

MEASUREMENTS OF THE LENGTH OF SOME JAPANESE VOWELS WITH SPECIAL REFERENCE TO THEIR DEVOICING

Hideo Mase

0.1. Introduction

It is known that in Japanese unaccented i and u¹ are devoiced when they occur between voiceless consonants.² At the same time, it is generally admitted that the length of a syllable³ - a mora - is fairly constant in Japanese.

In my previous experiment (5) I found the following tendency in ka-syllables: the vowel is shorter in a low-pitched syllable than in a high-pitched one, while the length of the 'open interval (= the distance from the explosion to the beginning of the vowel) + vowel' is almost the same in both high- and low-pitched syllables. (See Fig. 1 in Mase (5), p. 154.)

- 1) "Accented" and "unaccented" means "phonologically accented" and "phonologically unaccented". An accented vowel is phonetically a high(er)-pitched vowel before word border or immediately followed by a syllable with a low(er)-pitched vowel, the latter being an unaccented vowel. An unaccented vowel is any vowel which does not fulfill the above condition of the accented vowel. An unaccented vowel is low(er)-pitched if it occurs in the initial syllable or in the syllable after the syllable whose vowel is accented; it is high(er)-pitched if it occurs in a non-initial syllable before the syllable whose vowel is accented.

The present investigation is confined to vowels which occur in the initial syllable, so that high- and low-pitched vowels are phonetic manifestations of phonologically accented and unaccented vowels, respectively.

- 2) Besides, unaccented u after a voiceless consonant and in phrase- or utterance-final position is often devoiced.
Details about conditions on devoicing are found in M.Han (2), p. 17-34. A summary of her description is found in my previous report(5), p. 146-47.
- 3) A phonological syllable consists of one or two morae and of one phonetic syllable. (Cf. Hattori (3) and (4), p. 751.) When a phonological syllable consists of two morae, the second mora is always a bound mora, since it cannot take accent, and since it always occurs after a free mora. For the experiment reported here only syllables consisting of one mora were chosen, so that a phonological syllable accords with a phonetic syllable and a 'free' mora. I sometimes use 'syllable', sometimes 'mora' - a unit of length - depending upon the situation.

This lengthening of the open interval in low-pitched syllables is presumably partly due to the devoicing of the following vowel. It is not possible to distinguish such a devoiced phase from the aspiration on the mingograms, and hardly on the spectrograms either. Therefore, the devoiced beginning of the vowel will be taken to belong to the preceding consonant, or rather to the open interval of the preceding consonant. This is also true of the devoiced beginning of the vowel after the spirant ʃ: it is not possible to distinguish the devoiced part from ʃ. It is therefore interesting to investigate whether there is a vowel-shortening in low-pitched position, which corresponds to a lengthening of ʃ and of the open interval of a stop consonant, or whether there is also a lengthening of the closure period, i.e. lengthening of the whole consonant. In m + V-syllables, for example, there is no devoicing of the vowel, and the syllable length is fairly constant. Accordingly, a lengthening of ʃ and of the open interval of stop consonants and a shortening of the vowel can be considered as a compensation between the consonant and the vowel, though the principle of compensation must be applied with care, because there may be various heterogenous factors involved⁴.

The present experiment was undertaken to investigate the influence of initial voiceless consonants on the devoicing of vowels. Completely devoiced vowels will, however, not be considered.

0.2. Material and informant

The stop consonant k and the spirant ʃ (hereafter written S) are chosen here as initial consonants in a CV-syllable. The disadvantage in choosing k is clear, because the consonant is accompanied by much aspiration and affrication, especially when it is combined with high vowels, viz. i and u. But at

4) Cf. Eli Fischer-Jørgensen (1, p. 200 ff).

the same time, since the open interval is longer in k than in t and p (cf. 1.0.), the difference in length of open intervals is easiest to observe in k. The choice is also due to a phonological restriction. The phoneme (in the traditional sense) /t/ is phonetically manifested as [t] when it occurs before /a/, but as [tʃ] before /i/ and as [ts] before /u/. Word-initial /p/ ([p]) occurs only in loan words. s was chosen because it is found before i, a and u ([s] does not occur before i). Some other consonants have also been used as initial consonant. t is taken up in 1.0. and others in 1.2.

The vowels a, i and u have been used. The reason why a was chosen is that the vowel is the longest of the 5 Japanese vowels (cf. Han (2), p. 16), and is least influenced by surrounding consonants. i and u were used for comparison with a in the same consonantal environment.

The investigated high- and low-pitched syllables are followed by another CV-syllable where C- is t, d, k, s or m. These CVCV-syllables make up a phrase. In the text the phrase is preceded by a frame 'Sorewa' ("It (is) ...") - a frame which does not destroy the pitch contrast in the first syllable of the following phrase. Test words were arranged rather in random, but were ordered in such a way as to prevent unnatural lengthening or shortening of syllables or words. Two words, kù-t and kù-d⁵, have no low-pitched counterparts, since I could not find any suitable examples.

Two standard-dialect speakers were chosen as informants. One is a male speaker (MM) (post-graduate student) and the other is a female speaker (NF) (born in the 1930ies). In MM's speech devoicing is so conspicuous that i and u between voiceless consonants are almost completely devoiced both in high- and low-pitched environments. (But devoicing of an accented vowel in this consonantal environment does not on the whole seem to be rare in Japanese. My previous investigation (5) was

5) The symbol ` denotes accent.

about this very phenomenon observed in three male speakers other than MM.) NF has a flapped r, and MM's /r/ is either r-like or l-like. MM has an intervocalic g where NF has ŋ. u may be slightly rounded in the speech of both persons. (Japanese u is unrounded and tense, but is not so back and high as Cardinal No. 16.)

0.3. Recording and registration

Each test word was spoken 10 times by each person. The recording was made in the recording studio of the Phonetics Institute of the University of Copenhagen.

An 8-channel mingograph was used to register the test material. On the mingogram are shown: 1. (= the first channel from the top) duplex oscillogram; 2. logarithmic intensity curve, high-pass filtered at 500 cps.; 3. logarithmic intensity curve without filtering; 4. fundamental frequency curve; 5. logarithmic intensity curve, high-pass filtered at 2000 cps.; 6. oscillogram; 7. logarithmic intensity curve, low-pass filtered at 500 cps. But there was something wrong with this last curve (No. 7).

The integration times used for the intensity curves are 2.5 ms (for NF) and 5.0 ms (for MM) for the curves 2, 5 and 7, and 5.0 ms (for NF) and 10.0 ms (for MM) for the curve 3. Measurements are estimated to be precise within ± 0.25 cs. But since the border between the preceding vowel (in the frame) and the beginning of the closure period of the stop consonant is sometimes not so exact, there may in some cases be a little more uncertainty (about ± 0.5 cs).

1. Results

1.0. General comments on the length of the open interval

It is generally said that the open interval in Japanese voiceless stops is not very long, and that the aspiration is weak. According to Han (2, p. 57):

"Voiceless stops show a slight aspiration. This is most notable with the release of [k] followed by [i] and [ɨ]. The duration of the aspiration in this position is two to three centiseconds. Other stop consonants, and [k] before other vowels, show aspiration of one

to two centiseconds. The amount of aspiration is not a distinctive feature in Japanese, and it differs greatly from individual to individual and from dialect to dialect."

As will be shown in the following section, the stop consonant k of both NF and MM shows a longer open interval (except NF's ka's) than that which is described in Han (2). The same tendency, i.e. longer open interval, was also observed in other persons' speech in the previous investigation⁶. (Cf. the above-mentioned Fig. 1 in Mase (5).)

But a longer open interval does not necessarily mean a stronger aspiration. (Cf. 3.1.) In both persons' speech a longer open interval is caused partly by aspiration, partly by affrication, and partly by an interval of 'voicing lag' whose intensity is very much weaker than that of aspiration and affrication noise. By the way, the duration of the explosion is not stated in Han (2).

For comparison some ta-combinations (10 examples of each word) are shown here. (See Fig. 1.) The open interval of t is between 1.15 and 1.75 cs (NF's average value), varying from 0.5 to 2.0 cs (individual values), and for MM it is on an average between 2.65 and 2.95 cs when followed by a high-pitched a, and 5.77 cs when followed by a low-pitched a. Here the length of the open interval seems quite normal, except for the last sample of MM's recording, where the tendency is the same as in ka.

1.1. Average length of k- and S-syllables

The following figures are the average values of 10 or 9 examples of each word. (In individual words there is some overlapping.) Only the pairs where the vowel is partly (de)voiced are included.

(The symbol ">" is to be read as "longer in high-pitched environment", the symbol "<" means "longer in low-pitched environment", and the symbol "=" means "the same in both environments".)

6) Spectrograms in the paper by Torii (6) show, as far as I can see, longer open intervals for k than that which is described in Han (2).

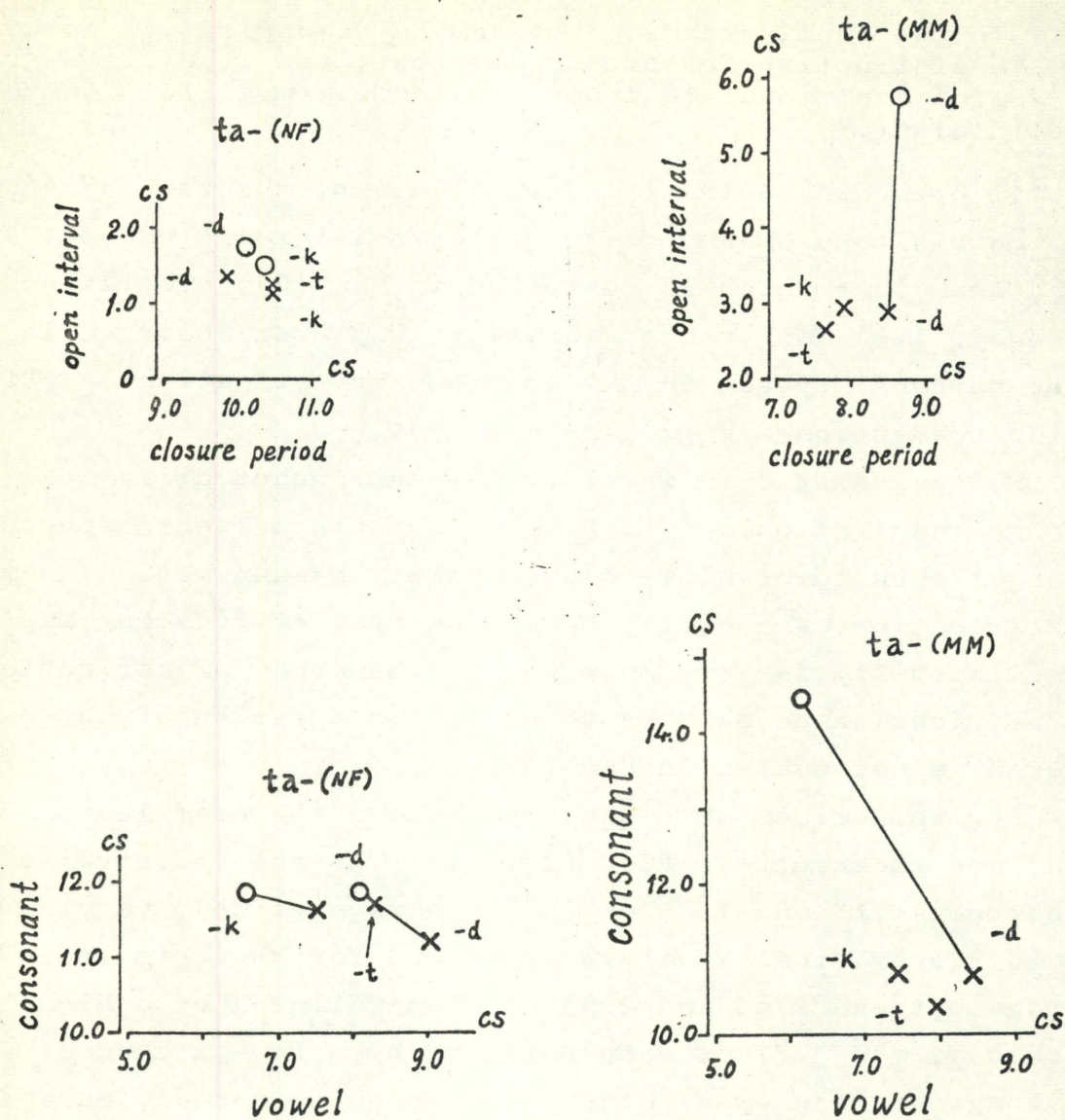


Fig. 1.

Average durations of closure period and open interval, and of consonant and vowel in ta-syllables.

X = in high-pitched environments, O = in low-pitched environments. -C = consonant following the syllable. Paired words are connected by lines.

1.1.1. The syllable (See Fig. 2)

In NF's speech the high-pitched syllable is in general (11 out of 15 pairs) longer than the low-pitched one ($>0.7 - 0.9$ cs, except ka-m (>0.15 cs)). In ka-s both syllables have the same length. In 3 cases, the low-pitched syllable is longer. All of these are examples of ki- and ku- syllables, i.e. ki-d (<0.85 cs), ki-m (<0.15 cs) and ku-m (<1.9 cs). - The general tendency in MM's speech is just the opposite. The low-pitched syllable is longer than the high-pitched one (11 out of 13 pairs) except for ka-d (>0.3 cs) and ka-m (=). The difference is between 0.5 and 1.2 cs, except for ka-s (<0.2 cs) and Sa-k (<2.5 cs). This seems to be due to extra-lengthening of the consonantal part.

1.1.2. The consonant (See Figs. 2 - 6 and 9)

The total duration of consonants (k, (t), S): All the consonants are longer in the low-pitched environment.

The closure period of stop consonants: except for NF's ka-k (>0.45 cs) (and ta-k (>0.1 cs)), and MM's ka-d (=), the closure period is longer before the low-pitched vowel than before the high-pitched one (cf. 1.4.).

The open interval (from the explosion to the beginning of the vowel): Except for NF's ka-t (>0.2 cs), the open interval is longer in the low-pitched environment, though the difference is not so great in NF's speech as in MM's.

1.1.3. The vowel (See Figs. 2 - 6 and 9)

The high-pitched vowel is longer than the low-pitched one, even though the difference is in some cases very small. In ku-m, the only example of u, the vowel duration is almost the same in both pitch environments. This is also true of MM's i's, where only the lengthening of the consonant in the low-pitched syllable is remarkable. Further, in individual cases of i and u, there is much overlapping. This is true of both persons' speech.

1.1.4. C/V ratio in the syllable

As is clear from 1.1.2. and 1.2.3., the ratio of the con-

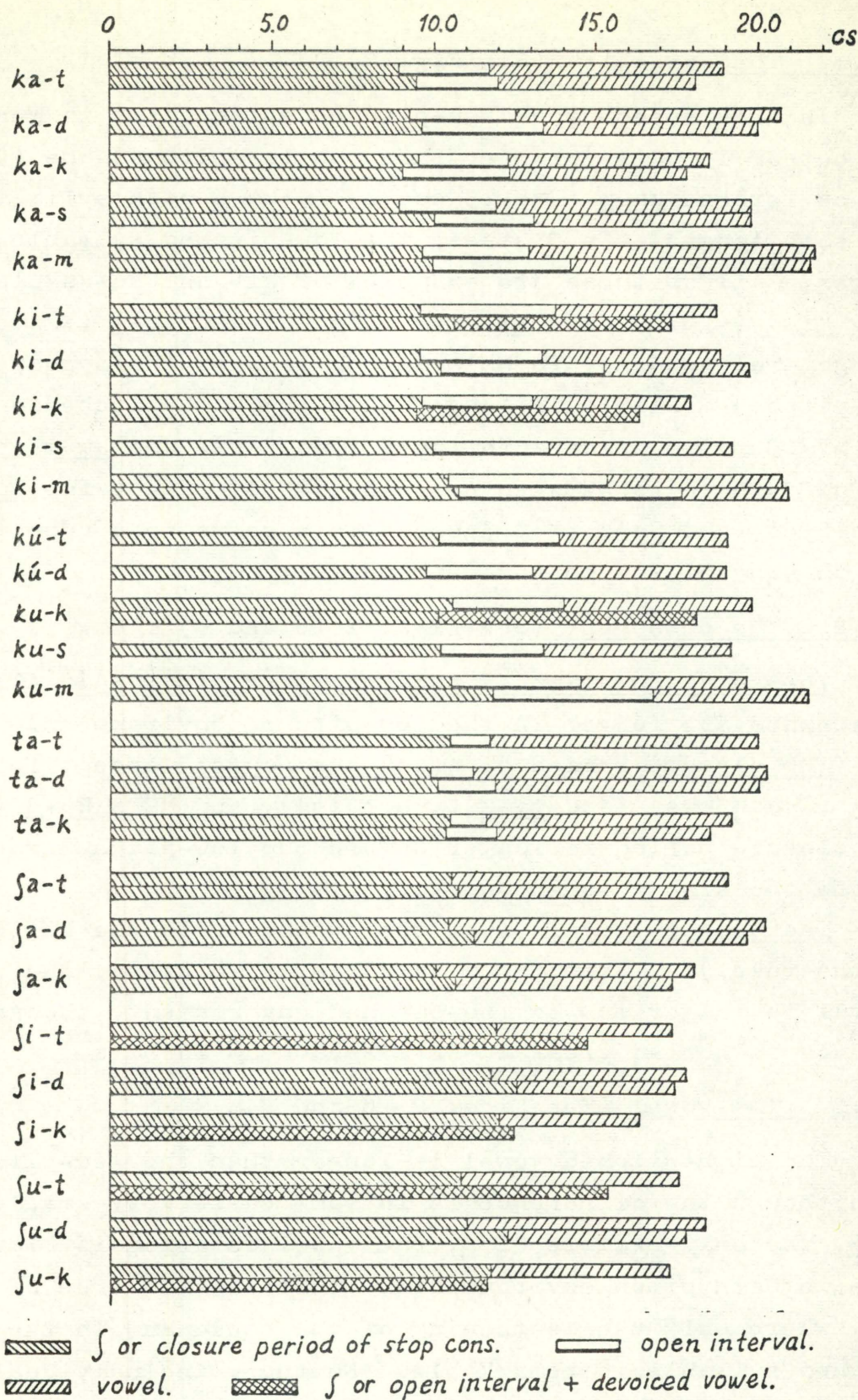
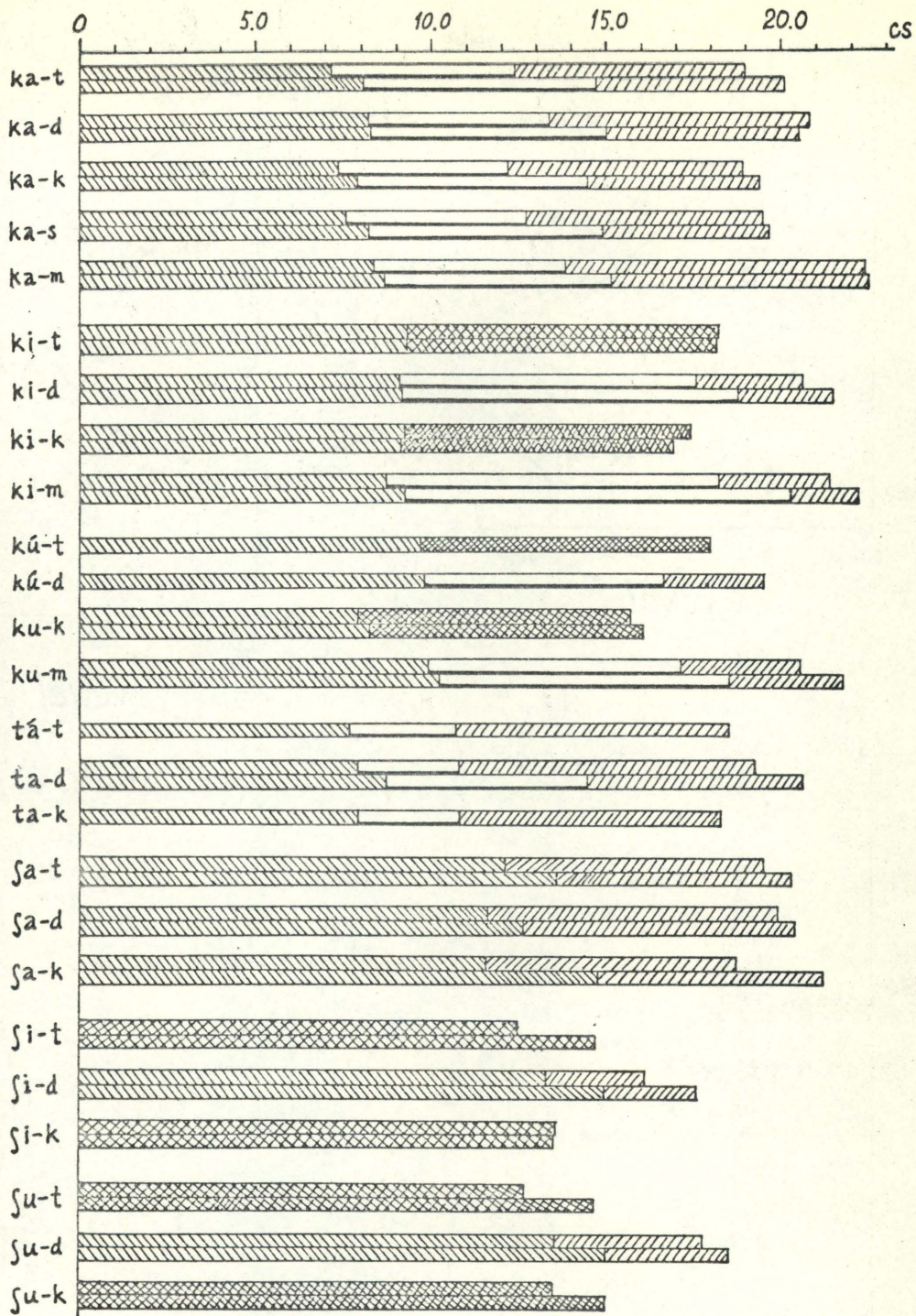


Fig. 2-a. Average durations (NF).

The first one in each pair is the high-pitched one.



▨ or closure period of stop cons. ▬ open interval.
 ▩ vowel. ▤ or open interval + devoiced vowel.

Fig. 2-b.

Average durations (MM).

The first one in each pair is the high-pitched one.

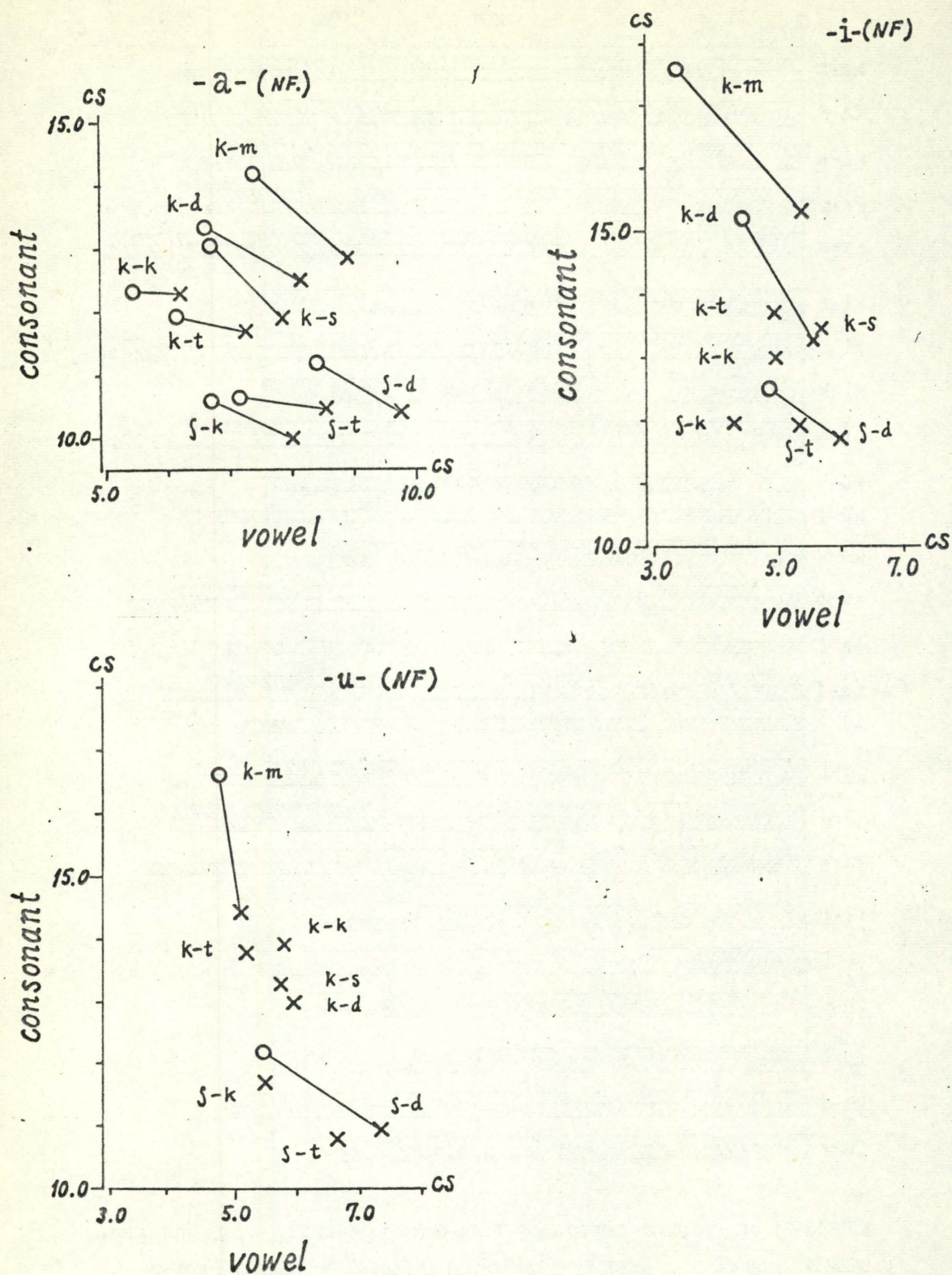


Fig. 3-a.

Average durations of consonant and vowel (NF).
(For symbols, see Fig. 1.)

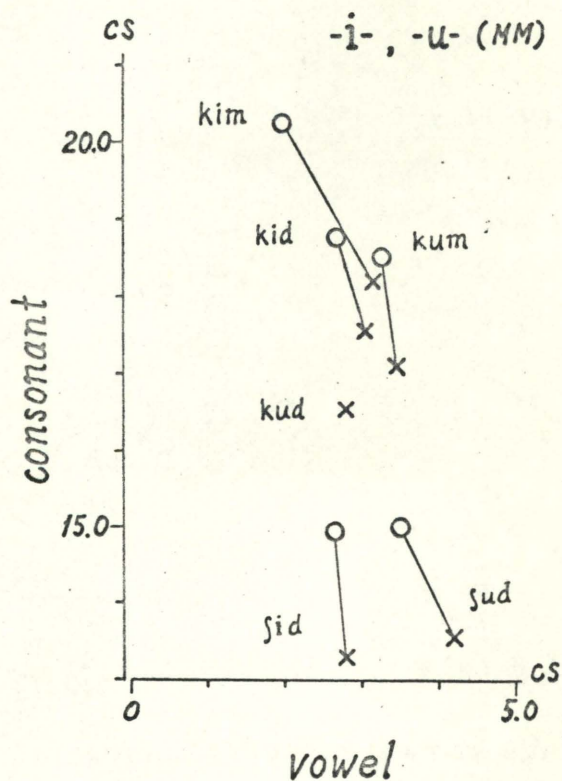
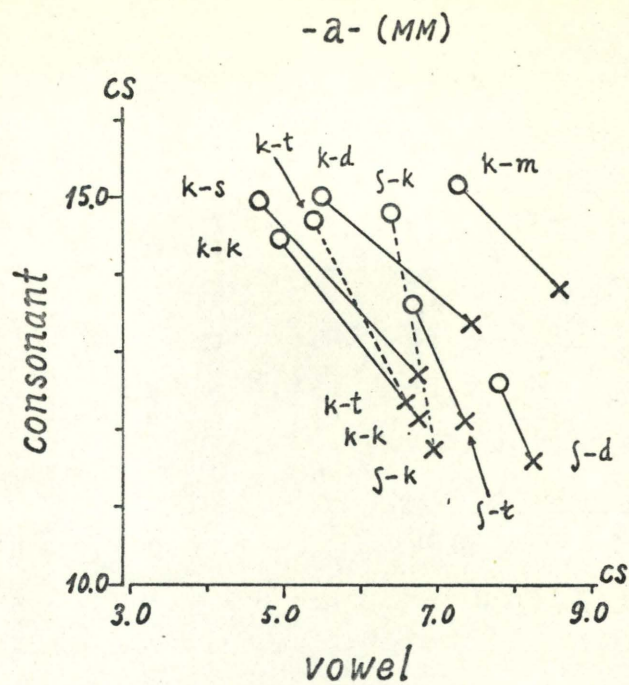


Fig. 3-b.

Average durations of consonant and vowel (MM).
(For symbols, see Fig. 1.)

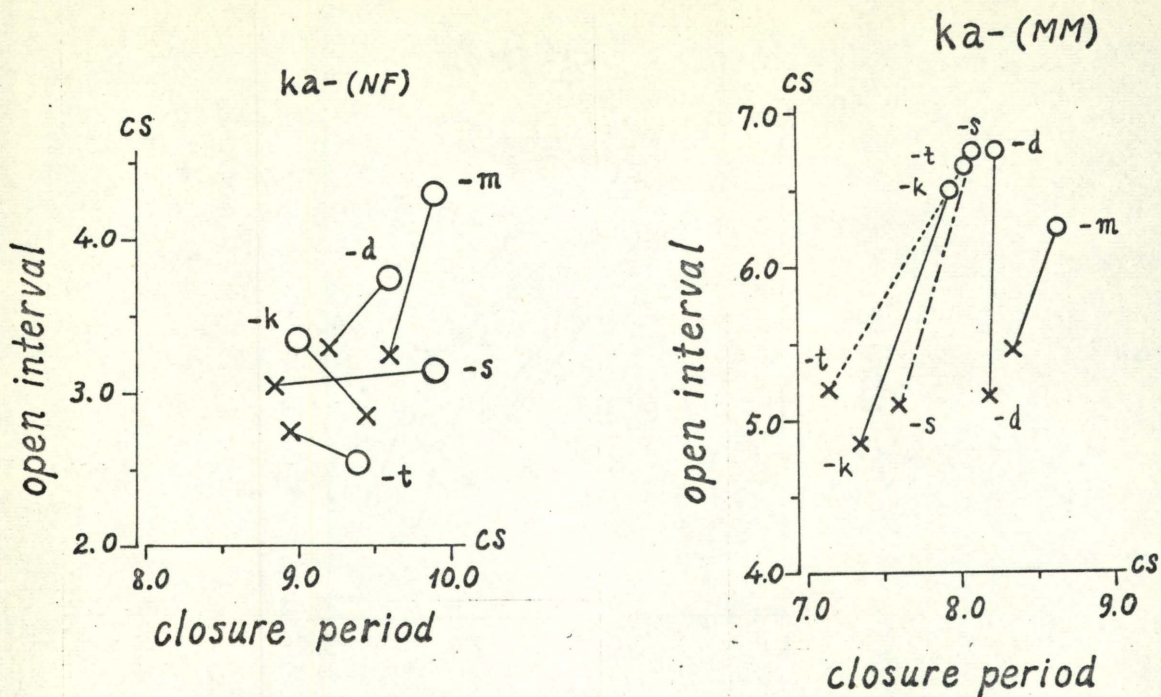


Fig. 4. Average durations of closure period and open interval in *ka*-syllables. (For symbols, see Fig. 1.)

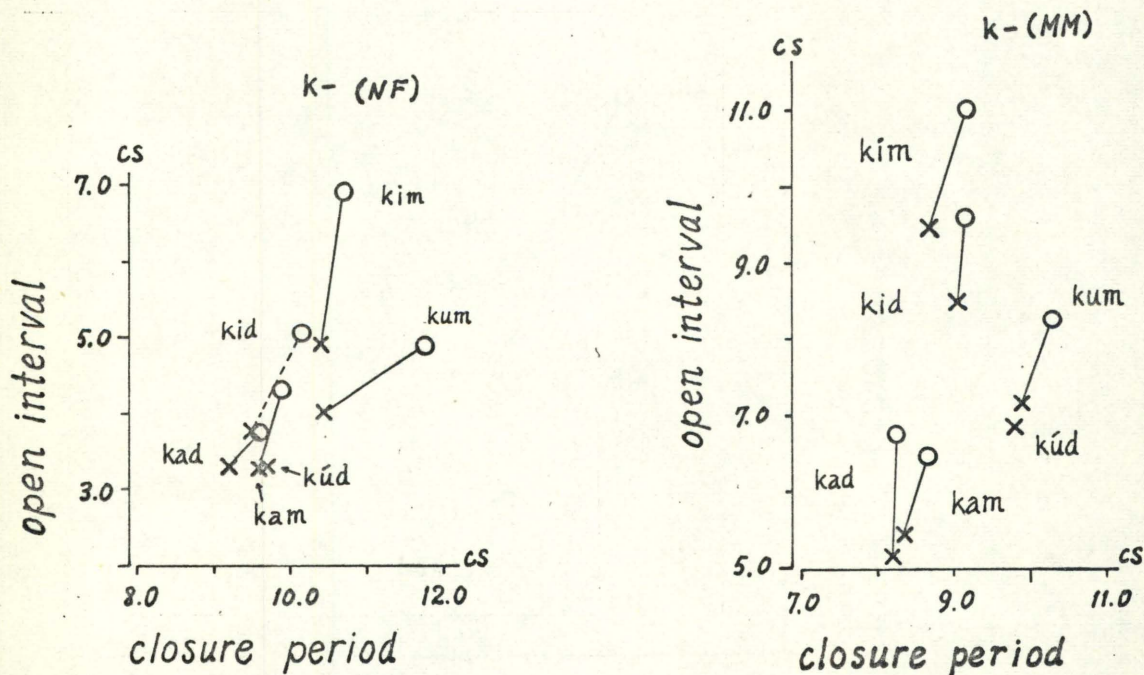


Fig. 5. Average durations of closure period and open interval in *ki*- and *ku*-syllables compared with those in *ka*-syllables. (For symbols, see Fig. 1.)

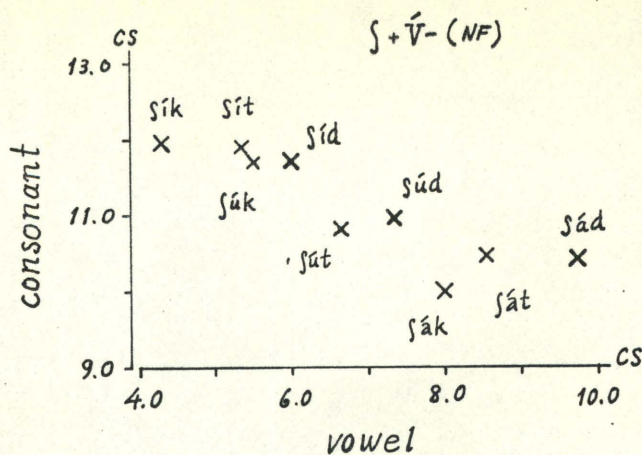


Fig. 6.

Average durations of consonant and vowel in high-pitched \int -syllables (NF). (For symbols, see Fig. 1.)

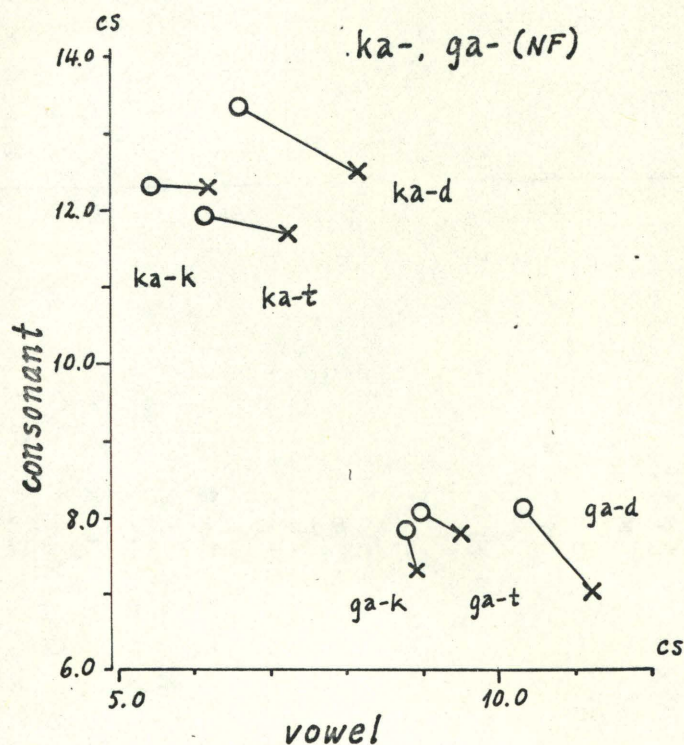


Fig. 7.

Average durations of consonant and vowel in ka - and ga -syllables (NF). (For symbols, see Fig. 1.)

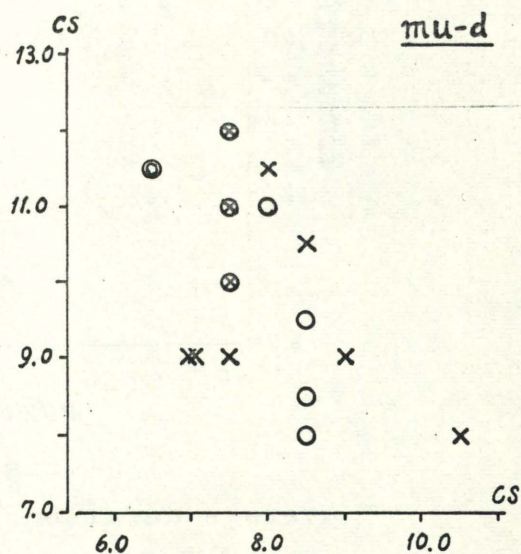
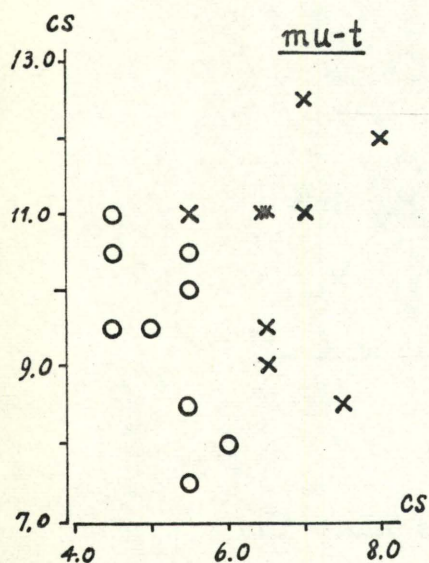
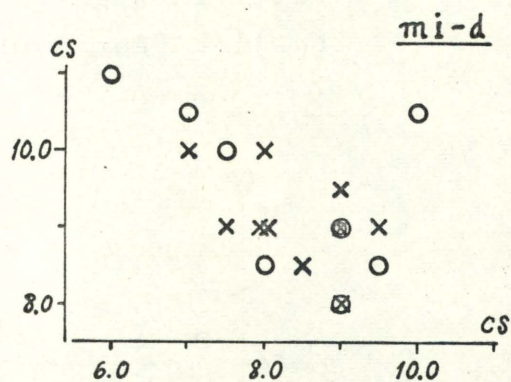
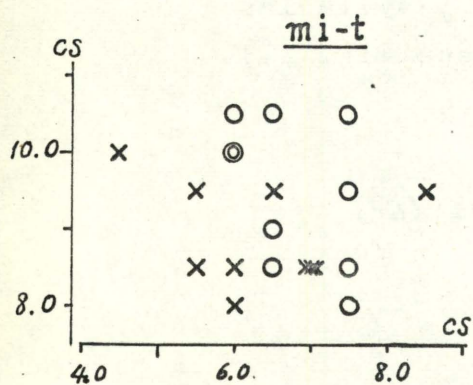
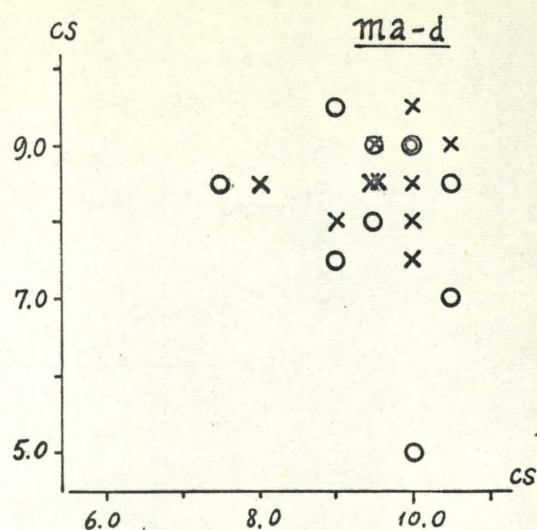
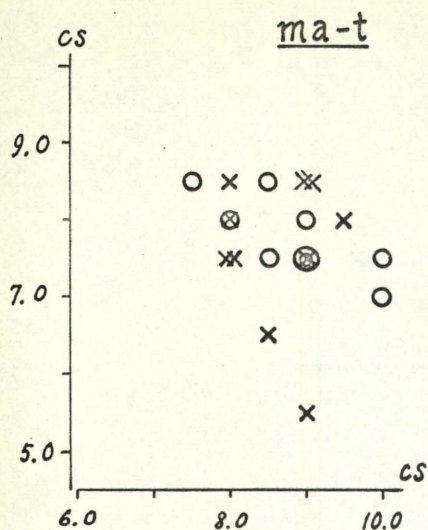


Fig. 8-a. Durations of consonant and vowel in words beginning with m- (NF) (single measurements).
x-axis = vowel, y-axis = consonant.

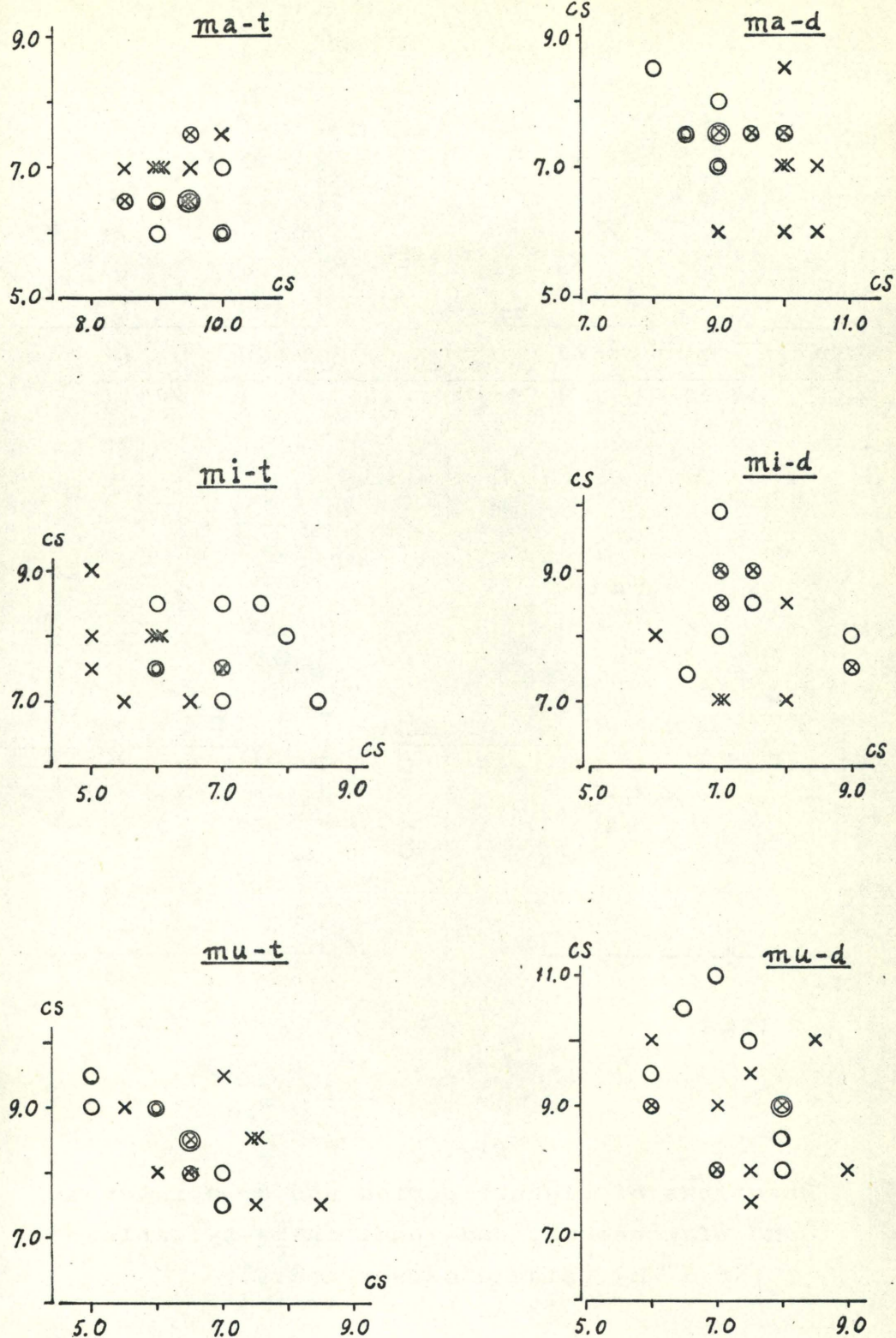


Fig. 8-b. Durations of consonant and vowel in words beginning with m- (MM) (single measurements).

x-axis = vowel, y-axis = consonant.

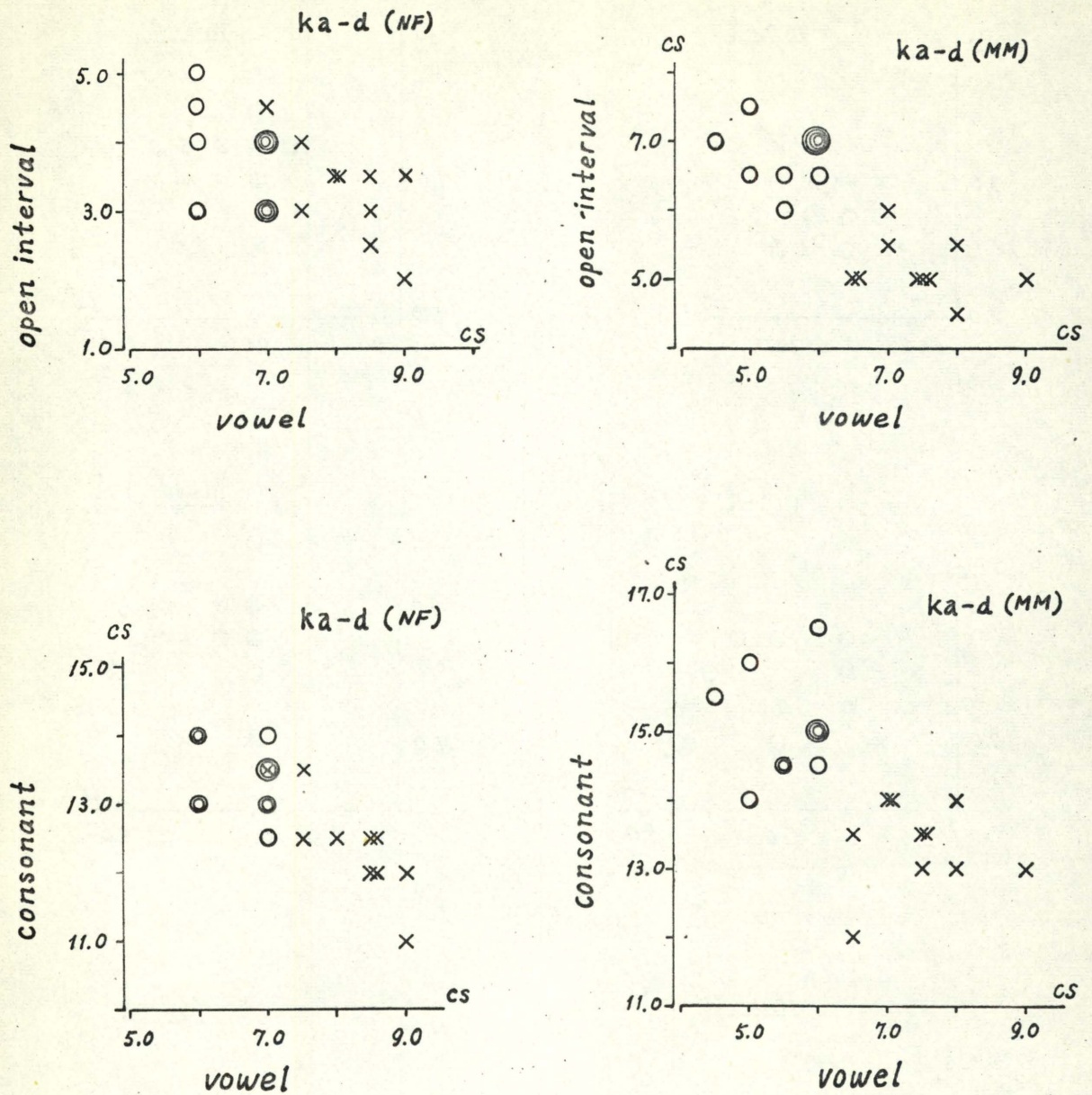


Fig. 9.

Durations of closure period and open interval,
 and of consonant and vowel in *ka*-syllables
 (NF & MM) (single measurements).

sonant to the vowel is greater in the low-pitched syllable. This is due either to the shortening of the vowel (NF's a-syllables, especially Sa-syllables), or to the lengthening of the consonant (MM's words as a whole, and NF's ki's and ku's).

1.1.5. Closure period/open interval (See Figs. 4 and 5)

Now, as was mentioned in 0.1, when the beginning of the vowel becomes devoiced, the devoiced part must appear as the last part of the open interval of the consonant. That is, the open interval must be longer before the low-pitched vowel than before the high-pitched one. As was said in 1.1.2., the length of the open interval is longer in the low-pitched environment, except for NF's ka-t. As for the ratio of the open interval to the closure period, it is greater in the low-pitched environment, except for NF's ka-t, ka-s and ku-m. (The difference between the durations in high- and low-pitched environments is, however, small in ku-m.)

The general tendency is clear: the open interval is longer in the low-pitched environment.

1.2. Syllables with other initial consonants

1.2.1. ga (NF. 10 examples of each word). (See Fig. 7)

The difference in vowel length caused by a pitch difference can also be observed in syllables beginning with g, the voiced counterpart of k. But, in reality, NF's g in word-initial position is almost voiceless. Fully voiced g's are scarce: one instance of gà-d, one of ga-d and one of ga-t. The first half of the closure period is voiced in 3 instances of gà-t, 3 of ga-t, 6 of gà-d, 4 of ga-d, 5 of gà-k and 6 of ga-k. How we get an auditive impression of g is not certain. NF's g has a shorter closure period than k, and the following vowel is longer after g than k.

1.2.2. m-syllables (See Fig. 8)

m+V-syllables do not show a relevant difference of vowel length in different pitch environments. What seems more constant is the total syllable length: when the consonant is shorter, the vowel is longer, and vice versa.

(NB: The low-pitched mu-t is not a two-syllable, but a three-syllable phrase.)

1.3. The influence of surrounding consonants

The vowel a is, of course, much longer than i and u. As for the influence of the following consonant on the vowel, the tendency is as follows. In Sa-syllables, the vowel is longer before d than before t, and is longer before t than before k (this applies both to NF and MM). In ka-syllables the vowel is longest before m, and longer before d than before s, t and k, except NF's low-pitched a, which is on an average 0.1 cs shorter before d than before s. In general, the vowel is shorter before a voiceless consonant (i.e. between voiceless consonants) than before a voiced consonant.

The following consonant is sometimes longer and sometimes shorter (in NF's speech), and is often longer (in MM's speech) after a low-pitched vowel than after a high-pitched one. In many languages the postvocalic consonant is said to influence the length of the preceding vowel more than the prevocalic consonant does. The influence of the postvocalic consonant may also be found in Japanese. This influence may, however, be weaker in open syllables, as Japanese has, than in closed syllables. The problem is not taken up here.

The vowel u is said to be the shortest and i the next shortest of the five Japanese vowels.⁷ But both NF's and MM's u is longer than i. This may be due to the fact that the following consonant is a dental or labial consonant. In the case of k + V, the place of articulation of k may have had some influence, too.

1.4.

A longer open interval in the low-pitched environment is much more clearly seen in isolated words, i.e. in absolutely initial position. But the beginning of the closure phase cannot, of course, be seen on the acoustic curve.

7) N. Torii (6), M. Han (2).

2. Final remarks

From this restricted material I cannot draw any decisive conclusion. The following tendency is, however, clear. After k, ſ (and t, g), the vowel is evidently shorter in low-pitched syllables, but this is not true after m. This fact indicates that the shortening of the vowel is not entirely due to the pitch environment. The longer ſ and the longer open interval of the stop consonant can be interpreted as reflecting the devoiced beginning of the vowel. But at the same time the closure period is also longer in the low-pitched environment, which is against the supposition that the devoiced beginning of the vowel is manifested only in the open interval of the preceding consonant. We cannot tell exactly in which part of the consonant the devoiced part of the vowel should appear. What is more evident is that compensation takes place between the consonant and the vowel in such a manner that the total syllable length is kept almost constant. Thus, it can be said that the vowel is shorter after a voiceless consonant. That is the reason why i and u after and between voiceless consonants become devoiced, especially when they are unaccented.

3. Appendix

3.1. Aspiration and open interval

A stronger aspiration is sometimes perceived when the open interval is longer, but often this is not so. In order to see the relation between the auditory impression of aspiration and the duration of the open interval, a listening test was given to 1 Norwegian and 6 Danish phoneticians. Each of them listened to the tape through ear-phones, and answered questions concerning the degree of aspiration (and affrication); moreover they were asked to make comparisons with k-sounds in other languages. They were allowed to listen to the tape as many times as they wanted. They said that it was sometimes difficult to compare Japanese k with the k-sounds of other languages, since a low-pitched syllable often gave an impression similar to that of a weakly stressed syllable in other languages. If a low-pitched syllable gives such an impression, one would expect Japanese k to be heard as having

a rather strong aspiration compared with k in other languages in the same (or a similar) accentual environment. But this does not appear from the answers. Since each of the listeners judged independently, there is, of course, no standard degree of 'aspiration', so comparisons should be made among k's within each person's judgment.

As for NF's speech, aspiration is seldom heard when the open interval is below some 2.5 or 3.0 cs, except when it precedes a. In some cases a longer open interval is heard as having a stronger aspiration, but just the opposite response is also found with the same example. As for k before i, a stronger aspiration is often heard when the open interval is longer, especially when i is devoiced. The longest open interval in NF's paired words is 7.0 cs. One listener heard the consonant in this case as being more strongly aspirated than the other member of the pair (4.5 cs), but the opposite answer was also given, and one person did not find any aspiration at all in the 7.0 cs example. Nobody recognized a particularly strong aspiration there. - The tendency of judgment is the same as regards MM's speech, though the degree of aspiration was heard as being much stronger than in NF's speech. In MM's speech, a difference in the duration of the open interval of about 3.0 cs did not favour the judgment that the consonant with the longer open interval has stronger aspiration.

From the above test, it seems that the judgment of strong(er) aspiration is not necessarily proportional to longer duration of the open interval. At least in the case of i and u, the judgment of stronger aspiration is more dependent on stronger energy than on a longer duration of the open interval, though the latter factor is of course important, too. It is probable that a longer open interval in the low-pitched environment often reflects an interval of 'voicing lag' in which the energy is not so strong.

3.2. Syllables including a completely devoiced vowel

The tendencies observed for two persons are a little different, but the tendency for a given syllable with a de-

voiced vowel is just parallel to that for the same syllable with a voiced vowel. That is, in NF's speech shortening of the syllable clearly takes place, while in MM's speech lengthening of the consonant is dominant.

Another matter is that the total duration of a syllable consisting of [ʃ] + a voiceless vowel, i.e. [ʃi̥] and [ʃu̥], is often not so much longer and sometimes even shorter than that of [ʃ] occurring before a voiced vowel (see Fig. 2). But we can clearly hear the difference between [ʃi̥] and [ʃu̥]. The auditive difference between [ki̥] and [ku̥] is also clear. I am not sure whether the auditive difference is aroused by different qualities of [ʃ]'s and [k]'s, or by the difference between [ʃ] or [k] + voiceless [i̥] or [u̥]. Both factors may most probably be taken into account.

References

- (1) Eli Fischer-Jørgensen, "Sound Duration and Place of Articulation", Zs.f.Phon.Sprw.u.Komm., Band 17, Heft 2-4 (1964), p. 175-207.
- (2) Mieko Shimizu Han, Japanese Phonology (Tokyo 1962).
- (3) Shiro Hattori, Onseigaku ["Phonetics"] (Tokyo 1951).
- (4) Shiro Hattori, Gengogaku no Hoohee ["Methods in Linguistics"] (Tokyo 1960), the chapter: "Phone, Phoneme, and Compound Phone" (in English), p. 751-763.
- (5) Hideo Mase, "Acoustic Cues for the Perception of Accent in Japanese when the Accented Vowel is Devoiced", ARIPUC 3 (1968), p. 143-176.
- (6) Noriko Torii, "Nihongo Sonaguramu ni tsuite no Jakkan no Koosatsu" ["Some Observations on Japanese Spectrograms"], Denki-Tsuushin Kenkyuuzyo (Tokyo 1957).