers was based on the Japanese Dialect Map provided in [10]. All of the speakers in this experiment were in their 20s or 30s, and they have lived in these Kansai or Kanto districts for more than 18 years.

The recordings were conducted in the soundproof room. All the utterances were recorded onto digital audiotapes at the sampling frequency of 48 kHz with 16 bit resolution.

The speakers uttered the words shown in Table 2 for three times, by answering the questions whose answers would be the target words [11]. So these utterances were quasi-spontaneous. Non-Type-1 filler words were mixed for camouflage. Read speech for all words were also recorded.

With Tokyo speakers, recordings of the spontaneous utterances of the words in Table 2 were carried out first, followed by those of the mimicking speech. They listened to each of the words recorded by Osaka speakers twice, then repeated them three times. The presented speech was Osaka speaker's read speech. They were instructed to copy the utterances of Osaka speakers.

### 3.2. Analysis of the pitch patterns

Before the analysis, the experimenter confirmed that all the speech samples were uttered with consistent accent patterns within one dialect. The analysis targets were the four of the five four-moraic words with Type 1 accent shown in Table 2. The word /o:kami/ was omitted from the analysis, since the boundary between the first and second morae was difficult to determine. Two tokens for each word were analysed. The frequencies for each cycle in the speech waveforms were calculated by using the identical methods as introduced by Sugito [12]. An example of the analysis is shown in Figure 1. The experimenter measured the pitch period on the basis of the zero-crossings visually and manually. All the analyses were conducted by using a computer software Praat [13].

Then the maximal F0 values for each mora were found, and we compared them among the first three morae. In order to care for the speakers' sex differences in F0, we calculated the ratios of the F0 peaks between the adjacent morae, and between the first and the third morae.

## 4. Results

The pitch (F0) contours for the utterances by the speakers of both dialects were drawn and comparison was made between

Osaka spontaneous speech and Tokyo spontaneous speech as well as Osaka read speech and mimicking speech by Tokyo speakers in Figures 2 and 3, respectively. The word /o:kami/ was omitted from the analysis, because the moraic boundary was difficult to determine. In spontaneous speech of the two dialects, which is shown in Figure 2, there was a tendency that Osaka speakers have sharper F0 decrease than Tokyo speakers. In Figure 3, we can see that this tendency is still retained. The utterances of Osaka speakers accompanied the sharp F0 decrease from the first to the third morae. On the other hand, Tokyo speakers could not copy it.

In order to show these differences in a more quantitative way, the ratios of the first-second, second-third, and first-third morae were calculated and the results of the spontaneous speech and of the read or mimicking speech are shown in Tables 3 and 4, respectively.

Table 1. List of speakers.

ID	Sex	Age	Place of birth
O1	Female	25	Kyoto
O2	Male	34	Hyogo
T1	Female	26	Tokyo
T2	Female	27	Tokyo
T3	Female	25	Tokyo
T4	Female	24	Tokyo
T5	Female	23	Tokyo
T6	Female	25	Tokyo
T7	Male	25	Saitama

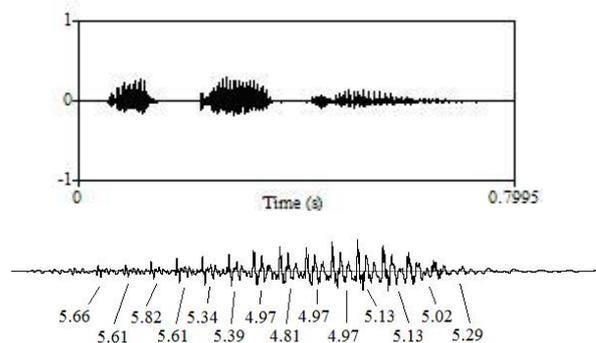


Fig.1. An example of the period measurement: the waveform of the word /akatʃan/ "Baby" uttered by Speaker O2 (upper), and the analysis on the first mora of this word (lower).

Table 2. Recorded speech materials: "H" stands for the high pitch and "L" the low pitch.

Target words	Accent patterns	
	Osaka	Tokyo
/akatʃan/ "Baby"	HLLL	HLLL
/daiʒin/ "Minister"	HLLL	HLLL
/ɸuʒisan/ "Mt. Fuji"	HLLL	HLLL
/tamaɕi/ "Soul"	HLLL	HLLL
/o:kami/ "Wolf"	HLLL	HLLL
Filler words	Osaka	Tokyo
/katamari/ "Chunk"	HHHH	LHHH
/tomodatʃi/ "Friend"	HHHH	LHHH
/kaminari/ "Thunder"	HHHL	LHHL
/murasaki/ "Purple"	LHLL	LHLL
/benjogoci/ "Lawyer"	LLHL	LHHL

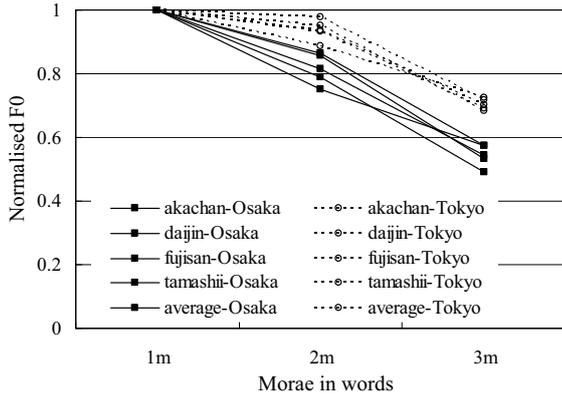


Fig.2. The pitch contours of the four-moraic words uttered spontaneously by speakers of Osaka and Tokyo Japanese. The peak F0s for the first three morae are plotted, where the values are normalised by the first one.

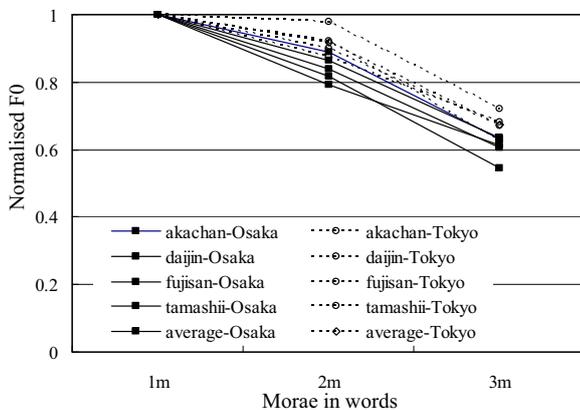


Fig.3. The pitch contours of the four-moraic words in Osaka read speech and in Tokyo mimicking speech. The peak F0s for the first three morae are plotted in the same manner as in Figure 2.

In spontaneous speech, the decrease of the F0 peaks from the first to the second morae was significantly different between Osaka and Tokyo Japanese in all words in a  $t$ -test ( $p < .01$ ). Also, the first-third pair showed a significant difference between the two dialects, but not for all words ( $p < .01$ ).

In the comparison between Osaka read speech and Tokyo mimicking speech, shown in Table 4, we can see that the differences between the two dialects as for the degree of the F0 decrease became smaller. In a  $t$ -test, the difference between the two dialects was significant only for the F0 ratios of the first and second morae ( $p < .05$ ).

## 5. Discussion

In all F0 ratio analyses in this study, inter-speaker differences, that is, intra-dialectal differences were not significant in ANOVA. Therefore, we can say that the great decrease of the peak F0s, especially from the first to the second morae, in four-moraic Type 1 words is a dialect-specific feature of Osaka Japanese, and this feature cannot be mimicked by the speakers

of another dialect, in this study, by the speakers of Tokyo Japanese.

In Japanese linguistics and dialectology, research on Osaka Japanese is important in the sense that it inherits the linguistic structures of the Old Japanese [14]. Some studies have focused on the acoustical differences between Osaka and Tokyo Japanese [12, 15-16]. Sugito [15] reports that Osaka Japanese has less “Ososagari” phenomena than Tokyo Japanese. The “Ososagari” is the phenomenon where the beginning of the F0 peak for a pitch accent occurs after the end of the associated mora [15]. Moriyama et al. [17] explains this phenomenon as “the delay of the laryngeal movement” in generating the pitch. Kori [16] investigated the utterances of Osaka and Tokyo Japanese, and pointed out that the positions of the F0 peaks in each mora of the words with Type 1 accent differed significantly between the two dialects. Sugito [18] found that there is a dialectal difference in the laryngeal controls. In the utterances of Osaka Japanese we often encounter voicing of the phonologically voiceless consonants, but this rarely occurs in Tokyo Japanese. She examined the laryngeal movements by using electromyography, and found that the movements of the sternohyoid muscle can be seen only in the speakers of Osaka Japanese. Fujimoto [19] investigated vowel devoicing in Osaka and Tokyo Japanese by using electroglottography. She spotted the frequency differences of vowel devoicing between the two dialects, that is, Osaka Japanese has much less vowel devoicing than Tokyo Japanese, and concluded that approaching movements of the vocal folds were observed in Osaka speakers even in the devoicing environments.

## 6. Conclusion

In summary, the laryngeal controls and the movements of the certain muscles differ between Osaka and Tokyo speakers, and this brings out the acoustic-phonetic differences of the two dialects. This is why Tokyo speakers could not completely mimic the utterances by Osaka speakers. They did not notice the differences in the laryngeal controls, or even if they did, they could not mimic them. On the other hand, Osaka speakers may not be aware of this dialectal feature. The dialectal characteristics in the muscle movements may be achieved involuntarily.

Unconscious and involuntarily habits in the speech production can be strong clues in forensic speaker identification, because they are often uncontrollable by the speakers [20]. Voice disguise and dialect mimicry are among the notorious problems in forensics [21]; therefore, it is important to know the segmental and suprasegmental features of a dialect comprehensively, and to understand these features in terms of the relationship between phonetic and phonological levels. Nowadays, some people speak other dialects of Japanese quite fluently without making any errors in accentuation. However, the results of this study showed a possibility that even these people can be recognised by using phonetic-acoustic differences among the dialects in question.

In future, we will further investigate the characteristics of other Japanese dialects and other accent types. We should also increase the number of the speakers in order to validate the data. We also continue looking for dialectal features that appear in other aspects of speech. The methodology of the analyses should be refined and made to be more efficient and effective. It is also required to make clear the influence on the perception of the dialectal differences. Further investigations will be reported in near future.

Table 3. Results of the F0 ratio analysis of the spontaneous speech by Osaka and Tokyo speakers: ratios of the peak F0s among the first three morae are indicated.

	2 <sup>nd</sup> / 1 <sup>st</sup> Morae		3 <sup>rd</sup> / 2 <sup>nd</sup> Morae		3 <sup>rd</sup> / 1 <sup>st</sup> Morae	
	Osaka	Tokyo	Osaka	Tokyo	Osaka	Tokyo
/akatʃan/	0.85	0.98	0.63	0.73	0.53	0.72
/daiʒin/	0.86	0.95	0.66	0.71	0.57	0.68
/ɸuʒisan/	0.75	0.88	0.76	0.82	0.57	0.72
/tamaçi/	0.79	0.93	0.62	0.74	0.49	0.69
Average	0.82	0.94	0.67	0.75	0.54	0.70

Table 4. Results of the F0 ratio analysis of the Osaka read speech and Tokyo mimicking speech: ratios of the peak F0s among the first three morae are indicated.

	2 <sup>nd</sup> / 1 <sup>st</sup> Morae		3 <sup>rd</sup> / 2 <sup>nd</sup> Morae		3 <sup>rd</sup> / 1 <sup>st</sup> Morae	
	Osaka	Tokyo	Osaka	Tokyo	Osaka	Tokyo
/akatʃan/	0.89	0.98	0.71	0.74	0.63	0.72
/daiʒin/	0.86	0.90	0.74	0.74	0.64	0.67
/ɸuʒisan/	0.79	0.87	0.78	0.78	0.61	0.68
/tamaçi/	0.82	0.92	0.67	0.68	0.54	0.62
Average	0.84	0.92	0.72	0.73	0.61	0.67

## 7. Acknowledgments

This research was partly supported by a Grant-in-Aid for JSPS Fellows 17-6901 and by Sophia University Open Research Centre from MEXT.

## 8. References

- [1] Furui, S., Acoustic and Speech Engineering (Onkyo Onsei Kougaku), Kindai-Kagakusha Publishing Company, Tokyo, 1992.
- [2] Niimi, Y., Speech Recognition (Onsei Ninshiki), Sakai, T. [Ed.], Kyoritsu-Shuppan Publishing Company, Tokyo, 1979.
- [3] Ladefoged, P., and Broadbent, D., "Information conveyed by vowels," J. Acoust. Soc. Am., 29:98-104, 1957.
- [4] Traunmueller, H., "Modulation and demodulation in production, perception, and imitation of speech and bodily gestures," Proc. FONETIK, 98:40-43, 1998.
- [5] Sjoestroem, M., Eriksson, E., Zetterholm, E., and Sullivan, K., "A switch of dialect as disguise," Lund Univ. Linguistic and Phonetics Working Papers, 52:113-116, 2006.
- [6] Ladefoged, P., A Course in Phonetics, 4<sup>th</sup> Edition, Heinle and Heinle, Boston, 2001.
- [7] Tsujimura, N., The Handbook of Japanese Linguistics, Blackwell, Oxford, 1999.
- [8] Kubozono, H., Mihara, K., Sugimoto, T., Kageyama, T., Takami, K., Nishimura, H., and Nishimitsu, Y., Introduction to English Linguistics: in Comparison to Japanese (Nichiei Taisho niyoru Eigogaku Gairon), Kuroshio Shuppan Publishing Company, Tokyo, 1999.
- [9] Matsumura, A. [Ed.], Daijirin 2<sup>nd</sup> Edition, Sanseido: Tokyo, 1995.
- [10] NHK Broadcasting Culture Research Institute [Ed.], Dictionary of Japanese Accents (Accent Jiten), NHK Broadcasting Culture Research Institute, Tokyo, 1998.
- [11] Chambers, J.K., and Trudgill, P., Dialectology, Cambridge Textbooks in Linguistics, 2<sup>nd</sup> Edition, Cambridge University Press, Cambridge, 1998.
- [12] Sugito, M., "Investigation of Osaka and Tokyo accents by the measurements of the pitch movements (Oncho doutai sokutei niyoru Osaka, Tokyo accent no ichikousatsu)," Studies in the Japanese Language (Kokugogaku), 79:71-89, 1969.
- [13] Boersma, P. and Weenink, D., 'Praat: doing phonetics by computer' (Computer program), URL: <http://www.praat.org/>, 2008.
- [14] Kinsui, S., "On the accent of special morae in Osaka Japanese (Osaka hougen no tokushu haku accent ni tsuite)," in Grammar and Speech II (Bunpou to Onsei II), Institute for Speech Communication [Ed.], 73-199, Kuroshio Shuppan Publishing Company, Tokyo, 1999.
- [15] Sugito, M., Research on the Japanese Accent (Nihongo Accent no Kenkyu), Sanseido, Tokyo, 1982.
- [16] Kori, S. "Acoustic correlates of typical Tokyo and Osaka pronunciation," J. Phon. Soc., 8(3):41-56, 2004.
- [17] Moriyama, T., Ogawa, H., and Tenpaku, S., "Analysis of fundamental frequency about the delay of the laryngeal movement," Info. Proc. Soc. Jpn., SIG Notes, 54:61-66, 2000.
- [18] Sugito, M., Mr. Shibata and Mr. Imada (Shibata San to Imada San), Izumi Shoin Publishing Company, Osaka, 1998.
- [19] Fujimoto, M., "Glottal opening pattern in devoiced tokens by an Osaka dialect speaker," J. Phon. Soc., 9(1):50-59, 2005.
- [20] Nolan, F., The Phonetic Basis of Speaker Recognition, Cambridge Studies in Speech Science and Communication, Cambridge, 1983.
- [21] Hollien, H., Forensic Voice Identification, Academic Press, San Diego, 2002.