

# WORD ACCENT AND RELATED FEATURES IN JAPANESE

—an Acoustic Analysis—

Toshiaki Fukasawa

## 1 Aim

To analyse acoustically the following three points:

- ( i ) What are the acoustic correlates of Japanese accent ?
- ( ii ) When there is more than one Head in a Tune, what are the acoustic features manifesting them, especially the Head other than the first one in a Tune ?
- ( iii ) What are the acoustic features of unvoiced morae ?

## 2 Method

### 2.1 Material

In order to examine the above points, four groups of words, eleven words altogether, were selected according to their segmental similarities and their accentual differences.

Each of the eleven words was put into the frame sentence, which was uttered with Tune 1 : // 1 moo icido/[ ] to iQte kudasai // 'say [ ] again please', literally 'more once [ ] say please'.

The use of a frame sentence for the examination of accent in Japanese is thought to be an important prerequisite for ( i ) achieving a close approximation to natural speech, ( ii ) minimizing extra-

neous variability which is expected to be greater when words are uttered in isolation than in a frame sentence and (iii) avoiding the influence of intonation on the duration of the segment in the utterance final position, since the segment in the utterance final position is known to be lengthened.

The phonological form of the frame sentence// / //was chosen here in order to examine the acoustic features of a Head other than the first one.

This particular segmental structure of the frame sentence was chosen in order to obtain, when the words are put into it, some unvoiced morae to be examined.

The procedures taken for the preparation of the material for recording was as follows:

- ( i ) Five tokens for each word, altogether fifty-five tokens were obtained in random order.
- ( ii ) The scripts were all written in Japanese orthography.

The eleven words selected are reproduced below in phonological transcription with English equivalents indicated.

#### Group 1

- / ikoo / 'since, afterwards'
- / ikoo / 'let's go'
- / ikoo / 'intention'

#### Group 2

- / jakimoki / 'impatiently'
- / kakimoci / 'broken rice-cake'
- / jakimoci / 'jealousy'
- / kakemoci / 'holding additional post'

#### Group 3

- / cikaku / 'neighbourhood'
- / cikaku / 'perception'

Group 4/ huku / 'blow'

/ huku / 'wipe'

**2.2 Recording**

Three tokens were found to be improperly pronounced, and so they were rejected.

Setting for recording

1. Distance from lips to microphone : 32 cm
2. Recording level volume : 7
3. V. U peak intensity reading : -5 dB
4. Recording characteristic : NAB
5. Output from oscillator : 48 mV (1 KHz tone)
6. Peak S. P. L. at microphone :  
     Before recording session ; 44 dB  
     After Recording session ; 44 dB  
     Filter ; LIN

**2.3 Acoustic analysis****2.3.1 Equipment used**

All the acceptable tokens were acoustically analysed using a sound spectrograph.

Both narrow-band and wide-band spectrograms were made, with amplitude displays on the narrow-band spectrograms. The trace was not calibrated.

**2.3.2 Measurement**

Tracings were made of the Fo curve and of the amplitude curve of every token of each set of utterances from the narrow-band

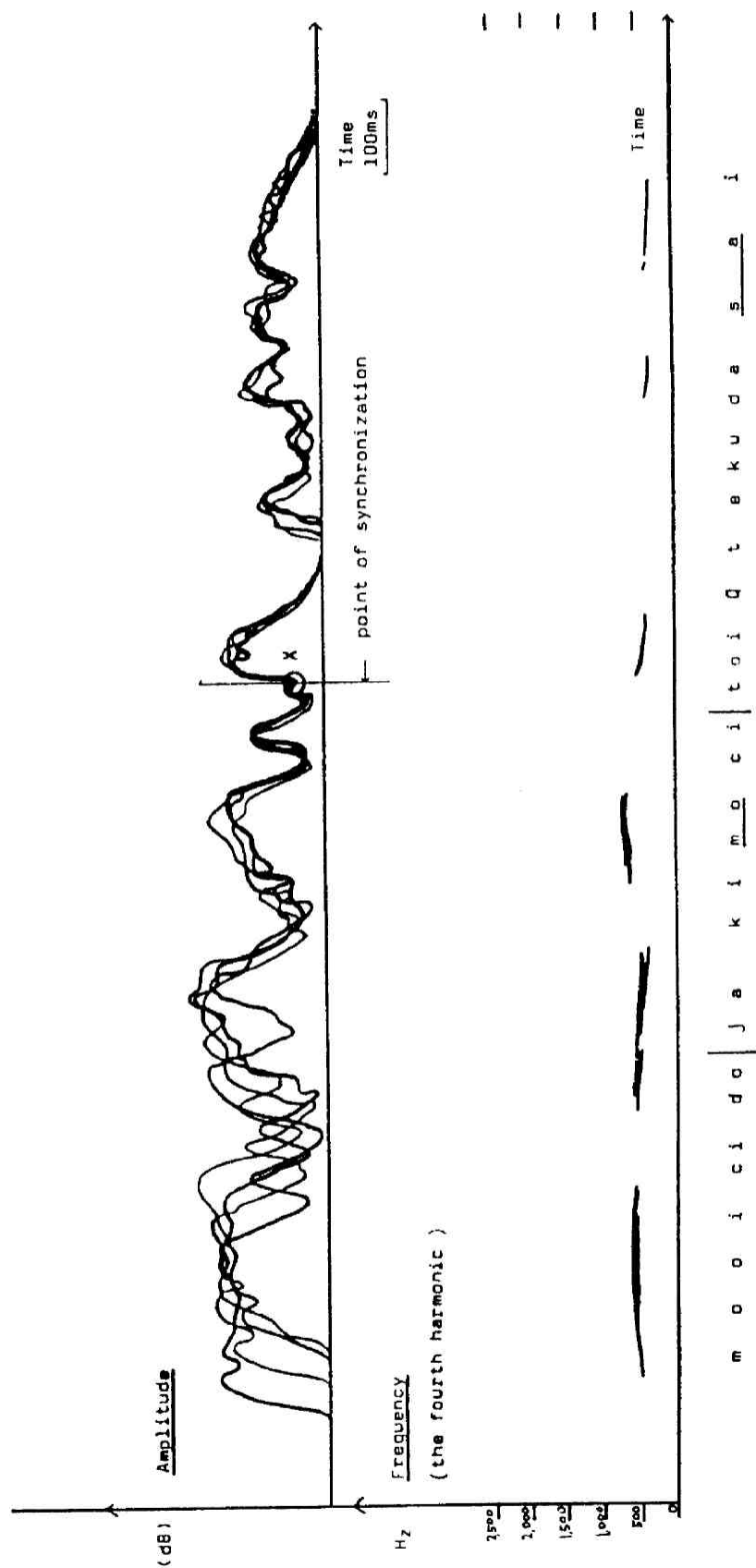


Fig. 1 Sample traces of  $F_0$  and amplitude for /jakimoci/ uttered in a frame sentence to illustrate

how tracings were made

x : Point of synchronization

Target word is enclosed by vertical lines.

This set-up applies to Figs. 2 to 16 which follow.

spectrograms.

Synchronization for the tracings was made at the onset of the vocalic element of [o] in [to], which was observed by the sudden rise of the amplitude trace which took place in every token immediately after the small rise of the trace which corresponded to the release of consonantal element of [t] in [to]. One such example of a trace is given below in Fig. 1, p. 4 to illustrate the point of synchronization.

### **3 Result and discussion**

Figs. 2 to 12 (pp. 6 to 16) show the traces of the Fo and the amplitude patterns of eleven sets of utterances which contain the eleven target words.

Each set shows certain variations in the Fo and the amplitude patterns, but there is no token that is noticeably different within each set, so that it can be said that the utterances in each set are homogeneous.

In order to compare the overall Fo and amplitude patterns of the eleven sets of utterances, an average line for each set of traces was estimated visually and drawn (Figs. 13 to 16, pp. 17 to 20): Figure 13 compares three sets of utterances which contain the words of Group 1. Likewise, Figures 14, 15 and 16 compare the average Fo and amplitude traces for the words of Groups 2, 3 and 4 uttered in a frame sentence respectively.

#### **3.1 Results on Accent**

Accent was defined phonologically in terms of the location of the “marked” mora in a lexical item.

#### Fundamental frequency

Figures 13 to 16 clearly show that the Fo traces for the eleven

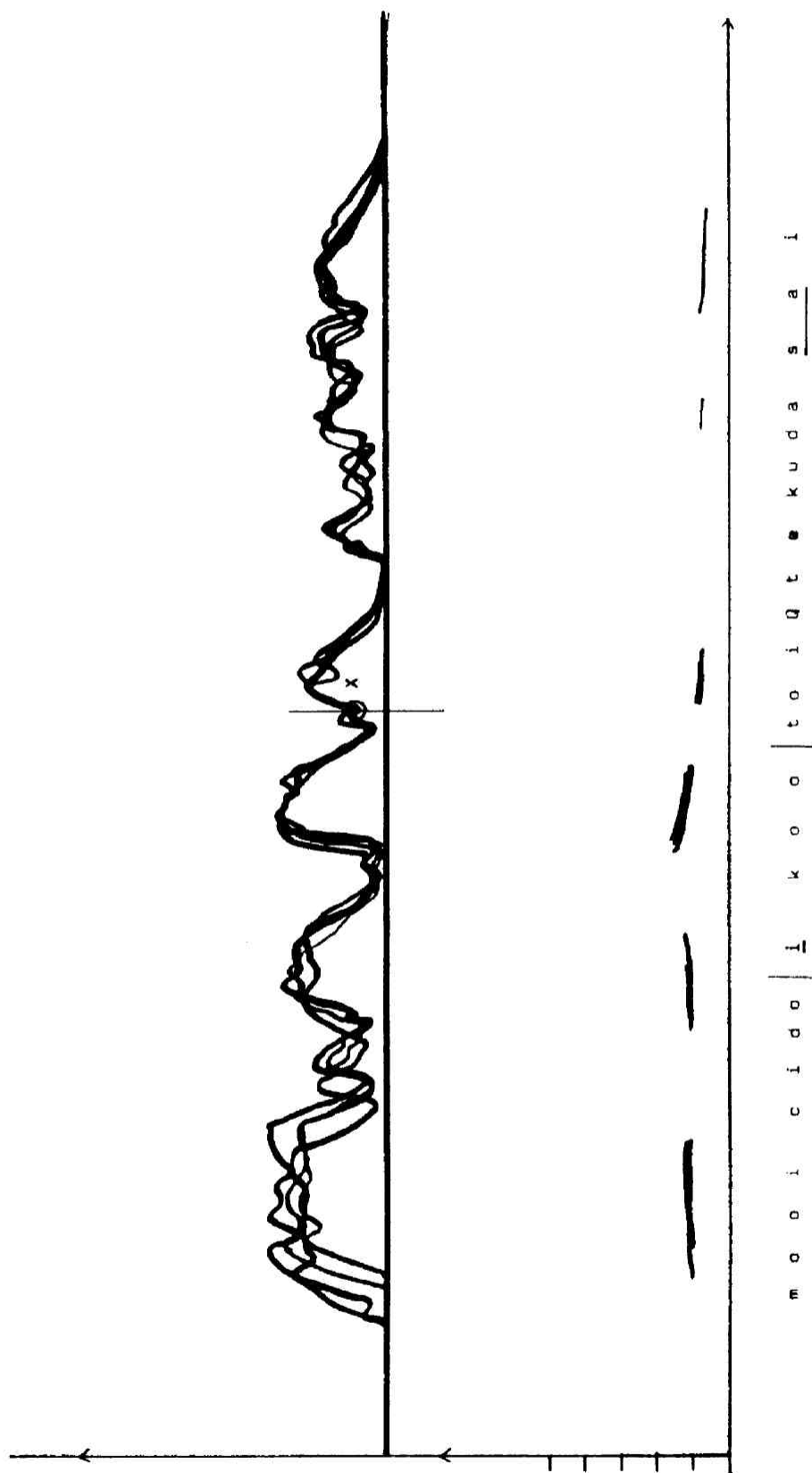
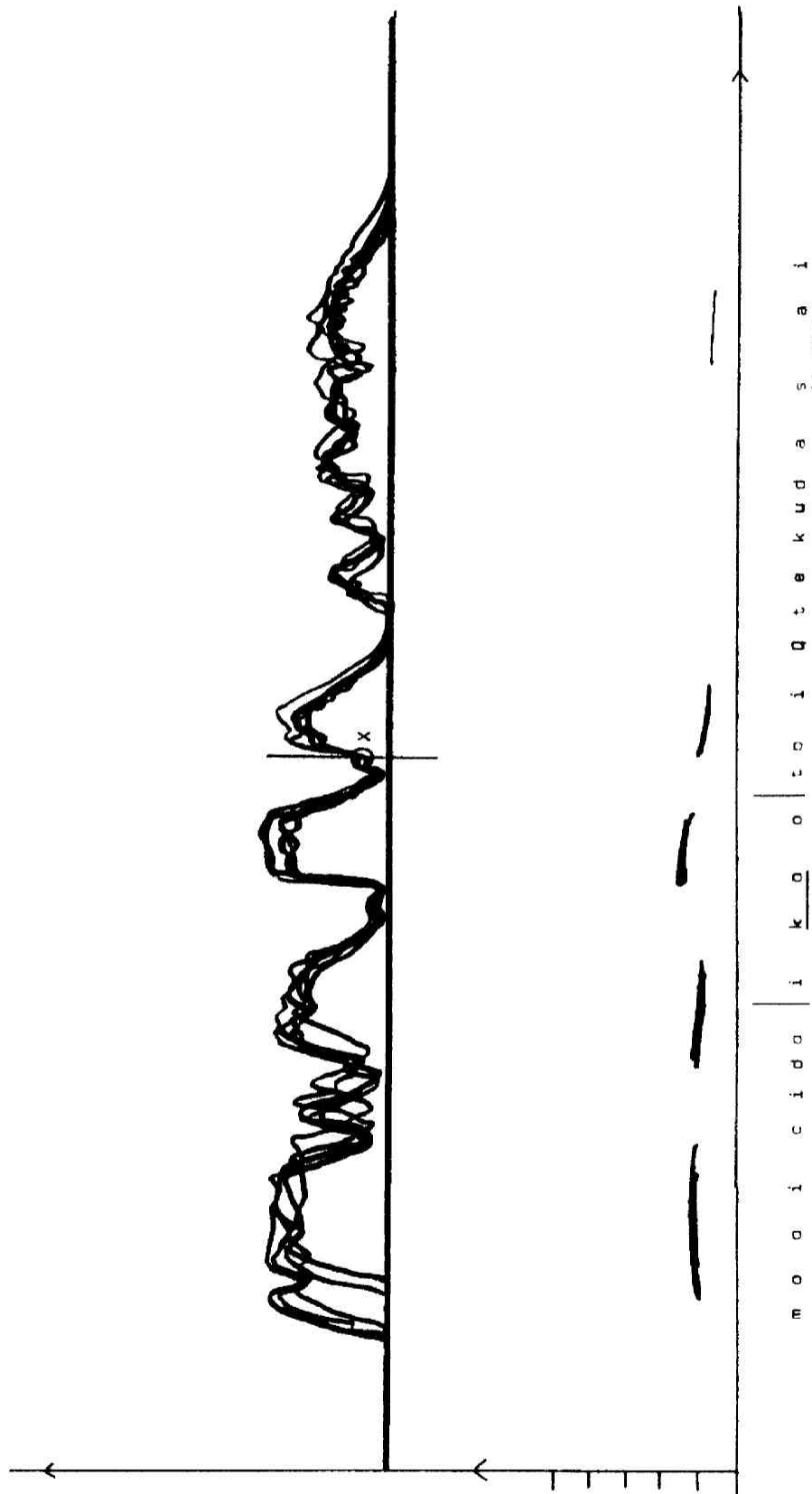


Fig. 2 Traces of F0 and amplitude for /ikoo/ uttered in a frame sentence

Four tokens are present on the Figure.



**Fig. 3** Traces of F<sub>0</sub> and amplitude for /ikoo/ uttered in a frame sentence  
Five tokens are present on the Figure.

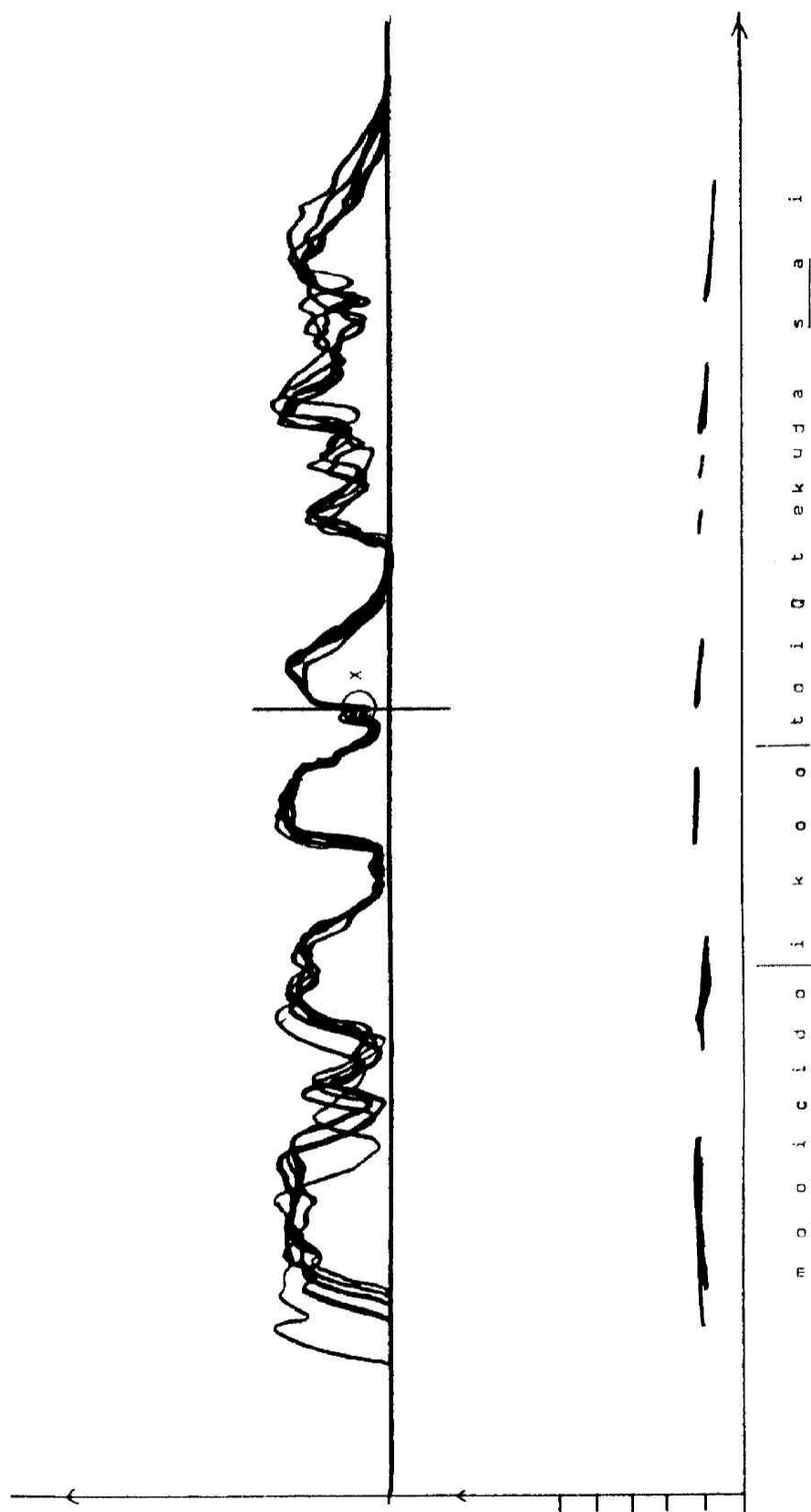


Fig. 4 Traces of F0 and amplitude for /ikoo / uttered in a frame sentence

Five tokens are present on the Figure.



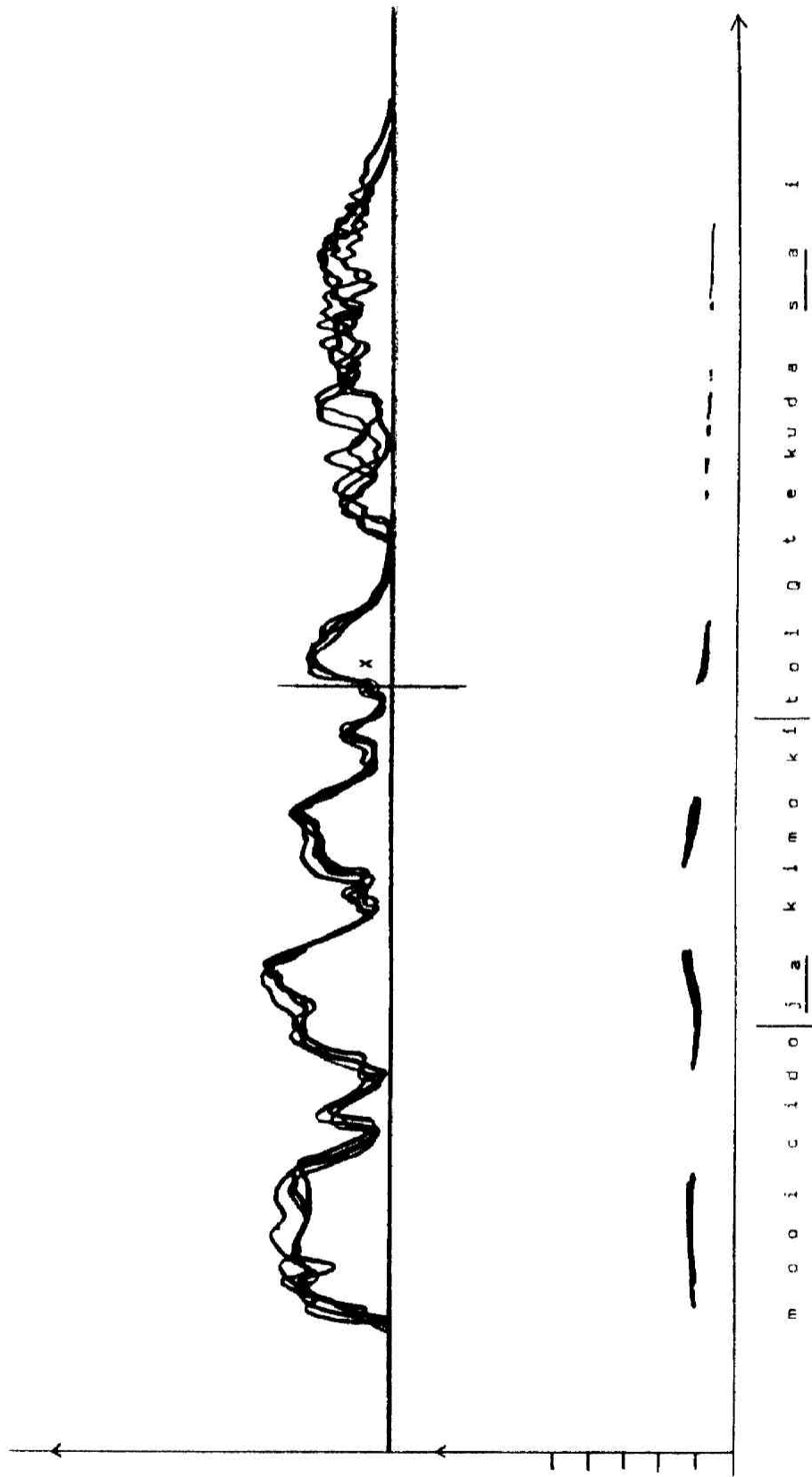


Fig. 5 Traces of  $F_0$  and amplitude for /jakimoki/ uttered in a frame sentence

Four tokens are present on the Figure.

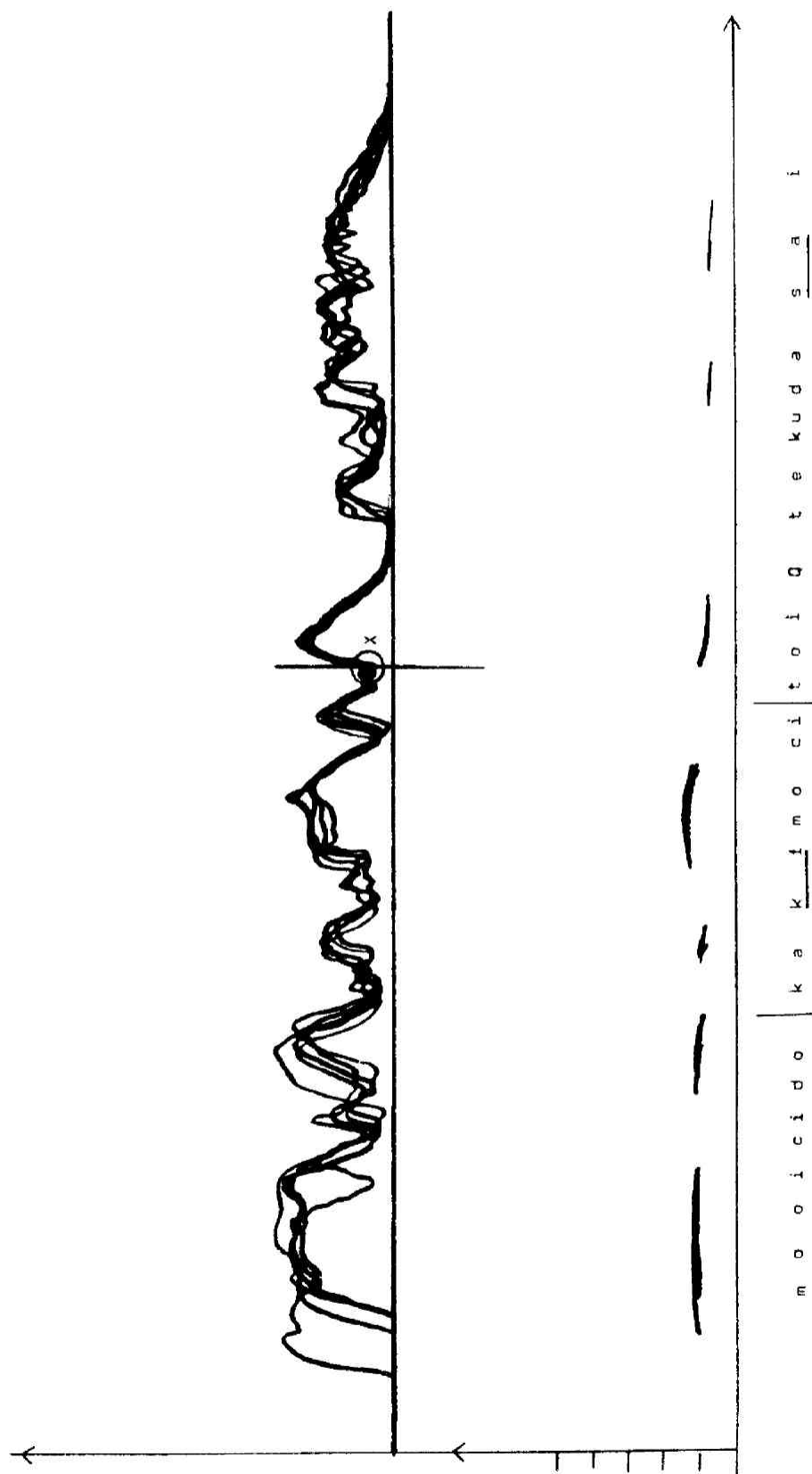


Fig. 6 Traces of F<sub>0</sub> and amplitude for / kakimoci / uttered in a frame sentence  
Five tokens are present on the Figure.

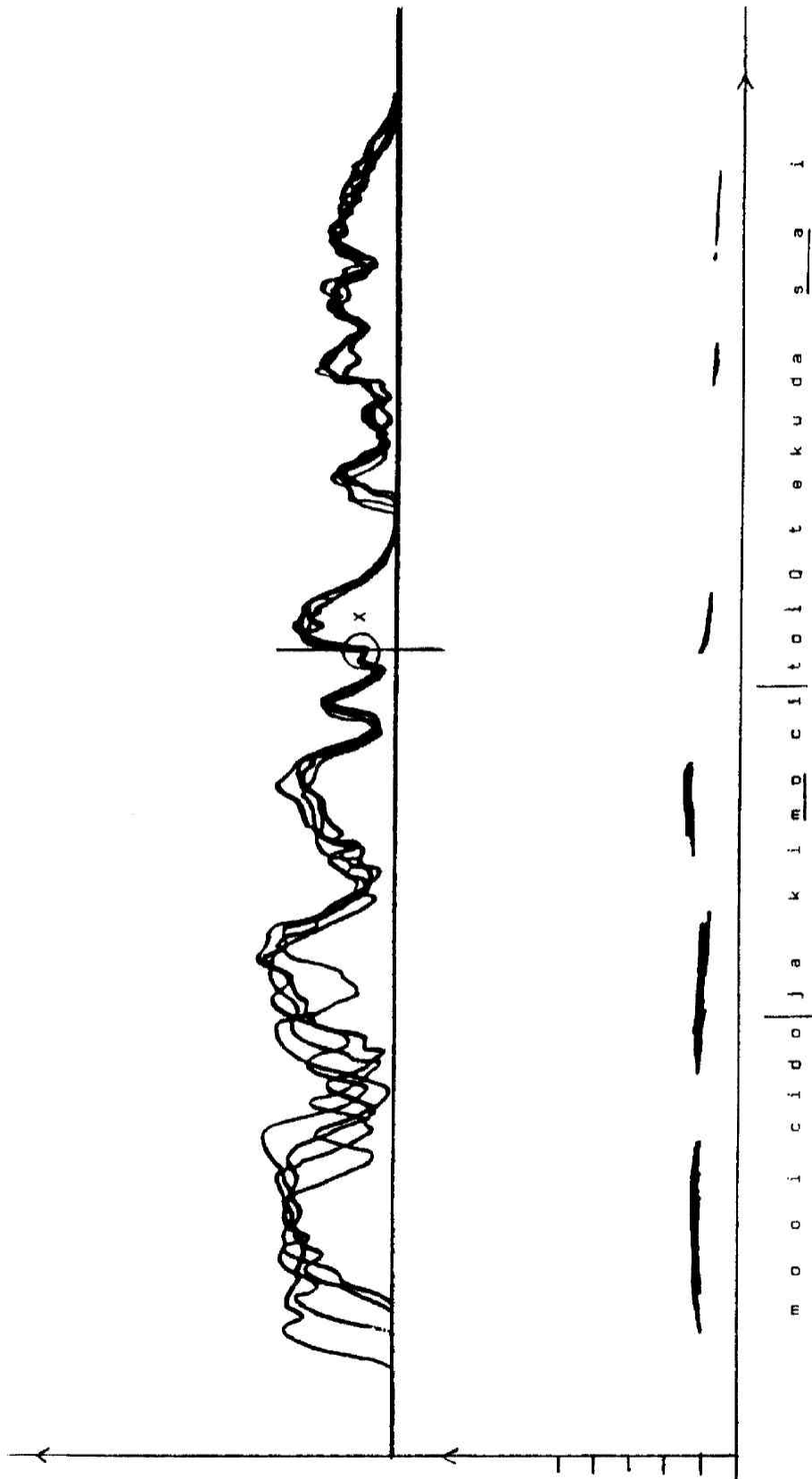


Fig. 7 Traces of  $F_0$  and amplitude for /jakimoci/ uttered in a frame sentence

Four tokens are present on the Figure.

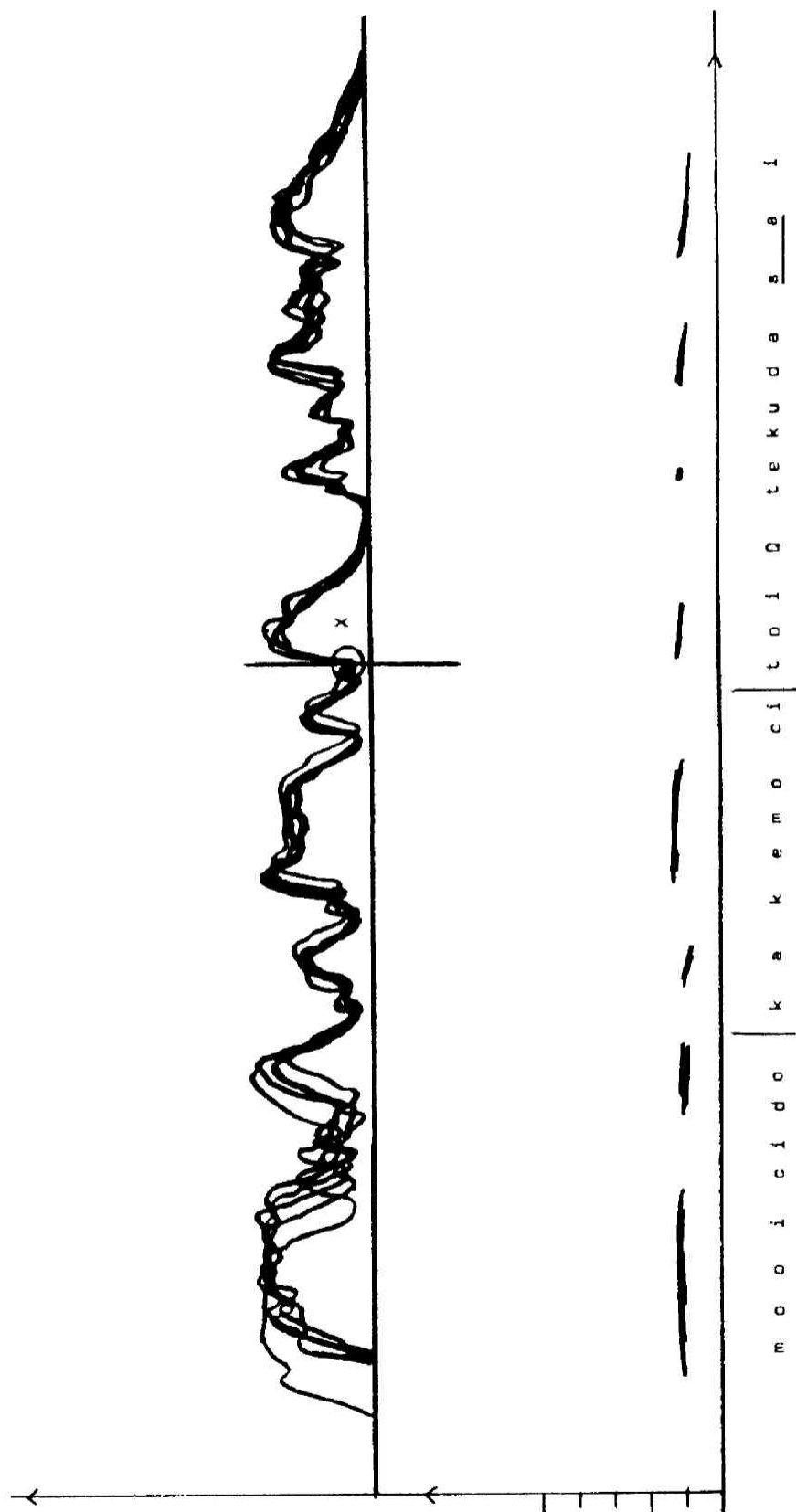


Fig. 8 Traces of F0 and amplitude for / kakemoci / uttered in a frame sentence

Five tokens are present on the Figure.

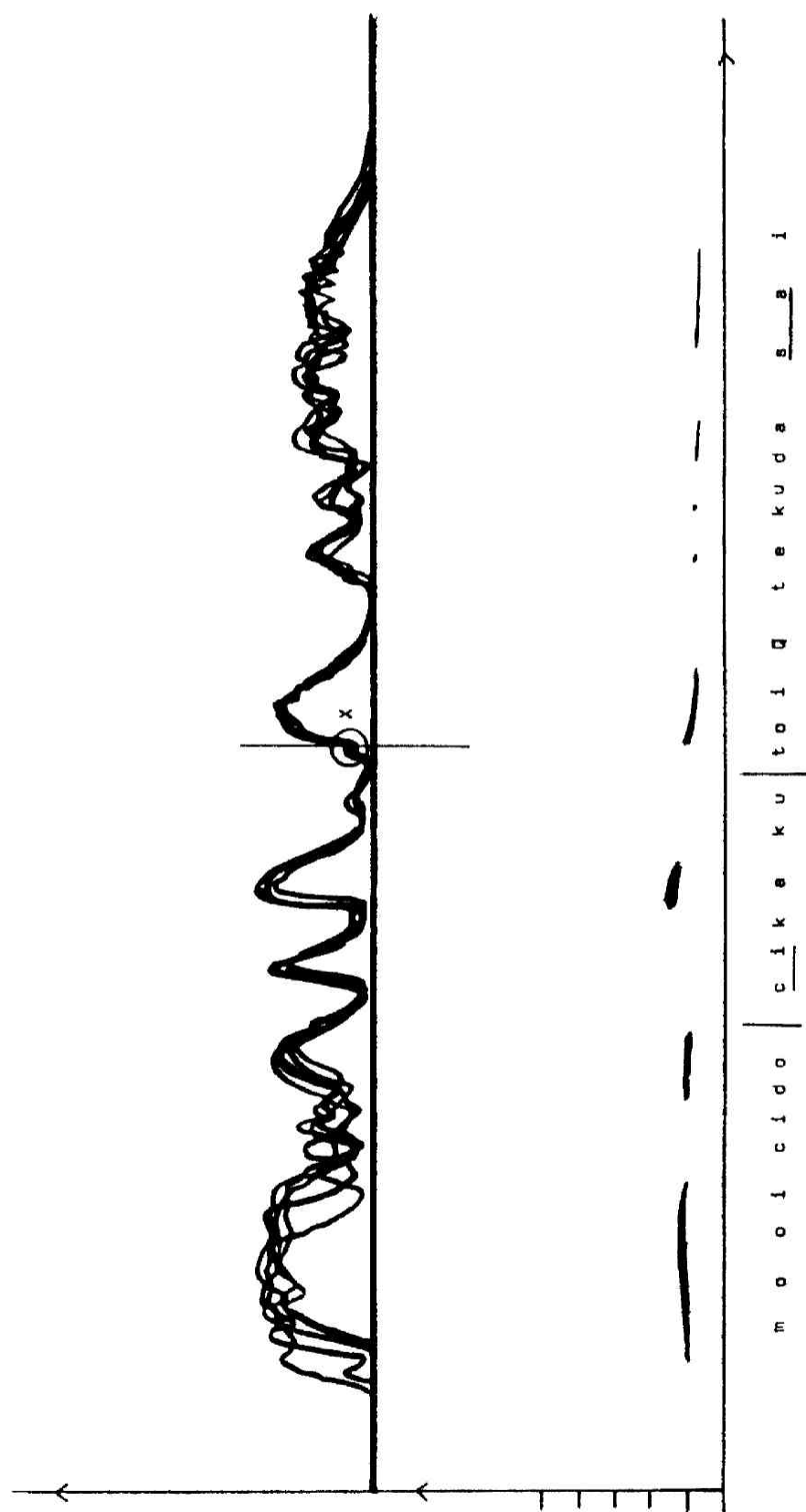


Fig. 9 Traces of F<sub>0</sub> and amplitude for / cikaku / uttered in a frame sentence

Five tokens are present on the Figure.

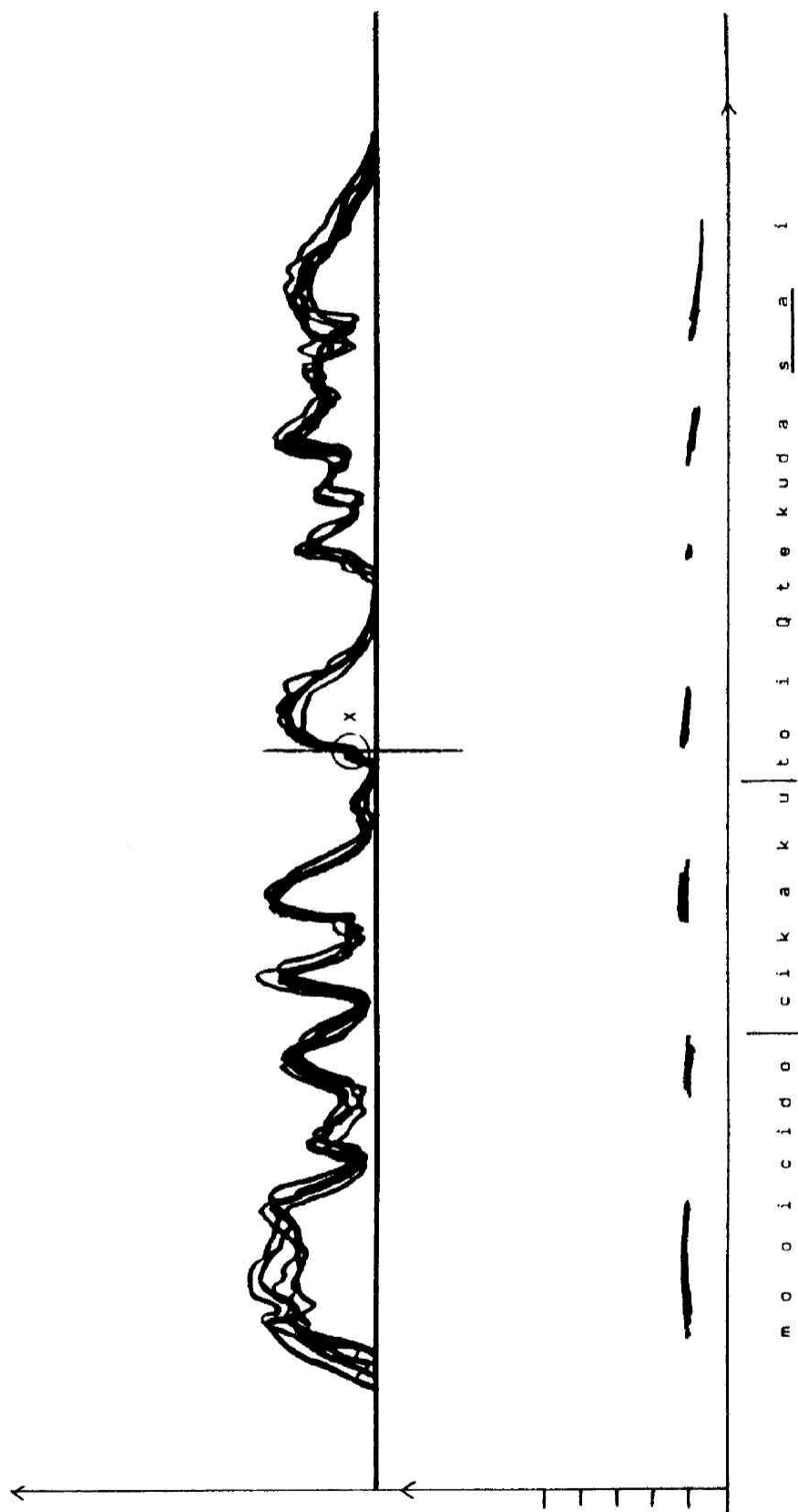


Fig. 10 Traces of  $F_0$  and amplitude for / cikaku / uttered in a frame sentence

Five tokens are present on the Figure.

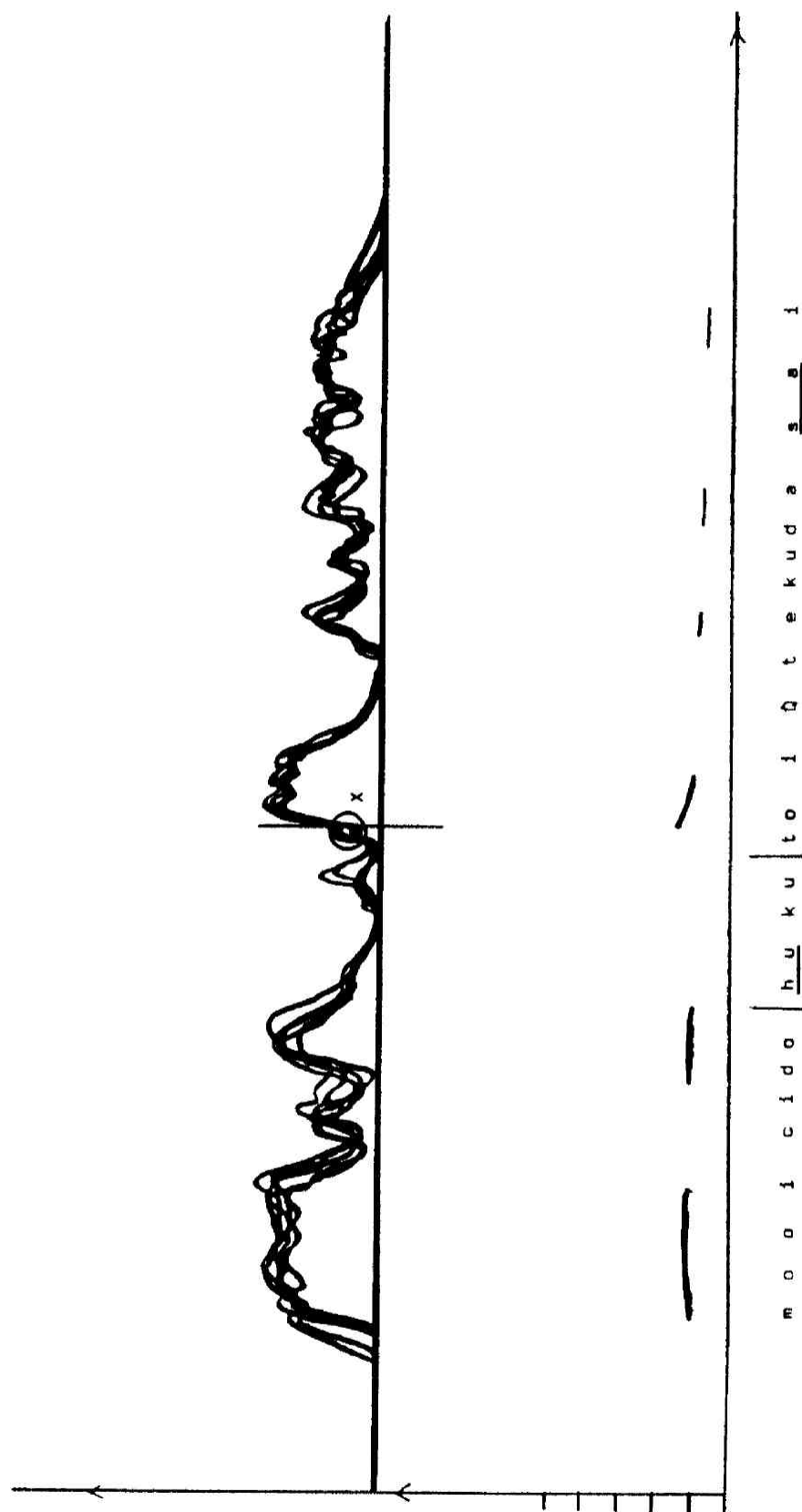
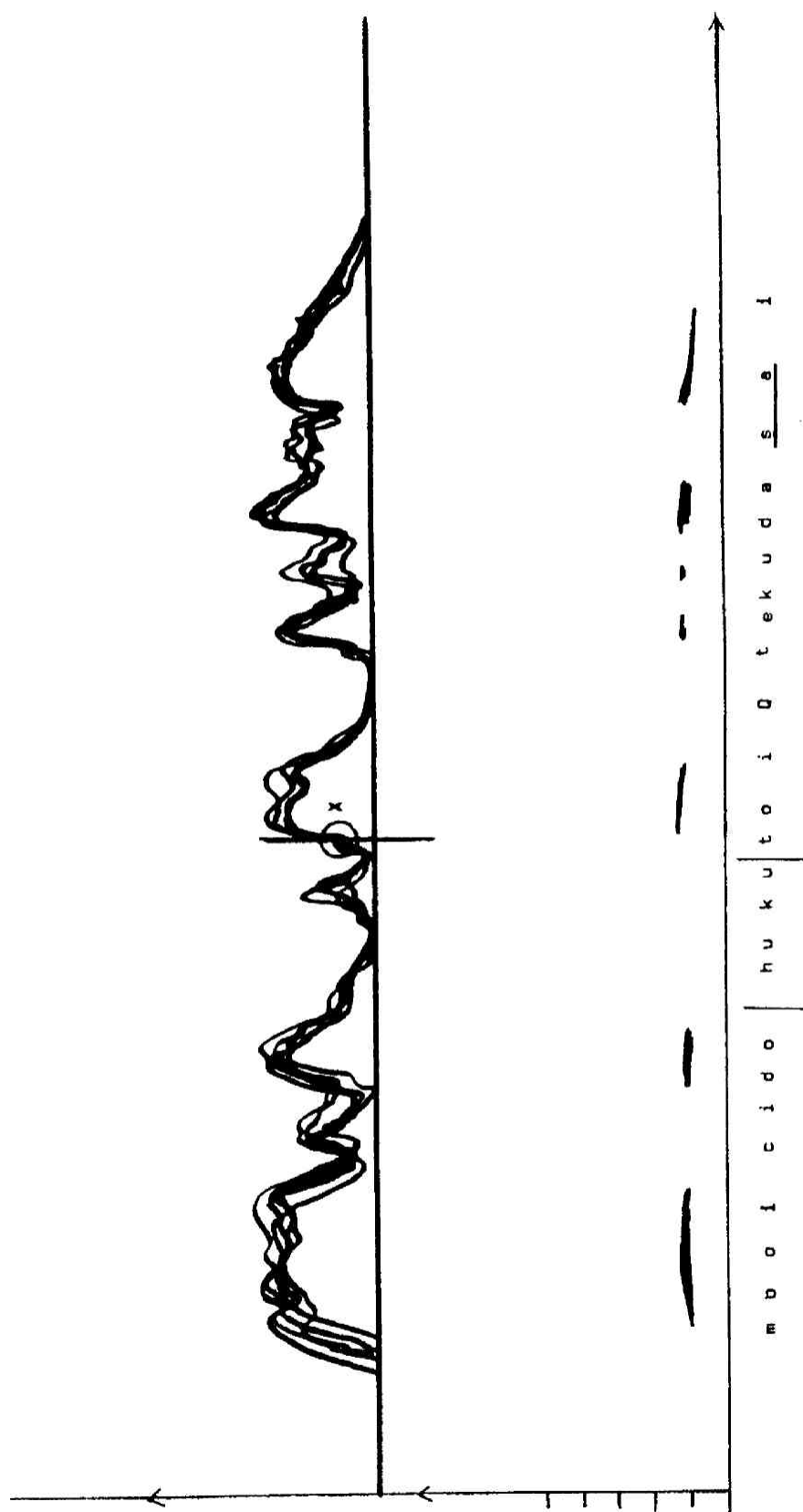


Fig. 11 Traces of  $F_0$  and amplitude for / huku / uttered in a frame sentence

Five tokens are present on the Figure.



**Fig. 12** Traces of F<sub>0</sub> and amplitude for /huku/ uttered in a frame sentence  
Five tokens are present on the Figure.



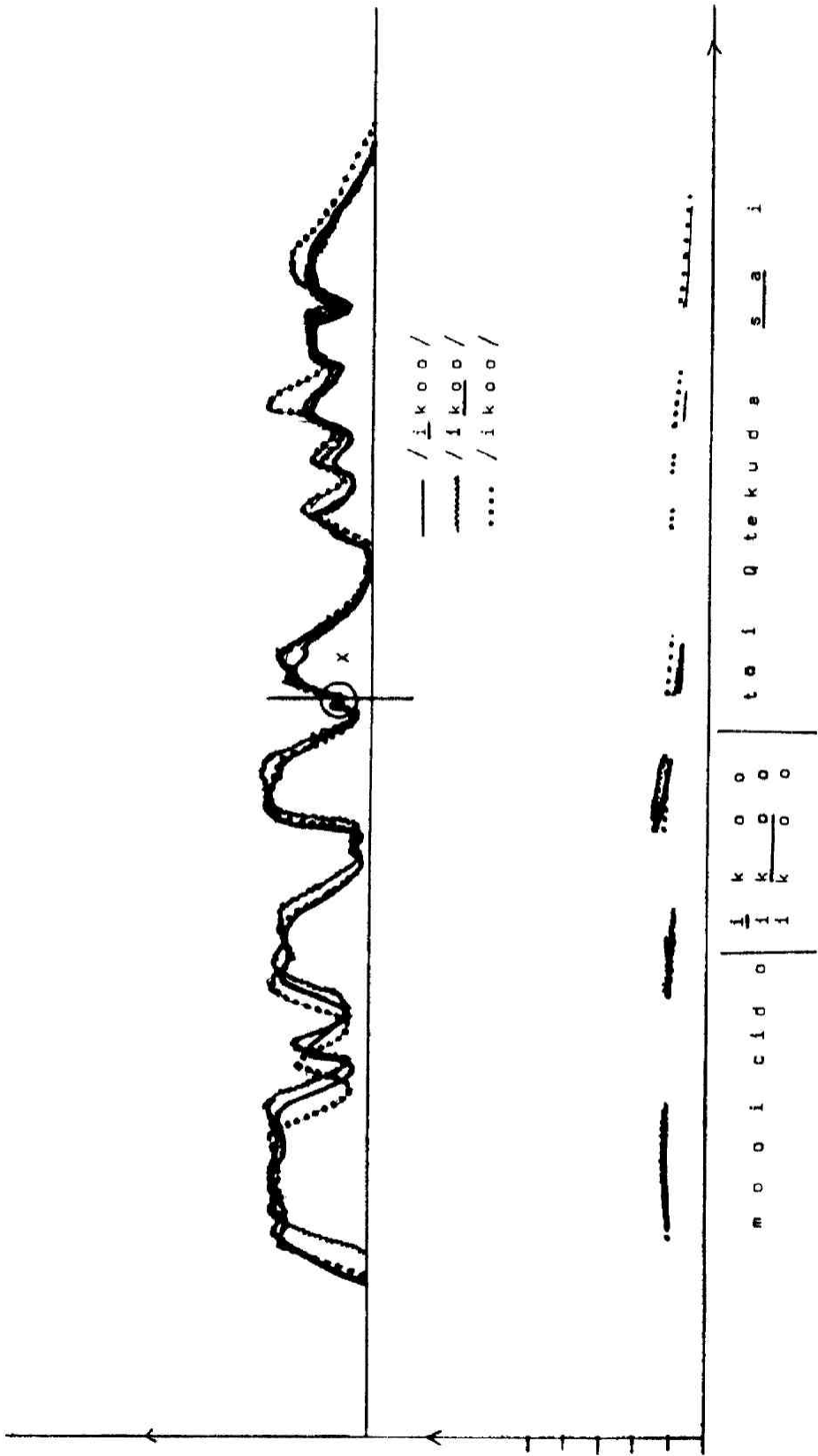


Fig. 13 Average F0 and amplitude traces for / i k o o , i k o o , i k o o / uttered in a frame sentence

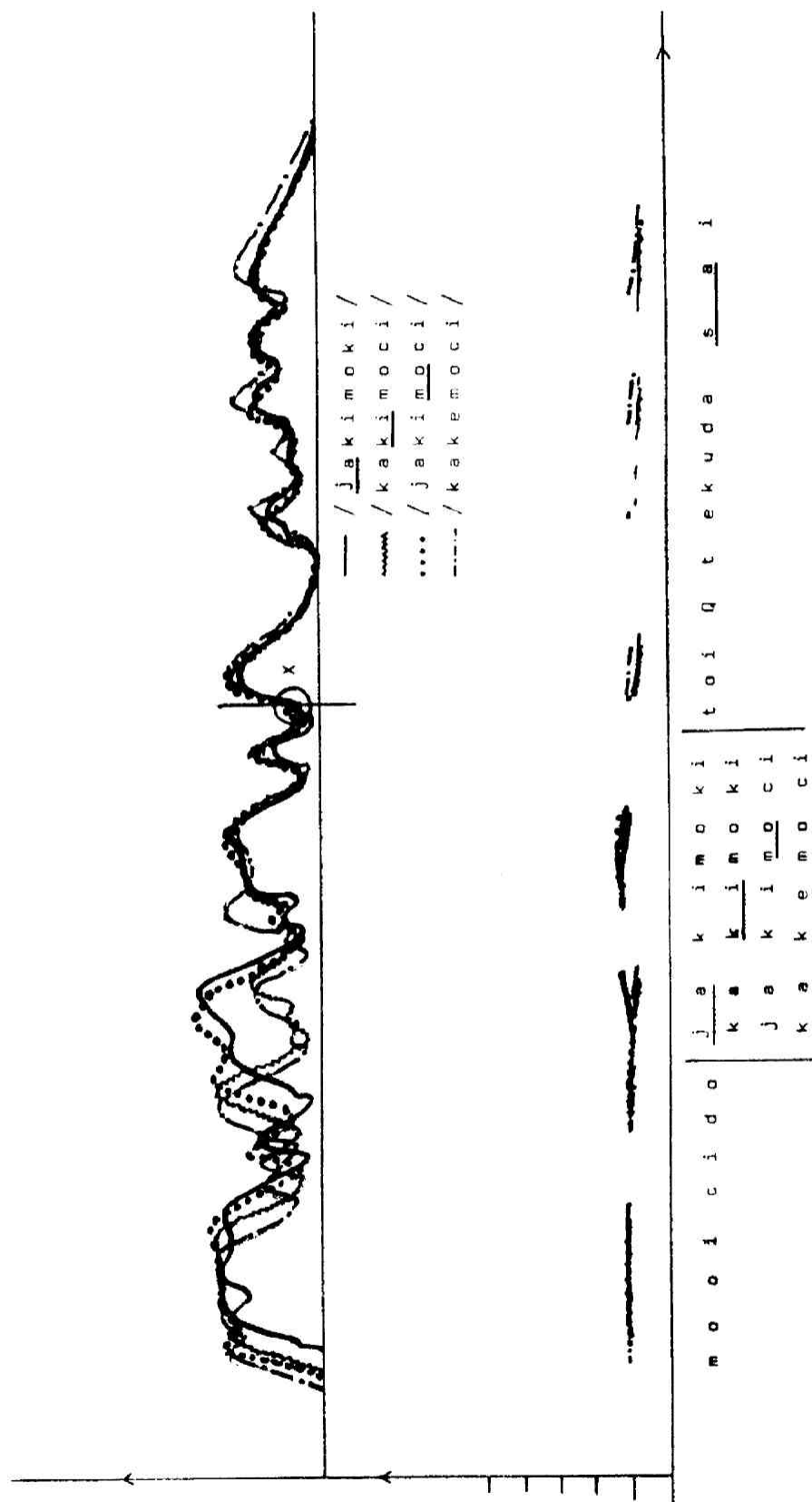


Fig. 14 Average F0 and amplitude traces for /jakimoki, kakimoci, jakimoci, kakemoci / uttered in a frame sentence

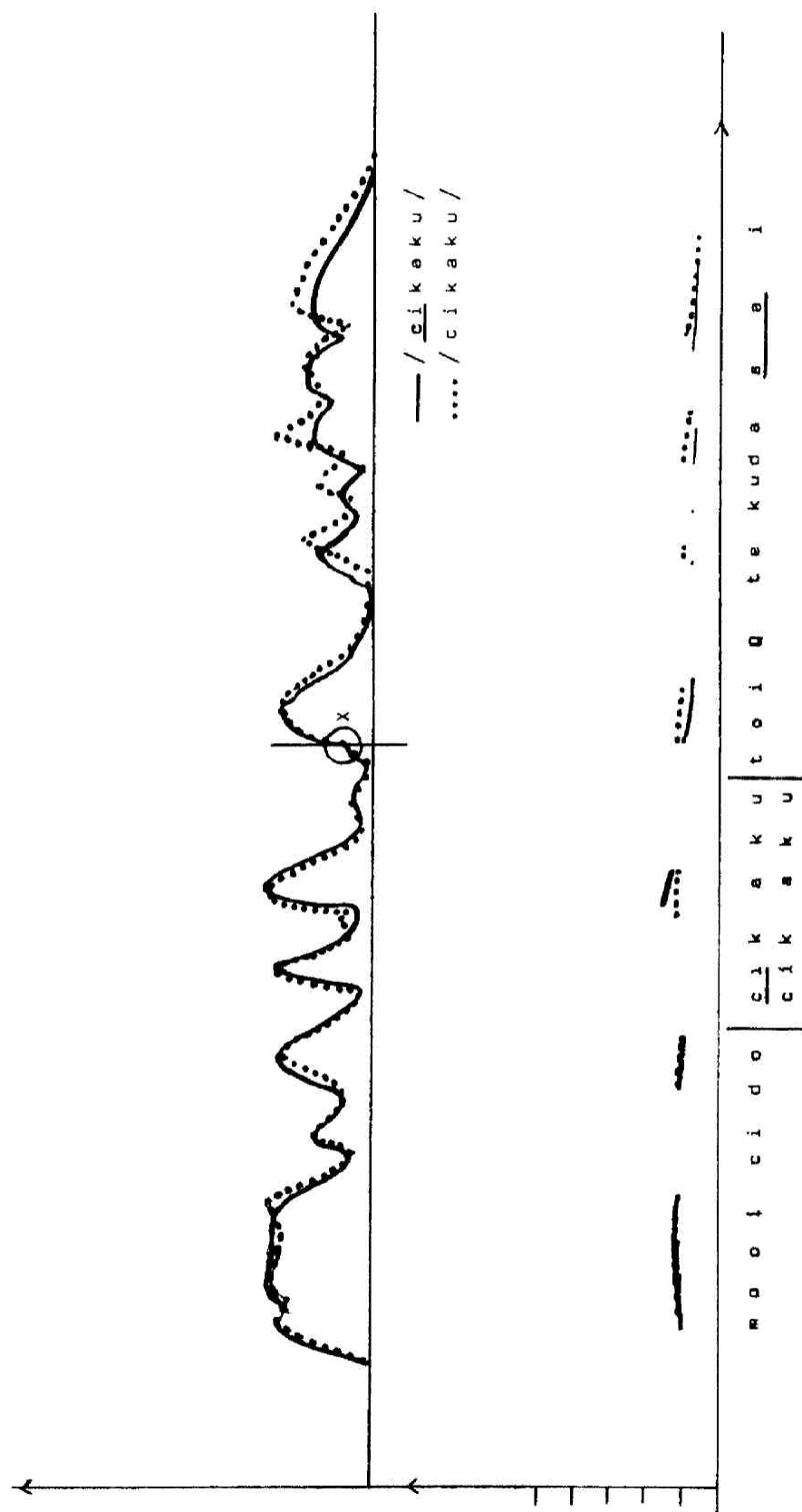


Fig. 15 Average F0 and amplitude traces for /cikaku, cikaku / uttered in a frame sentence

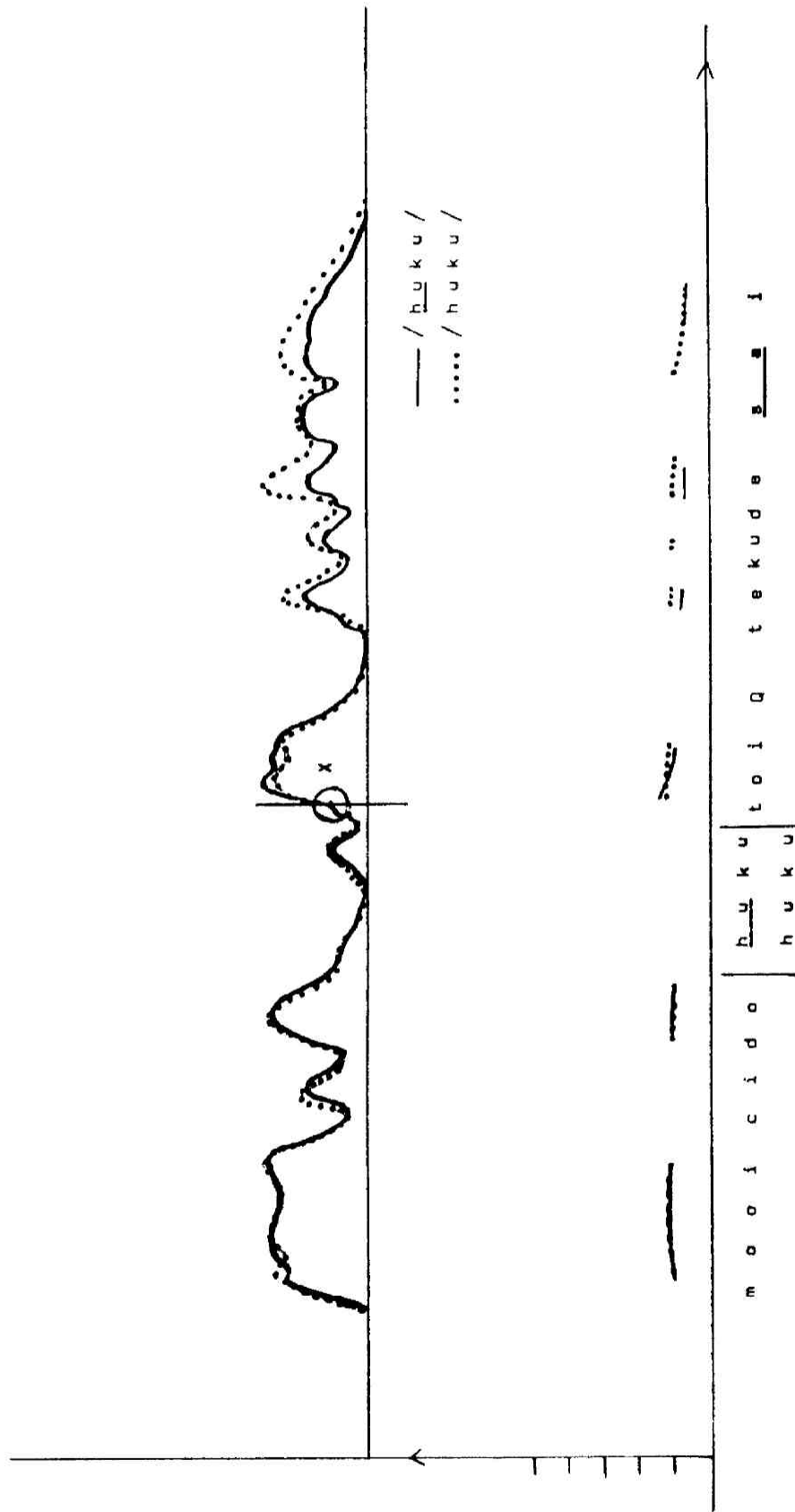


Fig. 16 Average F0 and amplitude traces for /huku, huku / uttered in a frame sentence

words are noticeably different, which suggests that the  $F_0$  is likely to be relevant to the acoustic description of Japanese accent. In order to study this point in detail, average  $F_0$  traces for these target words were visually estimated and drawn from Figs. 2 to 12 (pp. 6 to 16), and are given below (Figs. 17 to 20).

A thick solid line represents the  $F_0$  trace for the target word which is actually present on the spectrograms.

A thick dashed line represents the  $F_0$  trace for the morae preceding and following the target word which is actually present on the spectrograms.

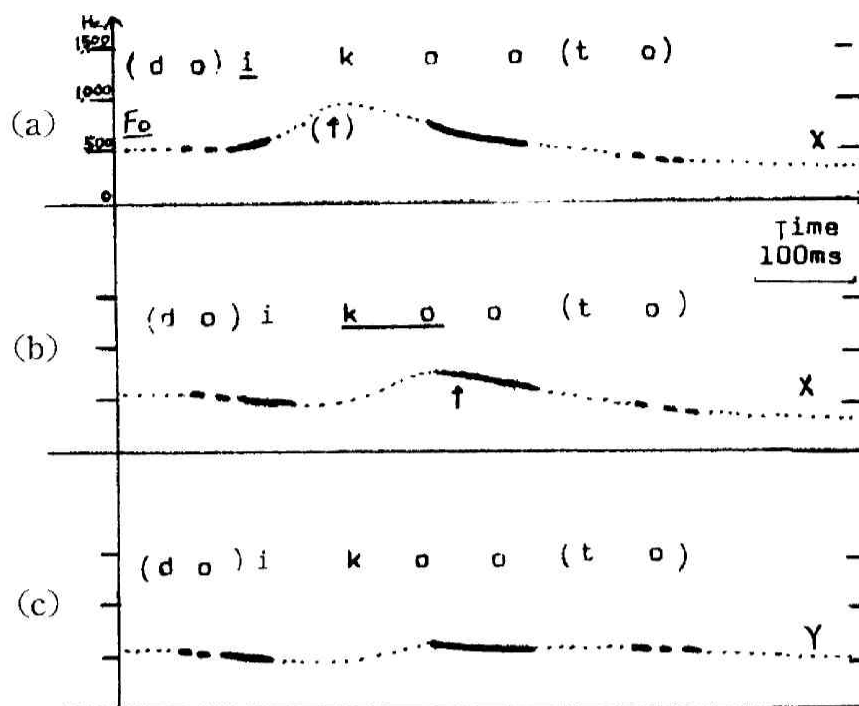
A thin dotted line represents a gap in  $F_0$  on the spectrograms which was filled in by inferring from the movements of  $F_0$  surrounding the gap (see below for more detailed descriptions of how the line was inferred in each case).

An arrow represents the approximate location where a sharp descent in  $F_0$  was actually observed on the  $F_0$  trace. When the location was inferred from the observation of the movements in  $F_0$ , it was shown by putting the arrow in brackets.

(1) Group 1 /ikoo, ikoo, ikoo / (See Fig. 17 below)

The  $F_0$  movements for / i / and / oo / for these words shown in Fig. 17 below suggest that

- (i) in the case of (a), a peak in  $F_0$  comes somewhere before / oo /, and that a sharp descent in  $F_0$  starts somewhere between / i / and / ko /,
- (ii) in the case of (b), a peak in  $F_0$  comes somewhere in the region of / o / after / k /, and that a sharp descent in  $F_0$  starts somewhere between / ko / and / o /,
- (iii) in the case of (c), a peak in  $F_0$  comes somewhere in the region of / o / after / k /, and that there occurs no sharp descent in  $F_0$ .



**Fig. 17 The average  $F_0$  traces for the words of Group 1, which were extracted from Figs. 2, 3 and 4 respectively**

- : The  $F_0$  trace for the target words, which is actually present on the spectrograms.
  - : The  $F_0$  trace for the morae preceding and following the target words, which is actually present on the spectrograms.
  - ..... : A gap in  $F_0$  filled in by inferring from the movements in  $F_0$  surrounding the gap.
  - ↑ : Location of a sharp descent in  $F_0$ .
  - (↑) : Location of a sharp descent in  $F_0$  inferred from the  $F_0$  movements surrounding the gap.
  - X, Y : The height of  $F_0$  after the target word, which shows a consistent difference (Y is higher than X) between a word with a “marked” mora (X) and a word without a “marked” mora (Y), which signals the presence or absence of sharp descent in  $F_0$  in the target word.
- The above settings apply to Figures 18 to 20 which follow.

The fact that the  $F_0$  is higher for (c) than for (a) and (b) in the trace after the target word, signals the absence of a sharp descent in  $F_0$  for (c), and the presence of it for (a) and (b). This feature is consistent throughout the traces examined below (see X and Y in the following figures).

- (2) Group 2 / jakimoki, kakimoci, jakimoci, kakemoci / (Fig. 18, p.24)

The observation of the Fo traces for these words, following the same method adopted above, suggests that

- (i) in the case of (a), a peak in Fo comes before / i / of the second mora / ki /, with a sharp descent in Fo which starts somewhere between / ja / and / ki /,
- (ii) in the case of (b), a peak in Fo comes somewhere in the region of / i / of / ki /, with a sharp descent in Fo which starts somewhere between / ki / and / mo /,
- (iii) in the case of (c), a peak in Fo comes somewhere in the region of / i / of / ki /, with a sharp descent in Fo which starts somewhere after / mo /,
- (iv) in the case of (d), a peak in Fo comes somewhere in the region of / e / of the second mora / ke /, with no sharp descent in Fo. This lack of sharp descent in Fo is also signaled by the higher Fo than the others after the target word. (see Y for (c) and X for the others in Fig. 18.)

- (3) Group 3 / cikaku, cikaku / (see Fig. 19, p. 26).

In both cases above, there is no voice present for the first mora / ci /. But in the case of (a), the Fo movement for / a / of the second mora / ka / suggests that there comes a peak in Fo somewhere before / a /, and that a sharp descent in Fo starts somewhere between / ci / and / ka /. In the case of (b), however, the Fo movement for / a / of / ka / suggests that a peak in Fo does not come before / a /, and that there occurs no sharp descent in Fo in the target word which is also signaled by the Fo height in the trace which follows the target word (see Y in comparison to X).

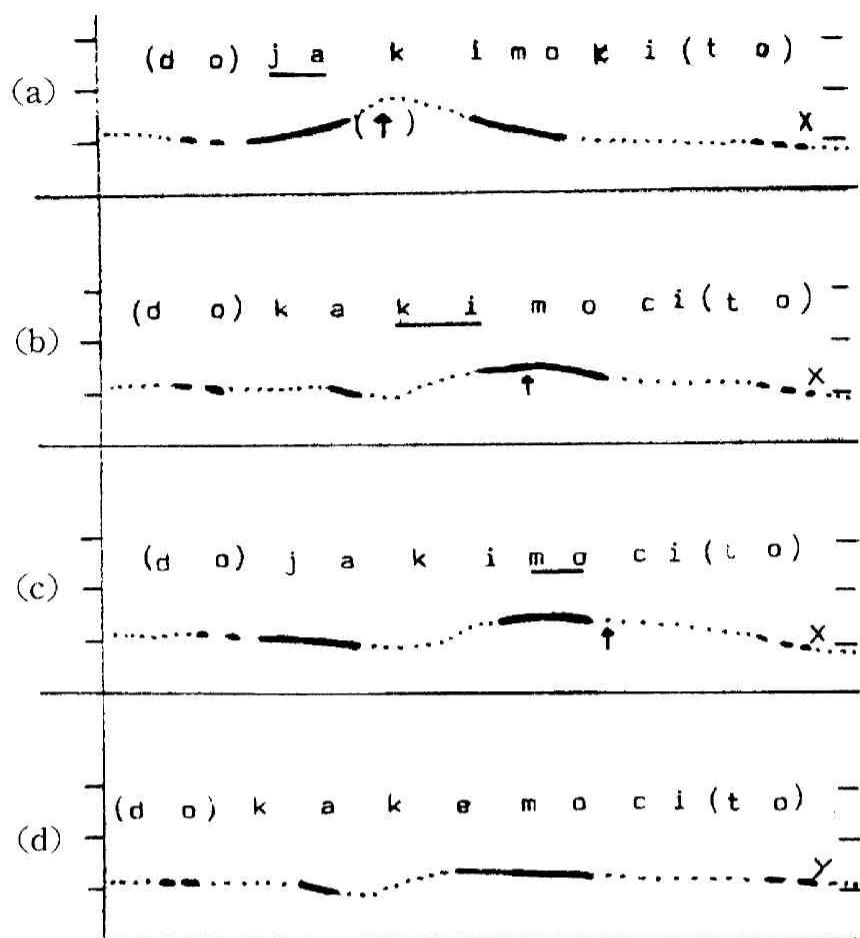


Fig. 18 The average F<sub>0</sub> traces for the words of Group 2, which were extracted from Figs. 5, 6, 7 and 8 respectively

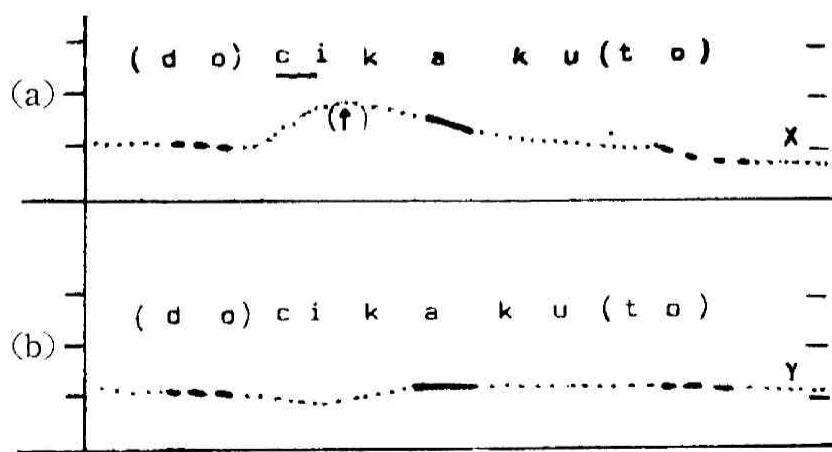


Fig. 19 The average F<sub>0</sub> traces for the words of Group 3, which were extracted from Figs. 9 and 10 respectively



(4) Group 4 / huku, huku / (see Fig. 20 below)

The traces for the above two words in Fig. 20 below show no voicing takes place for the entire words / huku /. However, the Fo movements which follow the target words suggest that, in the case of (a), a peak in Fo comes before / to /, and that a sharp descent in Fo starts somewhere within / huku /, whereas no such sharp descent in Fo occurs for (b).

As for the location of the peak in Fo for both (a) and (b), it is estimated as shown below according to the knowledge obtained in the examinations of other words in Group 1 and 2: The results in Group 1 and 2 have shown that the Fo for the first mora is always lower than for the second mora if the first mora is not "marked", and that if it is "marked", the Fo for the first mora is always higher than for the second mora.

Here again, the lack of a sharp descent in Fo is signaled by the higher Fo trace for (b) than for (a) after the target word (see X and Y in the figure below).

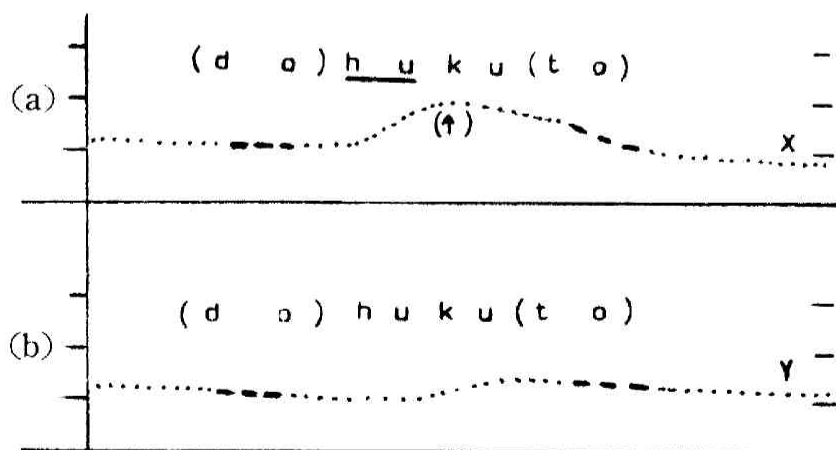
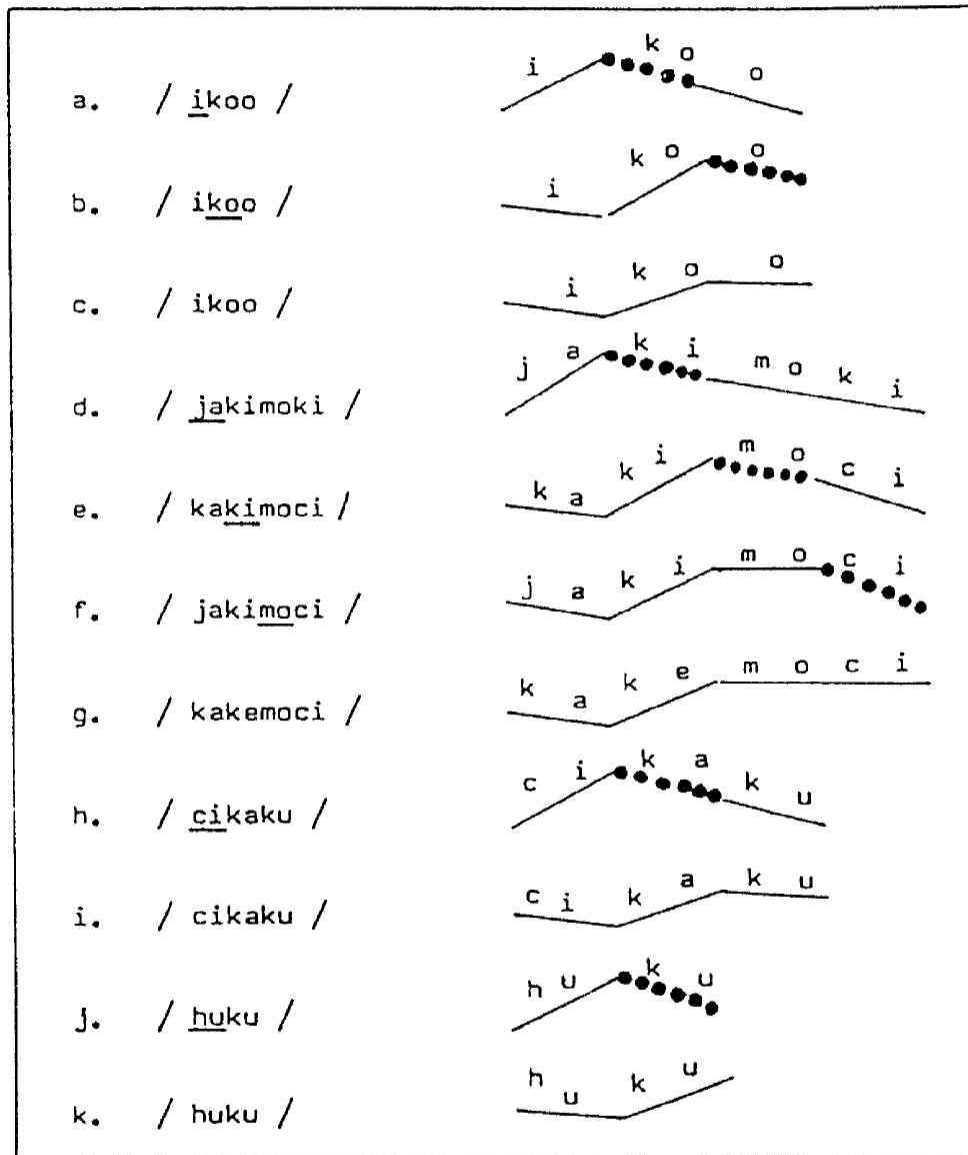


Fig. 20 The average Fo traces for the words of Group 4, which were extracted from Figs. 11 and 12 respectively

## (5) Summary (see Fig. 21 below)

Sketches are given below of the  $F_0$  traces for the eleven words examined above, which illustrate, among others, the  $F_0$  movement which follows the “marked” mora (if there is one) in the words.



**Fig. 21** Schematized  $F_0$  traces for eleven target words drawn from the average  $F_0$  traces for these words uttered in a frame sentence and shown in Figs. 17 to 20

- • • : Represents the  $F_0$  movement for the mora which follows a “marked” mora.

### Amplitude

Figs. 13 to 16 (pp. 17 to 20), which compare the average *F<sub>0</sub>* and Amplitude traces for the target words within each group, were examined to see whether the amplitude feature is also relevant in the description of Japanese accent.

Figs. 13, 15 and 16 show that, in spite of the clear differences in the *F<sub>0</sub>* traces for the three target words, there is no noticeable difference at all in the amplitude traces. In fact, they are more or less identical.

Fig. 14 shows that the *F<sub>0</sub>* traces for / jakimoki / and / jakimoci / are different, but the amplitude traces are not different (almost identical). The *F<sub>0</sub>* traces for / kakimoci / and / kakemoci / are different and the amplitude traces are also different. But the reason for this difference is thought to be due to the difference in the vocalic elements for the second mora: [i] for the former appears to be responsible for the lower amplitude and [e] for the latter for the higher amplitude. The fact that the amplitude traces for / jakimoki / and / jakimoci / are different from the traces for / kakimoci / and / kakemoci / is also thought to be due to the differences in their segmental structure: The voiced contoid [j] for the former appears to be responsible for the higher amplitude and the voiceless contoid [k] for the latter for the lower amplitude.

The Figures also show that the amplitude for “marked” mora is not necessarily higher than the amplitude for the following mora, where the *F<sub>0</sub>* starts falling sharply. In other words, a sharp descent in *F<sub>0</sub>* for the mora following a “marked” mora does not necessarily accompany a sharp drop in amplitude (for this point, see Figs. 2, 3, 6, 9 and 11 as well).

The above observations strongly suggest that the amplitude is not likely to be relevant in describing Japanese accent.

### 3.2 Head

By definition, the Head is either the first or the second mora of the first lexical item in a Tune.

There are two Heads in the Tune of this experiment. The first one is the second mora / o / of the first lexical item / moo /, the first mora / mo / being the Prehead. The second Head is either the first mora or the second mora of the eleven target words put into a frame sentence and uttered with Tune 1.

#### Fundamental frequency

As far as the first Head is concerned, it is clear from Figs. 2 to 12 that the Fo for the first Head / o / is higher than for the Prehead / mo /.

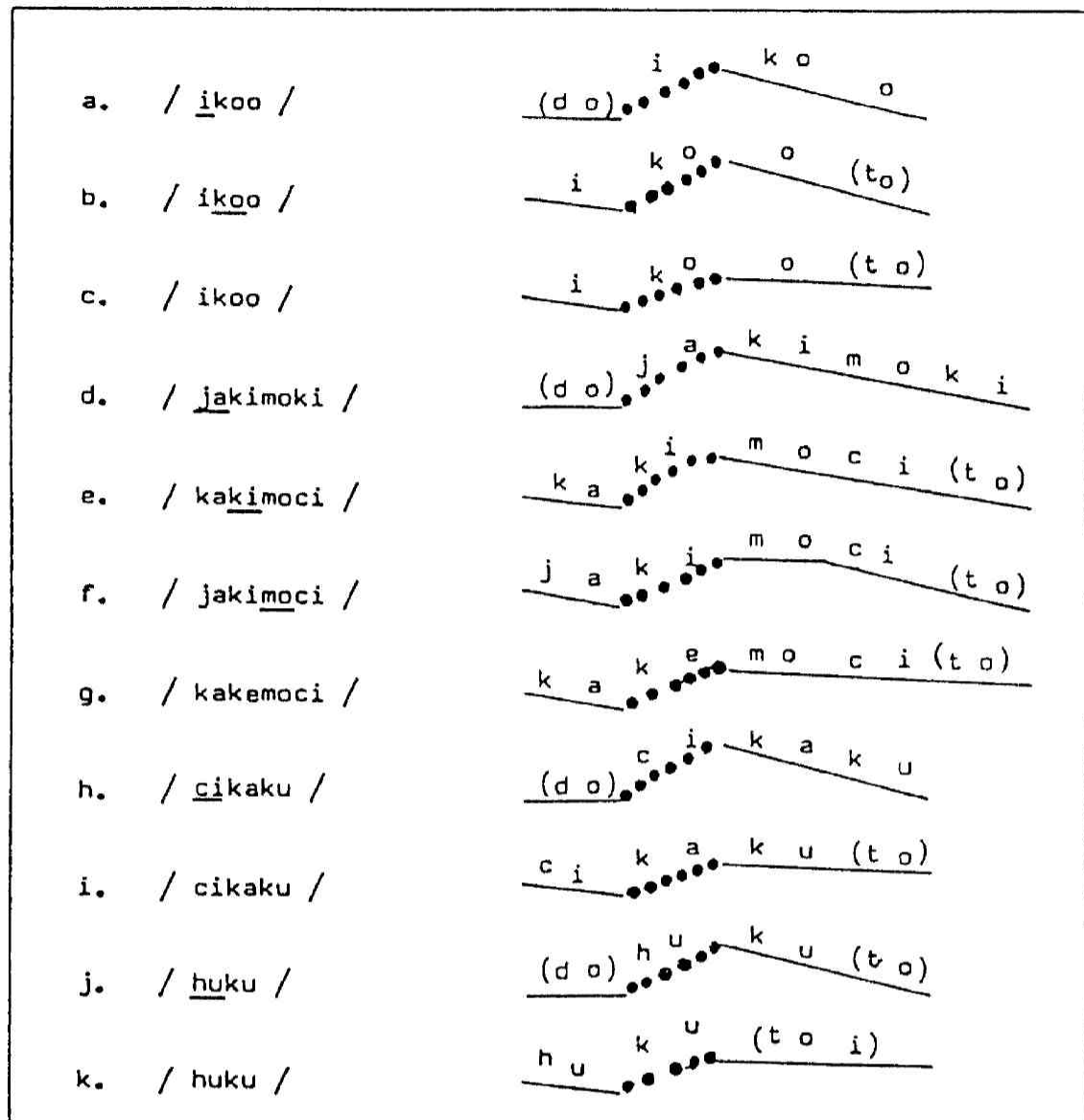
In order to examine the Fo feature for the second Head, the average Fo traces for the eleven words in Figs. 17 to 20 (pp. 22 to 25) were used. The traces were schematized focussing on the beginning part of the words and are given below (Fig. 22). A thick dotted line represents the Fo for the Head and a thin solid line for the morae preceding and following the Head.

The traces below clearly show that the Fo for the Head rises, without any exception, which means that the peak Fo for the Head is always higher than the Fo for the preceding mora.

#### Amplitude

The amplitude traces for a Head in relation to those for the preceding and following morae were examined. There was no consistent relationship of "higher amplitude for a Head than for the preceding mora." In other words, the increase in Fo for a Head, which is a consistent feature, does not necessarily accompany an increase in amplitude. In addition, no consistent relationship was found for a Head and the following mora in respect to amplitude.

The above observations suggest that the amplitude is not likely



**Fig. 22** schematized average  $F_0$  traces for eleven words which focus on the Head part in the Tune

The traces were drawn from Figs. 17 to 20.

..... : The  $F_0$  trace for the Head.

———— : The  $F_0$  trace for the morae preceding and following the Head.

( ) : The mora or morae preceding or following the target word.

to be relevant in the acoustic description of Head.

### 3.3 Unvoiced morae

In this experiment, some morae in the target words were found

to be completely lacking in voice, which was observed both in the wide-band and narrow-band spectrograms.

In the auditory analysis of intonation, intonation was perceived as continuous and unvoiced morae were plotted on the scale in brackets\*.

It is this problem of treating unvoiced morae in an intonation contour that is dealt with in this section. The morae lacking voice were studied from an acoustic point of view, and an attempt was made to see whether the above treatment was justified.

In the following, the lack of voice for certain morae in the target words is indicated by marking the vocalic elements with a circle underneath. A circle in brackets indicates that a little voice was present for some tokens.

moo	icido	ikoo	to iQte kudasai
		ikoo	
		ikoo	
		jakimoki	
		kakimoci	
		jakimoci	
		kakemoci	
		cikaku	
		cikaku	
		huku	
		huku	

As for the amplitude feature for unvoiced morae, this does not seem to serve to distinguish the accent of otherwise identical words. It was observed in 3.3.1 that the  $F_0$  for the unvoiced “marked” mora / ci / of / cikaku / (neighbourhood) was higher

---

\* ‘THE PHONOLOGY OF JAPANESE SENTENCE PROSODIES’ (1)-(4), *JINBUN KENKYU* 114-117 (KANAGAWA DAIGAKU JINBUN GAKKAI).

than for the unvoiced "unmarked" mora / ci / of / cikaku / (perception). However, as the amplitude traces for these words in Fig. 15, p. 19 clearly show (both are nearly identical in their traces), they are not differentiated from each other by the amplitude feature. The same applies to / huku / and / huku / (see fig. 16, p. 20).

As far as the Fo feature for unvoiced morae are concerned, we found in 3.3.1 that even when there were no harmonics present for certain morae, the Fo value for these morae was easily estimated based on the Fo characteristics for the surrounding morae, and that the whole Fo pattern thus worked out conformed to the general pattern for a Tune whose accentual patterns are the same, which means that the Fo pattern for a Tune is predictable if only we know the location of a "marked" mora in it.

Thus the way unvoiced morae were treated in the auditory analysis of intonation can be said to have strong acoustic evidence to support it.

#### 4 Conclusions

In this experiment Accent and Head were studied for some of their acoustic correlates, using eleven words which were grouped into four according to their segmental similarities and accentual differences, which were uttered in the frame sentence ;

/ / 1 moo icido / [ ] to iQte kudasai / /

Unvoiced morae in the Tune were also studied from an acoustic point of view. The following is the summary of the results obtained.

##### Unvoiced morae

Acoustically it was found to be possible to estimate the Fo movements for unvoiced morae by examining the movements for

**Table 1 Possible acoustic correlates for Accent and Head**

<div>Phonology</div> <div>Phonetics</div>	Accent : Phonologically defined in terms of the location of a "marked" mora
Acoustic correlates	<p>( i ) a sharp descent in Fo for the mora following the "marked" mora</p> <p>( ii ) no such sharp descent in Fo if there is no "marked" mora</p> <p>The presence or absence of such sharp descent in Fo for the mora following the "marked" mora is displayed by the degree of the steepness of the slope of the Fo trace : When there is a "marked" mora in a target word, the slope of the Fo trace is noticeably sharp, and when there is no "marked" mora, then the slope is noticeably mild. It is also signaled by the Fo height for the part of the Tune after the target word : When there is a "marked" mora, the Fo after the target word is constantly lower than there is no "marked" mora.</p>
<div>Phonology</div> <div>Phonetics</div>	Head : Phonologically defined as the first or the second mora of a certain lexical item in a Tune, which is governed by the location of the Accent of the lexical item.
Acoustic correlates	( i ) a rise in Fo for a Head ; the peak Fo for a Head is higher than the Fo for the preceding mora.

the preceding and the following morae. It was also found that the whole Fo pattern, the gaps of which were filled in, conformed to the general pattern for different utterances whose accentual patterns were identical or near identical in terms of the location of a "marked" mora.

## REFERENCES

東 淳一 (1992) 「日本語の韻律体系」『日本語音声の研究と日本語教育』



- 53-62 (「日本語音声」国際シンポジウム)
- 杉藤美代子 (1980) 「アクセント, イントネーションの比較」『日英語比較講座 1 音声と形態』107-183 (大修館)
- 杉藤美代子 (1982) 『日本語アクセントの研究』(三省堂)
- 杉藤美代子 (1990) 「日本語と英語のアクセントとイントネーション」『講座日本語と日本語教育 3 日本語の音声・音韻 (下)』349-378 (明治書院)
- ネウストuppニー, J. V. (1966) 「日本語のアクセントは高低アクセントか」『音声学会会報』121号, 1-7
- 藤崎博也 (1989) 「日本語の音調の分析とモデル化—語アクセント・統語構造・談話構造と音調との関係—」『講座日本語と日本語教育 2 日本語の音声・音韻 (上)』(明治書院)
- Abercrombie, D. (1967). *Elements of General Phonetics*. (Edinburgh University Press.)
- Fry, D. B. (1968). "Prosodic Phenomena", *Manual of Phonetics*, ed. B. Malmberg (North-Holland Publishing Co.), 365-410.
- Fry, D. B. (1970). "Speech Reception and Perception", *New Horizons in Linguistics*, ed. J. Lyons (Penguin Books), 29-52.
- Kubozono, H. (1992). "Japanese Prosody and Prosodic Theory", 『日本語音声の研究と日本語教育』175-185 (「日本語音声」国際シンポジウム)
- Ladd, D.R.(1992). "Current Theories of Prosody : an Introductory Sketch" 『日本語音声の研究と日本語教育』197-206 (「日本語音声」国際シンポジウム)
- Ladefoged, P. (1982). *A Course in Phonetics*. (Harcourt Brace Jovanovich, Inc.)
- Lehiste, I. (1970). *Suprasegmentals*. (M. I. T. Press)
- Lovins, J. B. (1976). "Pitch Accent and Vowel Devoicing in Japanese : a Preliminary Study", *Annual Bulletin of the Research Institute of Logopedics* (University of Tokyo), No. 10, 113-125.
- Weitzman, R. S. (1970). *Word Accent in Japanese*. (Acoustic Phonetics Research Laboratory, University of Southern California.)