Aspects of the Pitch-Accent System of a Japanese Dialect: The Case of Ibuki-jima, Kagawa-ken

Thomas Robb

This is a brief report on the accent system of a unique dialect of Japanese, that of Ibuki-jima, Kan onji City, Kagawa Prefecture, Japan, which was the object of study for the author's Ph. D. dissertation. In the mid-1960's this dialect was discovered to maintain the same 5-way distinction for 2-mora nouns that is documented in the Ruijuumyougisho, a dictionary compiled in 1081. As described in Wada (1966b) not only does a 5-way distinction exist, but there is also a difference in the manifestation of the pitches between different age groups. Using the model for Japanese Pitch-Accent proposed in Pierrehumbert & Beckman (1988) sequences of the 3-mora nouns were analyzed to determine the degree and manner that 'catathesis' applies in the dialect. We conclude that catathesis most likely does apply, but unlike Osaka, focus does not cause a lowering of the F_0 for words that start Low. In emulation of Pierrehumbert & Beckman, the dissertation presents an algorithm for generating pitch contours for Ibuki accent classes using a set of parameters including, declination, maximum and minimum F_0 levels, accent peaks as a percentage of available pitch range, focus increment, and catathesis. To demonstrate the validity of the algorithm, a computer program, IBUGEN, is used to produce contours which are matched with pitch extractions for actual utterances.

1. Introduction

This is a brief report on the accent system of a unique dialect of Japanese, that of Ibuki-jima, Kan'onji city,[1] Kagawa Prefecture, Japan, which was the object of study for the author's Ph. D. dissertation, herein referred to as Robb (1992). Shuuko Senno, at that time a graduate student at Kagawa University, originally discovered that this dialect preserved the origi-

[1] The name of the city is pronounced /kaNoNji/ with moraic nasals which is at variance with the standard reading pronunciation, /kaNnoNji/. The apostrophe will be dispensed with henceforth.

nal five classes of two mora nouns which were reported in the *Ruijumyoogisho*, a document of the Late Heian period (1081).[2]

My interest in this dialect stems from a summary of Wada (1967) found in a pre-publication version of Martin (1987). The fact that a dialect possessing the original five-way distinctions for five-mora nouns was still alive in Ibuki-jima appears to be the main point of interest for many scholars. For me, however, the same report presented more puzzling data. First it reported that the children on the island, while maintaining the 5-way distinction, had a different accent pattern for some of the classes, and second, it appeared that for two of the classes, traditionally labelled 2. 2 and 2. 3, the adult's and child's patterns were virtually the opposite. This state of affairs, going against my linguistic intuitions, resulted in this study.

Ibuki-jima, with a population of approximately 2,000, is a small island in the Seto Inland Sea, 8km due west of Kan'onjishi on the Shikoku mainland with which is politically affiliated. It can be reached by a boat which makes five daily round trips taking 40 minutes in each direction.

Linguistically, the Ibuki dialect shares most features of Kagawa dialects, syntactically, morphologically and lexically. Only the accent system and some fairly straight-forward phonological processes, separate it from its neighbors. There are also several hundred archaic words in active use in Ibuki, but these are most likely common in other Inland Sea dialects, as well. See Kubo (1974) for a compilation of these words produced for the Ibukijima Elementary School.

Minora Wada of Kobe University, realizing the significance of Senoo's discovery, made two trips to the island with Senoo and reported the findings in Wada (1966a) and Wada (1966b).

Below are the essentials concerning the 2-mora nouns:

Table 1. [1. 1][3]Summary of data presented in Wada (1966a).

CLASS	WORDS	ADULTS	BOTH	CHILDREN
2. 1	niwa, miti		HH	
2. 2	hasi ('bridge')	HL-L		HF/HH-L
2. 3	yama, tuki	HM/HH-M		HL-L
2. 4	sumi, sora		LH/LL-H	
2. 5	ido, mado	LH/LH-L		LF/LL-F
(H=high; M=mid; L=low; F=falling)				

[2] See Martin (1987: 167) for further details.

[3] The number in square brackets refers to the table or figure number in Robb (1992).

This provocative report raised a number of questions, some of which I attempt to address in Robb (1992):

1) How can the children have a system where the basic pattern of 2. 2 and 2. 3 words is almost the reverse of the adult system? Is it possible for two accent patterns to shift in opposite directions and cross without actually merging?

2) Can the difference in the adult's system and children's system be explained by a simple and coherent rule?

3) What are the rules for the production of the intonational contour for utterances in Ibuki?

Robb (1992) provides a definitive answers to questions 1) and 2) and goes as far as to suggest a mechanism for generating intonational contours for series of noun phrases, but not for complete sentences, which would be an endeavour of much larger scope and complexity.

2. How Can the Adults and Children Be So Different?

This question was largely answered in Robb (1986) which appeared before the dissertation itself. A number of phrases containing 2-mora nouns with similar segmental structure were elicited from 2 adults and four children (hereafter referred to as "JHS" for "junior high school"). Figure 1 shows the patterns that emerged. Three points are plotted for each mora, the start, middle and end points of the vowel (s, m and e, respectively.).

Apart from the variation in the location of peaks for some of the classes, a striking difference between the adult and JHS informants is the onset. Note that in all cases the onset for the adults comes from below, while for the JHS subjects, the onset in fairly level. We shall see below that this across-the-board rule (be it the adults who lower the onset or the JHS who raise it) is one of the factors which causes the pitch for some classes to be perceived as different for the two groups.

Consider the 2. 2 class, for example. Figure 2 illustrates the averaged contours for the adults (solid line) and the JHS informants (dotted line) for the class we shall call "H1" after Uwano (1985).

First note that for both age groups the slope from the peak to the end of the utterance is nearly identical. A glance below at Figure 3 (class 2. 3) reveals that for this class, too, the slopes from the peak onwards are identi-

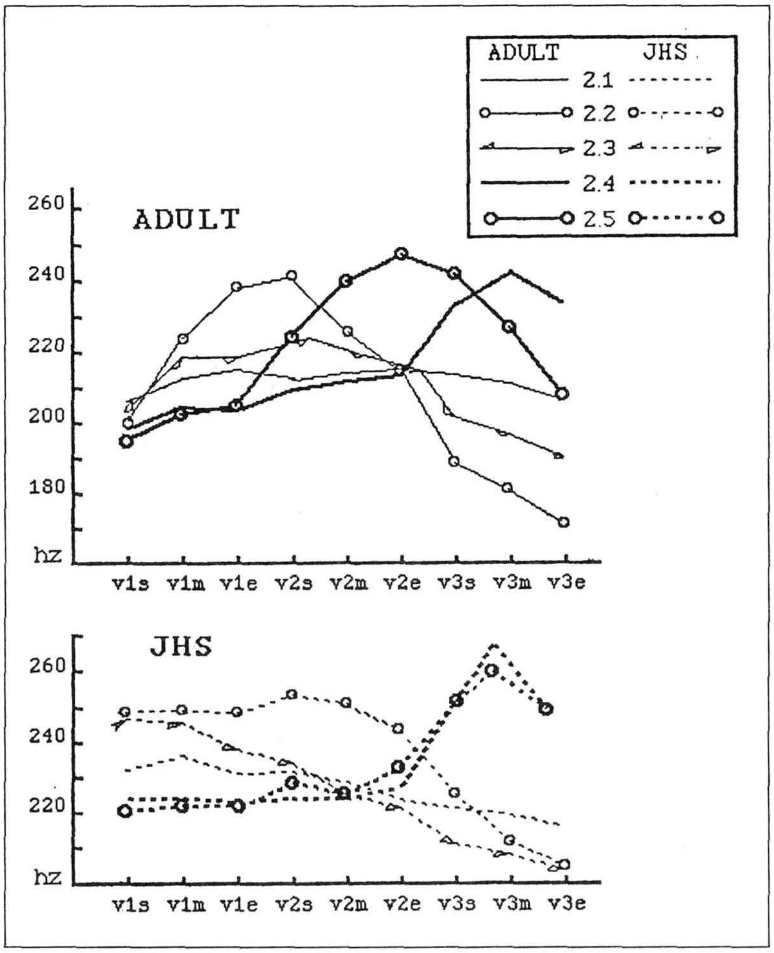

Fig. 1. [2. 6] Two-mora noun contours for Adults and JHS.

cal, although quite distinct from class 2. 2. The two charts are overlaid as Figure 4 for ease of comparison.

Next, returning our attention to class 2. 2, recall that Wada (1966) (Table 1) reported that the JHS pattern sounded High-Falling (HF) without a particle and HH-L with a particle attached. The JHS have delayed the fall so that the drop begins more slowly, dropping more precipitously as time increases. The net effect of this is that the end of mora-2 is still not low enough for the mora to be perceived as Low. See Sugito (1982) for experimental work on the relationship between F_0 change and perception of pitch.

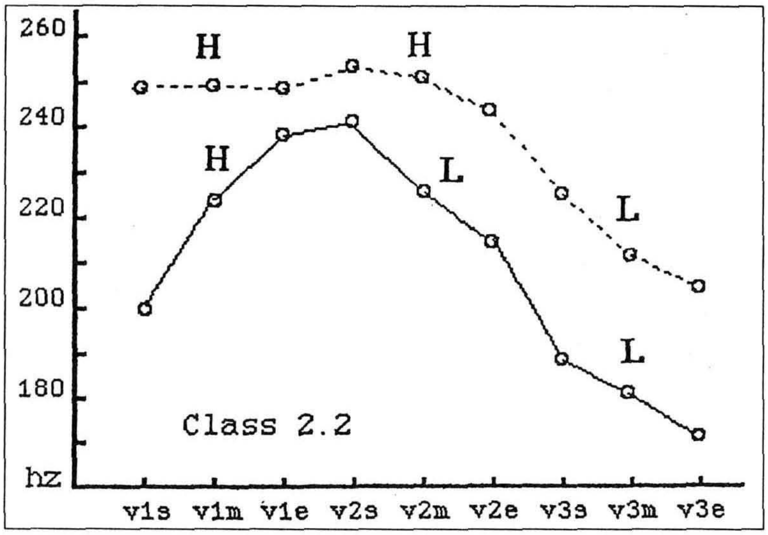

Fig. 2. [2. 7] Adult and JHS contours for noun class 2. 2.

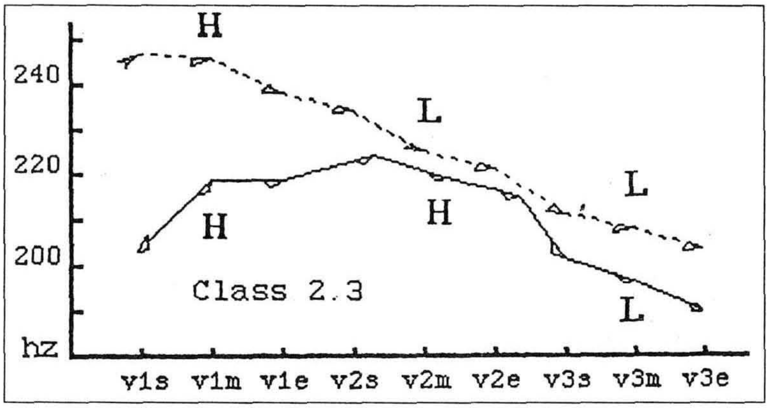

Fig. 3. [2. 8] Adult and JHS contours for noun class 2. 3.

For class 2. 3, one could say that the JHS utterance has no peak since there is fairly smooth downward slope from beginning to end. One could, however, posit a "peak" at a similar location to that in the Adult utterance and then claim that the mechanism which elevates the onset of JHS utterances has caused the peak to disappear. In other words, one could derive the JHS contour from the Adult contour by rotating the initial segment of the adult utterance upward.

Regardless of how this variation has arisen, the net effect of the change

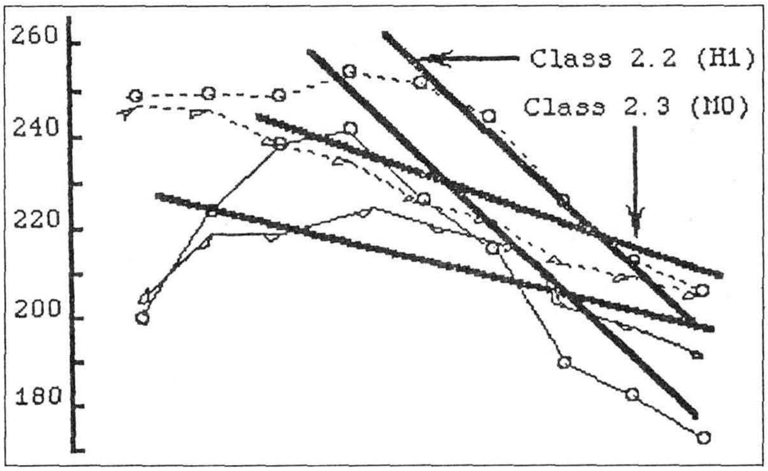

Fig. 4. [2. 9] Adult and JHS contours for noun classes 2. 2 and 2. 3 compared with slopes marked.

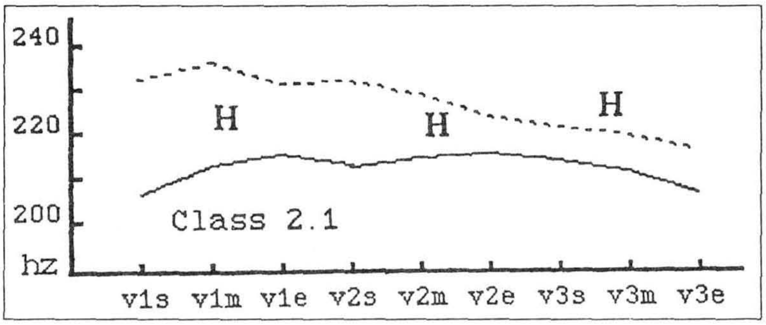

Fig. 5. [2. 10] Adult and JHS contours for noun class 2. 1.

is that the second mora is no longer heard as High. The slope contrasts with both class 2. 1 (H0) which has virtually no slope (Figure 5), and class 2. 2 (H1) which has a rather abrupt one. Thus, class 2. 3 is labeled as High–Mid.

In summary, we can see that the Adult and JHS utterances are not as different as one might be led to believe by looking only at the differences in their manifestations in terms of High and Low pitches. The JHS speakers have regularly raised the onset (relative to the adults) for each class.

Further, this raising combined with a modest shift towards the end of the word in the locus of fall in the H1 and L2 classes results in a major difference in the interpretation of the tones. It would appear that to speakers of this dialect, the *slope* of the phrase is a more important clue to the accent class than is the locus of fall.

3. Three-Mora Nouns

One aspect of three mora nouns which took up considerable space in Robb (1992) was the determination of number of accent classes for 3-mora nouns. Statistical tests and other arguments were advanced to show that there is not sufficient evidence to posit more than the same five classes which exist for two-mora nouns.

Of particular interest in the M0 group, which behaves similarly to the H1 group with the exception that the fall is not as abrupt. M0 words typically exhibit a fall from peak to word-end of approximately 40Hz as opposed to 60Hz for H1 items. The location of the fall in all cases appears to be identical with that for H1, i.e., at the beginning of Mora 2.

Of more interest than the distribution of class types in Ibuki is the role of "catathesis" or "down-step" in a series of two or more phrases, a phenomenon that occurs in the Tokyo dialect catathesis after the HL accent, but only if it has already been preceded by the word focus. Word focus following the HL drop blocks catathesis. Also, it must be the HL of an accented word, not an H followed by a boundary L. Pierrehumbert and Beckman (1988), hereafter "P&B", claim that catathesis occurs in Osaka, and under more liberal conditions than in Tokyo, occurring any time an H happens to conjoin an L.

In order to discover whether catathesis operates in Ibukijima, and, if it does, under what conditions, a set of utterances, called the "Kodomo ga kotori wo koota" set was elicited. See the Appendix for description of the set of utterances.

Figure 6 reproduces one of P&B's figures for the Osaka dialect which clearly shows the role of focus in shaping the pitch contour. Note that in all cases the word in focus has a clearly higher F_0 than its unfocussed counterpart.

In Ibuki, while we can observe a clear reduction of the ceiling for the following word for items with Word-1 focus, very interestingly, items with Word-1 focus do not necessarily have the higher F_0 that would be expected. Figure 7 presents representative forms with Word-1 focus (thin line)

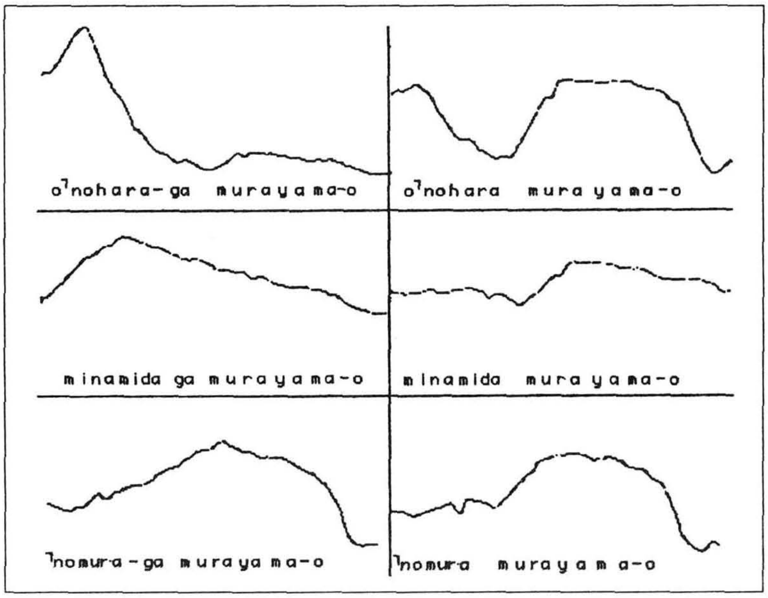

Fig. 6. [5. 2] Osaka F_0 Contours of unaccented high-beginning *Murayama-o* preceded by (top) accented O'nohara (-ga), (middle) unaccented high-beginning *Minamida*(-ga) or (bottom) unaccented low-beginning *Nomura-ga* with narrow focus either (left) on the preceding word or (right) on *Murayama*. (P & B p. 230, Figure 8.10)

and Word-2 focus (heavy line) using representative contours from the data. Our five accent classes are represented with the "W1" line representing 15 utterances with Word-1 focus, and the "W2" line representing 15 utterances with Word-2 focus. The W2-1 figures indicate the difference in the mean values. The bold faced "ALL GRP" values are means of all 5 rows above it, for a total of 15 utterances per word class.

Each line on the Table 2 represents the mean of 15 utterances which all have the same first element (the accent class under examination) followed by 3 instances each of our 5 accent classes.

The values for W2M1 (Word-2 Mora-1) onwards are thus "apples and oranges" to a certain extent. Since, however, each of the 5 classes contains exactly the same Word 2 representation of 5 accent classes times 3 utterances each, we can legitimately hypothesize that the values for each of the Word 2 sets should be the same if there is no influence from the preceding Word-1.

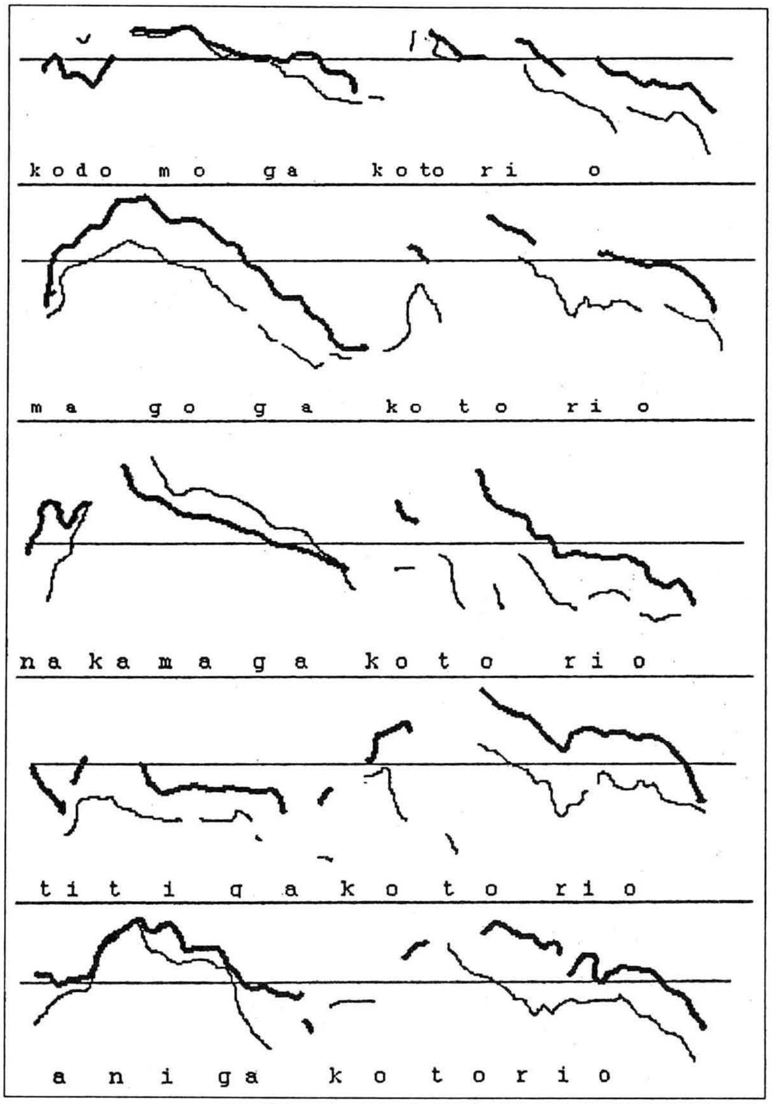

Fig. 7. [5. 3] F_0 contour traces for each Ibuki 3-mora word class followed by high-starting, unaccented *Kotori*-o. Thin line=Word 1 Focus, Thick line=Word 2 focus.

Of immediate interest to us now are the moras which show a significant difference between the two focus groups. In all three accented classes, H1, M0 and L2, there is a significant difference in the moras W2M2 and W2M3 which represent the peaks of the waveform. We can attribute this difference to the operation of catathesis.

Looking at the W2M2 column again, we can see that the two unaccented

Table 2. [5. 1] Comparison of Word- Focus (upper line) vs. Word–2 Focus (lower line) for MKn & FK Combined Data Ordered by Word–1 Class (*=p<0.05, **=p<0.01)

	W1M1	W1M2	W1M3	WIM4	W2M1	W2M2	W2M3	W2M4	W3M1
H0									
W1	201.5	205.3	200.7	192.7	178.9	181.0	170.5	155.2	152.9
W2	206.1	213.1	207.2	200.0	187.2	194.4	181.9	160.8	159.2
W2–1	4.6	7.8	6.5	7.3	8.3	13.4	11.4	5.6	6.3
H1									
W1	179.5	214.3	189.7	162.6	163.4	177.5	168.5	149.6	149.1
W2	193.8	217.5	190.5	161.3	170.5	194.2	186.7	164.4	163.1
W2–2	14.3**	3.2	0.8	–1.3	7.1	16.7*	18.2**	14.8**	14.0*
M0									
W1	186.1	212.8	196.8	176.4	167.9	170.3	163.0	149.7	151.2
W2	193.9	216.9	200.4	182.5	178.8	191.7	177.5	157.6	155.9
W2–1	7.8	4.1	3.6	6.1	10.9*	21.4**	14.5**	7.9	4.7
L0									
W1	177.1	190.3	196.8	197.6	184.6	196.4	181.9	165.6	155.0
W2	179.0	182.6	195.1	198.5	186.0	208.3	191.4	170.4	161.4
W2–1	1.9	–7.7	–1.7	0.9	1.4	11.9	9.5	4.8	6.4
L2									
W1	173.8	207.4	198.6	174.9	167.4	179.0	171.4	154.4	156.3
W2	181.4	217.9	212.2	185.2	183.3	201.0	190.8	168.6	162.2
W2–1	7.6	10.5	13.6**	10.3**	15.9**	22.0**	19.4**	14.2*	5.9

Table 3. [6. 3] Comparison of actual means (boldface) and values generated by IBUGEN for Osaka/Kobe data. Italicized values indicate poor predictions(> 10hz off).

Focus on W2 (Person)							Focus on W3 (Object)						
	W1	Valley	W2	Valley	W3	Verb		W1	Valley	W2	Valley	W3	Verb
aAa	**239**	**168**	**236**	**159**	**168**	**132**	aaA	**248**	**164**	**187**	**155**	**223**	**143**
	233	161	234	151	165	131		246	161	183	129	224	150
aUa	**232**	**169**	**219**	**195**	**206**	**148**	auA	**252**	**161**	**187**	**167**	**231**	**145**
	233	161	217	179	196	151		246	161	174	149	226	150
uAu	**219**	**201**	**240**	**165**	**169**	**156**	uaU	**218**	**195**	**204**	**161**	**203**	**182**
	215	190	229	150	152	144		228	189	209	148	202	183
uUu	**209**	**195**	**218**	**170**	**191**	**165**	uuU	**219**	**195**	**203**	**174**	**213**	**190**
	215	198	213	182	184	165		228	189	197	174	202	183

classes (H0 and L0) show a mean difference of 13.4 and 11.9hz, while the three accented classes (H1, M0 and L2) show differences of 16.7, 21.4 and 22.0, respectively. If we assume that all classes manifest Word–2 focus by raising the height of the peak, we can then determine that focus augments the height by 12-13hz with catathesis depressing the Word-2 heights by another 8-10hz.

Observations such as those above then led to the next logical step: to see if such rules could be systematized to the extent that a computer program could generate appropriate contours with "knowledge" of only the accent class of words and the word in focus. For this, the author developed a program called IBUGEN, some the results for which are presented below in Table 3. Figure 8 shows the IBUGEN-produced contours on the top with the extracted pitch contours from actual utterances from informant MKn on the bottom. Table 3 compares the values of actual utterances (bold-faced) with those generated by IBUGEN. Italicized items indicate those places where the algorithm failed to come within 10hz of matching the actual utterances.

Naturally, since these are individual utterances each with their own idiosyncratic elements, they will not match the generated utterances exactly.

While the exact formulation of the IBUGEN algorithm is beyond the scope of this paper, I will present below a list of the various parameters that were used to generate the utterances in Figure 8. Please see Chapter 6 of Robb (1992) for a complete description.

Declination — 3hz per mora

MAxhz — the maximum pitch level for a speaker's utterance. (250)

MInhz — the minimum pitch level for a speaker's utterance. (120)

Accent Peaks — the percentage of the current range utilized for the peak value of each accent type.

- HP — High-Accented peak (0.90)
- HOP — High-Unaccented peak (0.75)
- MP — Mid-Peak - For MO accents (0.50)

LMin — the base value for Low-starting utterances (0.40)

Catathesis — (0.40) - (The greater the value, the larger the drop)

F1 — Focus Increment - the degree that the peak value is incremented when the word is the point of focus.(0.10)

SH — Start High - position of the onset for a phrase for high-starting phrases. (0.80)

SL — Start Low - position of the onset for a phrase for low-starting phrases. (0.40)

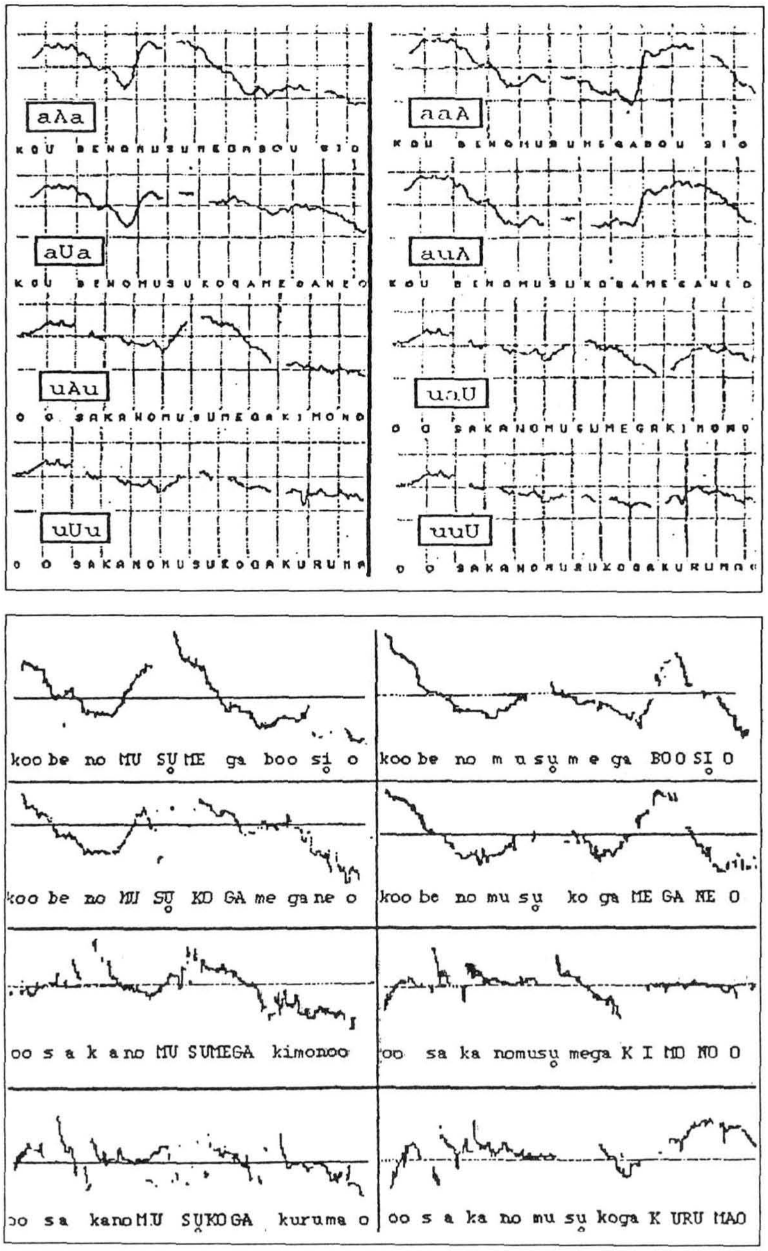

Fig. 8. [6. 7~6. 9] Output of IBUGEN (above) with acutal utterances (below).

4. Conclusion

As illustrated above, a system of parameters with numerical parameters can provide a close fit to actual utterances. Such a system, while much more complicated than the simple traditional descriptive systems which utilize marks to indicate changes in pitch height, is most likely much closer to the truth. The system can eventually result in changes at the phonemic level.

Appendix

Description of Major Data Sets Used in Robb (1992)

"Kodomo ga kotori wo koota" (6×6) set

Nouns representing 6 hypothetical accent classes were substituted in two slots for a total of 36 combinations. The purpose was to gather sufficient data to refute Uwano's claim for an H2 class of 3-mora nouns. Secondly, this set allows a detailed analysis of how the pitch contour is affected by the presence of a following or preceding word of another pitch class. Thirdly, it allows us to investigate the role of focus; i.e., the relative prominence of noun-1 vs. noun-2 under various elicitation conditions.

H0	kodomo ("child")	kotori ("bird")
H1	mago ("grandchild")	megane ("glasses")
H2	itoko ("cousin")	komugi ("wheat")
M0	nakama ("friend")	hasami ("scissors")
L0	titi ("father")	kabura ("turnip")
L2	ani ("older brother")	monaka ("bean cake")

Note that some 2-mora nouns were unavoidably used for the first slot despite the fact that 3-mora nouns would have been more desirable. The set, however, required a noun designating a kind of familial or inter-personal relationship which restricted the field from which lexical items could be drawn. Fortunately, the overall F_0 pattern for these classes is invariant; the contour is merely spread over more morae as necessary. For 3-mora nouns, one point was input for each mora, for a total of four points, including the final particle. For the two mora nouns, since we are 'one mora short,' four points were selected at approximately evenly-spaced intervals in a manner which most faithfully represented the accent contour.

The sentences were elicited in three ways:

1) A straight reading
2) As an answer to the question, XXXX ga nani o kootekita? ("What did XXXX buy?")
3) As an answer to the question, YYYY wo koota no wa dare? ("Who bought YYYY?")

Two complete sets were recorded from informant MKn (F, 83), and one additional set from FK (F, 88). (While a greater number of sets might have been desirable from a statistical standpoint, it was important not to overtax the informants.)

"Oosaka no musuko ga kuruma o koota" set (O/K Focus Set)

The purpose of this set was to gather multiple instances of a few crucial combinations in order to 1) check the effect of focus on the pitch contour and 2) check the interaction of declination and catathesis.

There were four sentences:

Accent Types, Word:	1	2	3
Oosaka no musuko ga kuruma o koota	H0	H0	H0
(My Osaka son bought a car.)			
Oosaka no musume ga kimono o koota	H0	H1	H0
(My Osaka daughter bought a kimono.)			
Koobe no musuko ga megane o koota	H1	H0	H1
(My Kobe son bought some glasses.)			
Koobe no musuko ga megane o koota	H1	H1	H1
(My Kobe daughter bought a hat.)			

Condition 1–Error in the thing bought

I made a false statement which the informant then corrected, for example,

"Oosaka no musuko ga KIMONO o koota yaroo"

("Your Osaka son bought a kimono, right?")

to which the informant replied,

"Oosaka no musuko ga KURUMA o koota"

("My Osaka son bought a CAR.")

There were 12 instances of each sentence in this set.

Condition 2–Error in the person who bought it

This run was similar to that above but the informant replied with the correct person. I asked,

"Kimono wo koota no wa, Osaka no musuko deshoo" ('The person who bought the kimono was your Osaka son, right?') to which the informant replied:

"Oosaka no MUSUME ga kuruma o koota"

("It was my Osaka DAUGHTER who bought the kimono")

The sentences in this set were elicited 5 times.

References

Martin, Samuel E. (1987) *The Japanese Language Through Time*, Yale University Press, New Haven.

Pierrehumbert, Janet B. and Mary E. Beckman (1988) *Japanese Tone Structure*, Linguistic Inquiry Monograph 15, The MIT Press, Cambridge, Massachusetts.

Robb, Thomas N. (1986) 'On the Use of the T-Test for Differentiating Accent Classes: The Case of Ibuki-Jima (Kagawa Pref.),' in *Kokusai Gengo Kagaku Kenkyunsho Shohoo* 7. 2, 196–213, Kyoto Sangyo University, Kyoto.

Robb, Thomas N. (1992) 'Aspects of the Pitch Accent System of Ibukijima,' Doctoral Disseration, University of Hawaii at Manoa, Department of Linguistics (unpublished).

Sato, Eisaku (1985) 'The Structure of Pitch Accent Patterns in the Ibukijima Dialect,' in *Kokugogaku*, March 1985, No. 140, 32–44.

Senoo, Shuuko (1986) 'Kagawa Ibukijima no Akusento (The Accent of Ibukijima, Kagawa Prefecture)' in *Kokugo Kenkyuu* #22, (May, 1966) Kokugakuin Daigaku, 22–3.

Sugito, Miyoko (1982) *Nihon Akusento no Kenkyuu* (Research on Japanese Accent), Sanseido.

Uwano, Zendoo (1985) 'The Accent System of Ibukijima Dialect,' in *The Transactions of the Japan Academy* (Nihon Gakushiin Kiyoo), February 1985, 40. 2, 75–179.

Wada, Minoru (1966a) 'Dai Itiji Akusento no Hakken (The Discovery of the First Stage Accent)' in *Kokugo Kenkyuu* #22, (May, 1966) kokugakuin Daigaku, 24–28.

Wada, Minoru (1966b) 'Nihaku-meishi itsutsu no rui o itsutsu niwakeru akusento (A Dialect Which Makes a 5–Way Distinction among the 5 Classes of 2–Mora Nouns)' in *Gengo Kenkyuu* 50, 128–130.

6–27 Hirakata–Motomachi
Hirakata 573
Japan