

# On the Accent of Japanese from the Phonological Point of View

著者(英)	Shiro HATTORI
翻訳者(英)	Wayne Lawrence
校正者(英)	John H. Haig
journal or publication title	Pioneering Linguistic Works in Japan
page range	1-35
year	2021-02
URL	<a href="http://doi.org/10.15084/00003111">http://doi.org/10.15084/00003111</a>

## On the Accent of Japanese from the Phonological Point of View

HATTORI Shirō

### Contents

- 1 Phonetic description and phonological interpretation
  - 1 Phones, phonemes
  - 2 Accent
- 2 Procedure for phonological interpretation, various operational principles
  - 1 Phones, phonemes
  - 2 Accent
- 3 Distinctive features and non-distinctive features
  - 1 Phones, phonemes
  - 2 "Accent kernel"
- 4 Kōichi Miyata's theory on Tokyo dialect accentuation
- 5 Yale University's Prof. Bloch's theory on Tokyo dialect accentuation
- 6 Pronunciation and the phonological system
- 7 Investigation into and definition of terminology related to accent
  - 1 Terms used in my book Phonetics
  - 2 "The prosodeme"
    1. Tokyo dialect, Tosa dialect
    2. Sendai dialect, Yamatohama dialect
    3. Prosodemic characteristics and points in common between the two dialect types
  - 3 Stress accent languages and "prosodemes"
  - 4 Conclusion

Notes 1-27

**1.1** As phonetic research into languages progressed, the number of cases where what was thought to be the same sound turned out to be somewhat different sounds becomes more frequent, and phonetic observation and description become more and more detailed. However, researchers started to feel that such detailed descriptions did not always correspond well with speakers' intuitions regarding their pronunciations, and this was one trigger which led to the development of the field of phonology. In brief, it became clear that there were not a few instances where what were objectively different sounds actually could be regarded as corresponding to the same phoneme.

However, as phonology develops, it is becoming clear that, to a greater or lesser extent, in phonetic descriptions, no matter how detailed they may be, there was always a phonological viewpoint which had unconsciously been introduced. I especially want to emphasise this point, as it is not sufficiently recognised by everyone.

For example, the “consonant of the *s*-line” in the Tokyo and Kyoto dialects is [sa ʃi su se so] in the most detailed representation (vowels are in broad transcription). It is correct that the consonant before the vowel *i* is markedly different from the other consonants, but because the consonants before the other four vowels are, strictly speaking, all slightly different from each other, it is inaccurate to represent them all using [s]. Phonologically these syllables are interpreted as /sa si su se so/, so representing the four types of *s* above with the same symbol [s] can be considered to be a case where, in a manner of speaking, a phonological viewpoint has dominated unconsciously even before phonology as a field existed. In other words, these consonants are viewed intuitively as the same consonants without adequate phonological analysis, and this result is phonologically correct.

Another example is the words ‘mosquito’, ‘paddy field’ and ‘name’ in the Tokyo and other dialects, which are transcribed as [ka], [ta] and [na] respectively when pronounced with the vowels cut short. If written in a narrower transcription they can be transcribed as [k<sup>h</sup>a], [t<sup>h</sup>a] and [nã] (the nasalisation on the vowel in ‘name’ is very light). However, in actuality in such a careful pronunciation, the vowel always ends with a glottal stop, so more precisely the words should be written as [k<sup>h</sup>aʔ], [t<sup>h</sup>aʔ] and [nãʔ]. However, this glottal stop never fails to appear with short vowels pronounced as above, so it is phonologically without significance. One could even say that the glottal stop is a way of cutting the vowel short. Taking this, as well as the phonological system as a whole, into consideration, the above words can be interpreted phonologically as /ka/, /ta/ and /na/. Even phoneticians, unless paying particular attention, may not notice the presence of the glottal stop. This is the reason why the glottal stop is not transcribed.

Generally speaking, phonetic transcription using a phonetic alphabet, no matter how precise it may be, cannot completely represent sound. To some extent it is a simplification which incorporates phonological elements. When it comes to a broad transcription, we can say that it is extremely simplified due to a phonological approach. One could possibly even go so far as to say, ironically, that the difference between a broad transcription and a completely phonological transcription is that the former incorporates a degree of deception. In phonology we start with a phonetic transcription which is as narrow as possible, and then, using that, we perform our phonological analysis, but actually that is what we do with as yet unclear areas of the phonological system of a language (strictly speaking, a homogeneous dialect), but with areas about which we have a good idea, we usually almost unconsciously use a phonological transcription. One could go so far as to say that phonetic transcriptions are, to some degree, phonological.

Phonetic description is also so, but to some extent it was normal for a phonological viewpoint to dominate unconsciously even before phonology existed as a field. Let us summarise this as “phonetic description contains many parts which are unconsciously phonological”.

**1.2** With very few exceptions, descriptions and transcriptions of Japanese accent have been unconsciously half phonetic and half phonological. Many do not distinguish between a phonetic approach and a phonological approach. (Here “accent” is used in the sense it is used in in Dr Kanae Sakuma’s examples, and does not refer to the “accent peak”.)

For example, the accent of Tokyo Japanese *āsāgao* ‘morning glory’ and *sorā mame* ‘broad bean’ is described in a number of ways, such as Low-High-Mid-Mid, L-H-L-L-, M-H-M-M, M-H-M-L.<sup>a)</sup> Why do scholars have such a range of opinions? One reason is that these opinions are semi-phonetic. Vocal pitch as an objective phonetic entity is complicated, and cannot be accurately described in terms of a 3-step scale such as H, M, and L. When pronounced in isolation, the fourth syllable of these words is pronounced on a lower pitch than the third syllable, and is much lower than the first. For this reason, of the descriptions given above, M-H-M-L is phonetically most accurate, but even this cannot be said to be sufficiently correct. On the other hand, although the pitch continues to fall from the third syllable to the fourth, giving these two syllables as M-M or L-L is mixing in a phonological interpretation. In other words, the fall in pitch contour (in this paper “pitch contour” is used with the meaning of “changes or lack of change in the voice’s pitch”) observed here is phonologically without significance, so this kind of interpretation is possible. Expanding a little on this point that this fall in pitch contour is phonologically without significance, according to the socially conventionalized pitch contours on words in the Tokyo dialect, the distinctions between M-H-M-L and M-H-M-M or M-H-L-L do not exist in the same phonological environment. In other words, the fall M-L in M-H-M-L never establishes a phonological contrast in this dialect.

Even so, in the above four different views of the pitch of *asagao* etc., there are points in common. All of them recognise that the second syllable is higher than the adjacent syllables, and we can surmise that this fact is phonologically significant.

**2.1** In order to carry out phonological analysis, we first investigate the distribution of individual sounds (phones). Phones themselves are actually a notion which has the phoneme as its backing, and so when we talk of phones we are often already unconsciously basing this on the phoneme. Strictly speaking this is not permitted, but if we are sufficiently aware of this while we operate, we can probably arrive at satisfactory results. However, we must be careful not to fall into the mechanistic trap of believing that phones can be objectively specified purely phonetically.<sup>1</sup>

Investigating the distribution of these phones, we answer such questions as which phones can appear in the “same environment” and establish a “phonological contrast”, and which phones display a “complementary distribution” and do not contrast phonologically. Phones which can contrast phonologically are different phonemes. On the other hand, phones which are in complementary distribution are possibly the same phoneme, but this is not always the case.

In order to inductively arrive at the phoneme system, a partial investigation will not suffice. Any inquiry must be based on a broad range of materials. When operating in this way, the “operational principle of the system”, the “operational principle of economy”, the “operational principle of assimilation” etc. are powerful tools. I briefly explained these operational principles in my 1953 paper “Kokugo-no on’in-taikei-to shin-nihonshiki-rōmaji-no tsuzurikata” [The phonemic system of Japanese and the New Nippon System of Romanised spelling] (*Kyōiku Gijutsu* 8(4): 94-102).

The “operational principle of the system” is the prospect that phonemes are distributed in a very systematic, balanced manner. Carrying out research based on this principle is proving to be advantageous.

The “operational principle of economy” is the prospect that the fewer phonemes which are postulated, the better. However, I have repeatedly explained that it is dangerous to follow this principle blindly.

The “operational principle of assimilation” concerns the case where two (or more) phones in complementary distribution can be explained as the same sound assimilating to an adjacent sound to take on a different form – in such a case the phones may be recognised as belonging to the same phoneme. However, it may be preferable to call this the “operational principle of environment assimilation”, so from now on I shall use this name.<sup>2</sup> For example, Tokyo and Kyoto dialects’ *shi* can be transcribed phonetically as [ʃi], but this [ʃ] can be explained as being /s/ which has assimilated to the following /i/, so phonologically this [ʃi] is interpreted as /si/.

**2.2** The co-called “accent patterns” of dialects such as those of Tokyo and Kyoto are widely recognised as being one component of the word form, and thus can form phonological contrasts. For this reason, they are grouped together with segmental phonemes under the generic term “phoneme”.

Strictly speaking, when inductively arriving at phonological accent patterns too, the same procedures used when inductively arriving at the segmental phonemes must be used. In actual research, however, we often observe methods which are very unsophisticated but are nevertheless sufficiently successful.

The distribution of accent patterns is not usually a problem. Looking at just one dialect, there are a small number of accent patterns and the patterns themselves can usually be classified extremely systematically so there was no need to be aware of the operational principles of system, economy, etc.<sup>3</sup>

The same goes for the “operational principle of environment assimilation”, but, by way of precaution, I give an example of how it applies. In the Tokyo dialect, not a small number of speakers pronounce the following words with the following pitch contours.

- (a) [komaŋjirē] 'finely chopped'  
       [komaŋonā] 'in fragments'
- (b) [ko:baN̄] 'police box'  
       [kondaN̄] 'friendly discussion'

The pitch contour on the (a) words appears when the first two moras are /CVCV/ or disyllabic, and the (b) pitch contour appears when the first two moras are monosyllabic /CVV/ or /CVN/. (Syllable in this case is a phonological syllable. In addition to the phonological unit of the mora, I recognise the phonological syllable. In the Tokyo and Kyoto dialects, the mora and phonological syllable are usually the same, but as in the examples above, there are instances where they do not correspond.<sup>4</sup>) These two pitch contours are therefore in complementary distribution, and moreover the (b) pitch contour can be explained as emerging as a result of the rising contour being avoided on a syllable of the shape /CVV.../ or /CVN.../, so the (a) pattern and the (b) pattern can be recognised as being phonologically identical.

Also, in the Tokyo dialect there are the following examples where the vowel in the first syllable is usually devoiced.

- [k̥ita] 'came'        [k̥ita] 'wore'

Compare these with the following:

- [mīta] 'saw'        [mīta] 'Mita (place name)'

The vowel [a] in 'Mita' and the vowel [a] in 'wore' have the same pitch contour, but the [a] of 'saw' and the [a] of 'came' have markedly different pitch contours. In *mita* 'saw', the *mi* is high and the *ta* is low. In the word *kita* 'came', there are some people who pronounce the *ta* high, but there are also people who pronounce the *ki* strong and the *ta* low. Here I will analyse the latter pronunciation. In this pronunciation too, the *ta* of *kita* 'came' has a markedly different pitch contour from the *ta* of *mita* 'saw'. Whereas the [a] in *kita* 'came' has a marked fall in pitch on it, the [a] of *mita* 'saw' does not have such a marked falling contour. The pitch contour observed in *kita* 'came' is only found when the vowel in the initial syllable is devoiced, and moreover moreover, different pitch contours are observed on 'saw' and 'Mita' and on 'came' and 'wore', respectively. It follows from these observations that the pitch contour on *mita* 'saw' and that on *kita* 'came' are in complementary distribution. Because we can thus attribute the particular pitch contour in *kita* to the devoiced vowel in the first syllable, *mita* 'saw' and *kita* 'came' have the same accent pattern phonologically, and both have an accent peak (as will explain below, more accurately this should be called the "accent kernel") on the initial syllable. The fact that in both verbs the first syllable is stronger than the second also supports this hypothesis. This interpretation also corresponds to the speakers' intuition.<sup>5</sup>

**3.1** Each individual phone has a range of phonetic features. One subset of these are features which are dependent on the environment, and the phones are unrelated to the corresponding phoneme. Consequently, they are of course not distinctive features of the phoneme. (In my aforementioned paper in *Kyōiku Gijutsu* I called distinctive features “distinctive characteremes”.)

For example, in Tokyo and Kyoto Japanese, the initial consonant in [ʃi] /si/ is palatalised. However this is because the consonant precedes /i/ and assimilated to that, and this process is unrelated to the phoneme /s/ itself.

Phonemes may also possess a range of phonetic features, but not all of them are necessarily distinctive features.<sup>6</sup> For example, the phoneme /m/ in the Tokyo and Kyoto dialects is a voiced labial nasal. However, in these dialects nasal phonemes do not phonologically contrast voiced/voiceless. Thus, for /m/, voicedness is not a distinctive feature, but is instead a “non-distinctive feature”. In other words, it goes without saying that labial nasal phonemes in these dialects are normally voiced. I shall refer to non-distinctive features of a particular phoneme as “determined features”.

In order to describe the phonemes of a certain language (dialect) phonologically, it is sufficient to list the distinctive features. However, in order to mimic the pronunciation of that language (dialect), it is not enough to pronounce only the distinctive features and ignore the non-distinctive features. For example, when mimicking the Tokyo dialect, it would not be acceptable to pronounce all nasals as voiceless, and although [ʃi] is /si/, pronouncing the initial consonant as [s], without palatalization, would sound odd. In order to perfectly mimic a language (dialect), one must pronounce in such a way as to produce all the phonetic features perfectly. This means that non-distinctive features are also necessary to produce the complete characterisation of a language’s phonetics. Linguists who say that non-distinctive features are unnecessary for the phonetics of a particular language have a biased view which will lead to misunderstandings.

For this reason, it is necessary to describe the non-distinctive features of phonemes in order to give an overall description of the phonetics of a language (dialect), and basing this on the description of the phoneme is a much more systematic and efficient way of doing this. Here lies another reason for the need for phonology in addition to phonetics.

**3.2** Environmentally determined features are part of the pitch contour of a particular word (or fragment corresponding to a word) in a particular utterance. Such features are not distinctive features of the word’s accent pattern. For example, in the Tokyo dialect, when the phrases *kono kodomo* ‘this child’ and *kono karada* ‘this body’ are pronounced in one breath, as long as *kodomo* ‘child’ and *karada* ‘body’ are not particularly emphasised, they are pronounced very close to [konokodomō], [konokarada], and not as [konokodomo], [konokarada]. The preceding and following syllables are high, and this causes the pitch of the syllables [ko] and [ka] to rise in these phrases. When *kodomo* and *karada* are pronounced in isolation, they become [kodomo], [karada]. The accentual

features of these words in the phrases [konokodomo], [konokarada] are features due to the environment.

The pitch contour pattern of [ko:baN] 'police box' and [kondaN] 'friendly discussion' given above is also due to the environment, and does not contrast phonologically with the pitch contour pattern of [komaŋire] 'finely chopped' and [konaŋona] 'in fragments'.

Accent patterns themselves may possess various phonetic features, but not all of these are necessarily distinctive features. For example, in the Tokyo dialect the word *karakasa* 'oil-paper umbrella' is pronounced as [kara:kasa] in isolation. However, in this dialect pitch contours such as [○○○○] and [○○○

○] never contrast phonologically with [○○○○] in the same environment, so the fact that in [karakasa] the first *ka* is lower than the *ra*, and *ra* is the same pitch as the second *ka* is, as far as the Tokyo dialect is concerned, a "determined

feature", and consequently this part of the pattern [○○○○] is a non-distinctive feature. The distinctive feature of this pattern is only that the third mora is high and the fourth mora is low (and thus the third mora is pronounced with somewhat more intensity than the fourth mora). The reason for this is that this

pattern contrasts phonologically with [○○○○] and [○○○○] (the L-H-H-H pattern, distinct from the L-M-M-M pattern) in the same phonological environment, and this difference between these patterns lies in the location of the "high" tone which accompanies (or may accompany) a "low" tone. That the initial mora is low and the second high is common to all three patterns, and is thus non-distinctive.

In order to phonologically describe the accent pattern of a particular language (dialect), it is only necessary to list the distinctive features. In the case

of the Tokyo dialect, in the word [kara:kasa] the high tone on the second *ka* accompanies (or may accompany) a following low tone, and if this high tone<sup>7</sup> is called the "accent kernel", the distinctive features of the accent pattern of this dialect are:

- 1) Is there an accent kernel or not?
- 2) If there is an accent kernel, which mora is it located on?

Words without an accent kernel have the so-called flat (or unaccented) pattern, and those with an accent kernel have a rising and falling pattern. For example, the accent patterns of this dialect can be transcribed as follows.<sup>8</sup> (/ ˀ / indicates that this position is immediately preceded by an accent.)

[çi] `day`	/hi/
[çi] `fire`	/hi˧/
[ūji] `cow`	/ˈusi/
[ūma] `horse`	/ˈuma˧/
[neko] `cat`	/ne˧ko/
[kārada] `body`	/karada/
[atama] `head`	/ˈatama˧/
[kokōro] `mind`	/koko˧ro/
[īnot̄ji] `life`	/i˧˧noci/

However, when mimicking the accentuation of the Tokyo dialect, it is not sufficient to pronounce only the distinctive features and completely ignore the non-distinctive features of accentuation. For example, pronouncing *atama-ga*

[atamāga] `head-SUBJ` as [atamāga] or [atamāga] would sound strange. In order to perfectly mimic the accentuation of a particular language (dialect), the pronunciation must be so that all of the features are completely realised. This means that non-distinctive features too are indispensable in producing a complete characterisation of a language's accentuation.<sup>9</sup>

For this reason it is necessary to describe the non-distinctive features which accompany accentuation in order to give an overall description of the accentuation of a language (dialect), and basing it on the phonological description is a much more systematic and efficient way of doing this. In terms of Japanese accentuation, ideally the description should not only be of pitch, but should also include intensity.

Incidentally, the distinctive features for the accentuation of dialects like those of Kyoto and Kameyama are the following (see *Phonetics* p.193).<sup>10</sup>

- 1) Does the word begin high or low?
- 2) Is there an accent kernel or not?
- 3) If there is an accent kernel, which mora is it located on?

In these dialects, the accent pattern of a word like *suzume* `sparrow` can be

roughly transcribed as [suzumē], but its distinctive features are "low-beginning and no accent kernel". It has been debated whether this pattern is not pronounced as L-L-H but rather as L-M-H or L-L-M, but this is a debate about the phonetics, and is irrelevant to the phonology. As with the debate referred to above (§1.2), there is a danger that it will turn into a fruitless argument.

4 On the subject of Tokyo dialect accentuation, some may wonder if there have been no transcriptions or descriptions which are similar to the above phonological transcription and description. The answer is that there have been. In the distant past, Bimyō Yamada's transcription using 'mid', 'high', 'all level', 'first high', 'second high'... in his *Nihon Dai-jisho* [Grand dictionary of Japan] (1892) is equivalent to our phonological transcription, as is the more recent ①②③... in Kyōsuke Kindaichi's *Meikai Kokugo Jiten* [Meikai Japanese dictionary] (1943).

When an academic field becomes very developed, there are cases which make us wonder whether we have returned to an undeveloped, naive way of thinking. In linguistics too, as a result of the development of phonetics and then phonology, there is a strong trend towards a phonological transcription which is close to the naive phonetic transcription of a time before phonetics rather than a complex phonetic description.

On the other hand, with regard to the accent patterns of the Tokyo dialect, there is a scholar who, through sharp intuition, reached a conclusion which corresponds with part of my aforementioned conclusion reached through phonological analysis. This is Kōichi Miyata. In his papers "A new view of accentuation and the annotation of accentuation" (*Onsei-no Kenkyū* [The Study of Sounds] 1 (1927)) and "My opinion concerning Japanese accentuation" (*Onsei-no Kenkyū* [The Study of Sounds] 2 (1928)), he calls something very close to what I have been calling the "accent kernel" the "accent core". He writes, about the [taraʃi] part of the Tokyo dialect word [a<sup>̄</sup>taraʃi:] 'new':

"The grouping of three moras cannot be viewed as the accent core. This is because even if we pronounce this word as [ataraʃi:], [ataraʃi:] or [ataraʃi:], [.....] we do not feel that they are fundamentally different accent patterns. From this we can think that the accent core is not on the three-mora sequence of the above forms, but is the change in pitch — the fall — in the transition from the fourth mora to the fifth mora. This point probably becomes clearer when we try pronouncing the same word as [ataraʃi:], [ataraʃi:] or [ataraʃi:] — not only do we feel that these are completely different accent patterns from [ataraʃi:], these different pronunciations are all recognisable as independent accent patterns."  
(*Onsei-no Kenkyū* [The Study of Sounds] 2 (p.32))

This account is somewhat subjective, but in fact it is an attempt to state the same as part of the phonological analysis given above. The only point of

difference is that, although the pronunciations [ataraʃi:], [ataraʃi:] etc. are not felt to be "fundamentally different accent patterns", I would say that Tokyo dialect speakers would find them odd. While recognising Miyata's theory as magnificent, the fact that it was not accepted in its entirety is for that reason. A more persuasive explanation was required. In this way, one reason why we

could not agree with Miyake's theory was because he insisted on a mentalistic phonology based on Prof. Arisaka's "phoneme" (defined as "abstract target for pronunciation motor activity"), but there were other minor faults in his theory.<sup>11</sup>

The view of accent which I proposed above is a view which can resolve the stand-off between the three-pitch view and the two-pitch view on the one hand and Miyata's theory on the other, but I wish to emphasise that it is neither a theory which sides with either of these nor a third theory which competes with them.<sup>12</sup>

However, it is necessary to give sufficient recognition to the immense value of Miyata's achievement of having, a quarter of a century ago, produced an antithesis to the three-pitch view etc. which were influential at the time.

**5** It would probably not be out of place to introduce here the theory of Dr Bernard Bloch (professor of Linguistics at Yale University, U.S.A.) regarding the accentuation of "the speech of educated persons native to Tokyo".<sup>13</sup>

Firstly, stating that "[f]or a complete account of the observable pitch variations in Japanese, four phonemically different levels are necessary and sufficient", he uses the following numbers to denote "pitch phonemes".

- /1/ highest pitch
- /2/ higher mid pitch
- /3/ lower mid pitch
- /4/ lowest pitch

In the phrase [kimonoŋayoŋoreta] 'the clothes became dirty', Bloch gives [ki] as /4/, [monoŋayoŋore] as /3/, and [ta] as /4/, and in the phrase [aoi] 'blue', [a] is

/4/, [o] is /2/, and [i] is /4/. In [kimonoŋayoŋoreta], the overlined part is not on a perfectly level pitch — strictly speaking, it actually gradually falls while wavering. However, there is not a marked decrease in pitch between [re] and [ta], and there is no accent kernel in this sequence of words. It is unjustified to say that the string from [mo] to [re] has the "pitch phoneme" /3/ (underlining added by Hattori), and [ta] has the "pitch phoneme" /4/. But this is not such a big mistake. What is important is whether it is valid to interpret each mora as being associated with one of four types of "pitch phonemes". Under this interpretation, the same word [yoŋoreta] 'became dirty' is /3333/ in the sequence [kimonoŋayoŋoreta], but has the "pitch phonemes" /4333/ when [yoŋoreta] in isolation. However, as shown above, whether the first mora and the second mora of this word are pronounced on a level pitch or whether there is a rise is determined purely by the environment, and the difference in pitch cannot form a phonological contrast. A "phonological" interpretation where such non-distinctive features are described as relating to distinguishing "phonemes" which are phonologically significant must be said to be unjustified.

According to Prof. Bloch, the sequence [soodes'ka] expresses different meanings depending on the "pitch phonemes" associated with each mora as follows:

/2, 3, 3, - , 4/	Is that so?
/2, 3, 3, - , 31/	Is that really so?
/1, 3, 3, - , 4/	Oh, so that's it?
/1, 3, 3, - , 31/	Oh, so that's it?

From our viewpoint, such differences in “pitch fluctuation” are differences in sentence intonation, and the common feature of the above four examples — “the first mora is high and the pitch on the second mora is lower” — alone is phonologically distinctive. In other words, the form (in this case, word combination) [soodes`ka] has an accent kernel on the first mora. The same form can take various sentence intonations, and the same sentence intonation can associate with a variety of forms. It can be said that Bloch’s theory as yet is unable to adequately analyse accent and sentence intonation.<sup>14</sup> This is immediately clear from the following type of example (p.119).

${}_4\text{mu}_3\text{ka}\dot{\text{s}}\text{i}\text{muka}_2\dot{\text{s}}\text{i} \text{ } {}_2\text{a}_3\text{rutokoro}_2\text{ni}$

Our view is that the rising pitch contour at the end of these phrases is not related to accent.

We can conclude that Bloch’s description is not phonological; it is, rather, phonetic.

**6** With regard to phonemes, it is possible for two dialects to have the same phoneme system yet have different pronunciations. For example, Tokyo dialect /su/ and Kameyama /su/ differ. The vowel in the former is a central vowel, whereas the vowel in the latter is pronounced further back. In spite of this difference, both dialects have the same system in the “s-line”.

/sa si su se so s̄ja s̄ju s̄jo/

Among the dialects spoken in Kyushu and elsewhere there are some where /se/ is pronounced as [ʃe]. In spite of this, the system of the “s-line” in these dialects is the same as that of Tokyo, and so it can be thought that [ʃe] corresponds to /se/.

In the dialect of Sada in Shimo-Kitayama village, Yoshino county, Nara prefecture, the accentuation is as follows (according to the pronunciation of Toshio Uenishi).<sup>15</sup> (Bold overlines indicate a higher and more intense pronunciation than the fine lines.)

'cow'	[u $\bar{f}$ i ]	[ $\bar{u}$ f $\bar{i}$ g $\bar{a}$ ]
'dog'	[i $\bar{n}$ u]	[i $\bar{n}$ u $\bar{g}$ a]
'trace'	[ $\bar{a}$ to]	[ $\bar{a}$ t $\bar{o}$ g $\bar{a}$ ]
'child'	[ $\bar{k}$ o $\bar{d}$ o $\bar{m}$ o]	[ $\bar{k}$ o $\bar{d}$ o $\bar{m}$ o $\bar{g}$ a]
'word'	[ $\bar{k}$ o $\bar{t}$ o $\bar{b}$ a]	[ $\bar{k}$ o $\bar{t}$ o $\bar{b}$ a $\bar{g}$ a]
'jellyfish'	[ku $\bar{r}$ a $\bar{g}$ e]	[ku $\bar{r}$ a $\bar{g}$ e $\bar{g}$ a]
'helmet'	[ $\bar{k}$ a $\bar{b}$ u $\bar{t}$ o]	[ $\bar{k}$ a $\bar{b}$ u $\bar{t}$ o $\bar{g}$ a]

The accent peak on the first syllable of the forms [u $\bar{f}$ iga], [kodomo], [kodomoga], [kotoba], [kotobaga] etc. does not contrast phonologically, and so it is a non-distinctive feature of these patterns. From this the accent patterns of these words can be transcribed phonologically as the following.

'cow' 'dog' 'trace' 'child' 'word' 'jellyfish' 'helmet'  
 /'u $\bar{f}$ i/ /'i $\bar{n}$ u/ /'a $\bar{t}$ o/ / $\bar{k}$ o $\bar{d}$ o $\bar{m}$ o/ / $\bar{k}$ o $\bar{t}$ o $\bar{b}$ a/ / $\bar{k}$ u $\bar{r}$ a $\bar{g}$ e/ / $\bar{k}$ a $\bar{b}$ u $\bar{t}$ o/

In other words, despite the marked difference in the pitch contours, this dialect's accent system, as far as these forms are concerned, is identical to that of the Tokyo dialect.

The Kyoto dialect accent pattern [ $\bar{\circ}\bar{\circ}$ ], [ $\bar{\circ}\bar{\circ}\bar{\circ}$ ], [ $\bar{\circ}\bar{\circ}\bar{\circ}\bar{\circ}$ ]...

corresponds to the accent pattern [ $\bar{\circ}\bar{\circ}$ ], [ $\bar{\circ}\bar{\circ}\bar{\circ}$ ], [ $\bar{\circ}\bar{\circ}\bar{\circ}\bar{\circ}$ ]... in the Tosa dialect.<sup>b)</sup> (Of course the transcription used here is broad.) Although the pitch contours differ, this pattern of the Tosa dialect, like those of the Kyoto dialect, are described phonologically as "low beginning and no accent kernel".

It is of the utmost necessity that the classification of Japanese dialects' accentuation be revisited in the light of this kind of phonological approach. Also, a phonological approach must also be taken to historical accentual changes. For example, in the Kyoto dialect the following changes took place between the 17<sup>th</sup> and 19<sup>th</sup> centuries.

[ $\bar{\circ}\bar{\circ}\bar{\circ}$ ] > [ $\bar{\circ}\bar{\circ}\bar{\circ}$ ]  
 [ $\bar{\circ}\bar{\circ}\bar{\circ}\bar{\circ}$ ] > [ $\bar{\circ}\bar{\circ}\bar{\circ}\bar{\circ}$ ]

This is a phonetic change, and not a phonological change. However, in the following change the accent kernel has moved resulting in the merger of patterns, so this is a phonological change'.<sup>16</sup>

$$[\overline{\text{○○○}}] > [\overline{\text{○}}\text{○○}]$$

$$[\overline{\text{○}}\text{○○}] > [\overline{\text{○}}\text{○○}]$$

**7.1** Finally, in this section I will discuss the functions of “accent” and investigate the notions expressed by this term.

It is widely known that the term “accent” is used with a number of meanings in Japanese. In one usage of the word, the Tokyo Japanese word

[ka<sup>̄</sup>ra<sup>̄</sup>da] ‘body’ has no accent, and [i<sup>̄</sup>no<sup>̄</sup>tʃi] ‘life’ has an accent on the first syllable. In this case, phoneticians usually use the term “accent peak” instead of

“accent”. Phoneticians usually say that the accent of [ka<sup>̄</sup>ra<sup>̄</sup>da] ‘body’ is flat, or

that [ka<sup>̄</sup>ra<sup>̄</sup>da] has a flat-pattern accent. As I have already warned, I have been using the term “accent” with this latter meaning. However, even with this meaning it is necessary to further examine the term.

In §9.2 of my book *Phonetics* I defined “accent” as “a pattern of intensity or pitch of a stress group established by social convention”, and I divided it into “stress accent” and “pitch accent”. However, as explained above, in a narrow phonetic transcription, in the case of “stress accent” a description of pitch and length is also necessary, and in the case of “pitch accent” a description of intensity and length is also required. Also, there is always one accent peak in a “stress group”, and that peak is almost always pronounced high, and when it is pronounced with somewhat increased intensity there may be cases where one cannot determine whether it is a “stress accent” or a “pitch accent”. In such cases, there is no need to decide either way.

In my book I defined “pitch accent” as “the phenomenon of the pitch contour pattern on a stress group or syllable being established by social convention (or the highest point (peak) of that pitch)”. The reason I added “the highest point of that pitch” in parentheses was out of consideration of the aforementioned idiomatic usage in Japan of the word “accent”. When defining “accent” according to the normal academic usage among Japanese phoneticians, the parenthesised portion should be omitted.

I further divided pitch accent defined as such into “syllable pitch accent” and “word pitch accent”. The former refers to the likes of the four tones of Mandarin Chinese, and the latter refers to the phenomenon observed in Tokyo and other Japanese dialects. This paper is predominantly concerned with the latter type. However, because these two types are qualitatively rather different, if the latter is to be called “accent”, it is desirable that the former be called by another term. In all languages, linguistic forms (such as words) are not a string of individual segmental phonemes of equal importance strung together in a line; some of them group together to form syllables. For example, Japanese (Tokyo dialect) /karada/ ‘body’ is a sequence of the phonemes /k/, /a/, /r/..., but these are not elements of equal importance lined up. Instead they are grouped into syllables of two phonemes each to form the phonological syllables /ka/, /ra/ and

/da/, and these three syllables combine to produce the above form. In other words, with regards to the linguistics form, apart from individual segmental phonemes it is necessary to assume a force which acts on sequences of these phonemes to group them into syllables. The four tones of Mandarin Chinese can be considered to be one part of this type of force. However, the phenomenon which I shall from now on call "accent" has a different function. For this reason I would like to call the four tones of Mandarin Chinese and the like "tonemes".

In my book I further divided "word pitch accent" into phonologically<sup>17</sup> significant accent and non-significant accent. The former is what is observed in dialects such as those of Tokyo and Kyoto, as outlined above. In contrast to this, I left the existence of "phonologically non-significant word pitch accent" as doubtful (p.194). In other words, although the pronunciation of a word (or combination of words) has a pitch contour which is fixed by social convention, there may be dialects where only one pitch contour occurs in one phonological environment and thus there is no phonological contrast, but it was doubtful whether such a dialect existed.

Recently, observing the pronunciation of Sumako Osada of the Yamatohama<sup>c)</sup> community (Yamatohama village, Ōshima county, Kagoshima prefecture), it has become clear that her dialect has this kind of "phonologically non-significant accent". In her pronunciation there are two types of pitch contour patterns on three-mora words. In the examples below, [nu] is a dependent word which marks nominative case, and syllables not marked for pitch have the same pitch as the preceding syllable.

'knife'	[ˈkʰa_tʰaːna]	[ˈkʰatʰa_nanu]
'mirror'	[ˈkʰagan]	[ˈkʰa_gannu]

'Knife' in isolation is pronounced H-L-M or H-L-H and when [nu] is attached it is pronounced close to H-H-L-L. 'Mirror' in isolation is pronounced high-level and when [nu] is attached it is pronounced close to H-L-L-L. However, as far as I have investigated, all words with the structure /CVCVCV/ have the same pitch contour as 'knife', and words with the structure /CVCVN/ have the same pitch contour as 'mirror', so these two patterns do not contrast phonologically. This difference in pitch pattern can be explained as being due to the phonological environment. We can interpret it as follows. [gan] is, phonologically, one syllable, and to avoid having a rising pitch contour appear on the syllable, the whole of [kʰagan] becomes high level, and to avoid a falling pitch contour on the same syllable, the [kʰa] of [kʰagannu] was pronounced high and [gannu] became low level. Thus, these two patterns can be viewed as phonologically identical. In this dialect there is also no phonological contrast in pitch contour patterns on bimoraic nouns.

**7.2** This kind of phenomenon can be called, using the terminology used in my book *Phonetics*, "phonologically non-significant word(-combination) pitch accent". However, there is some doubt over whether the phrasing "phonologically non-significant" (or "significant") is appropriate here, and I will consider this further.

There is certainly a big difference between “phonologically significant accent” and “phonologically non-significant accent”, but are there no points in common?

**7.2.1** In the Tosa dialect, the vowel immediately preceding [d] and [g] is nasalised.<sup>18</sup> However, comparing the following examples, the nasalisation in examples (3,6) tends to be weaker than that in examples (1,2,4,5).

- (1) [aĩda] ‘interval’ (2) [haĩde] ‘using ash’ (3) [takaĩdai] ‘high stand’  
 (4) [ãgo] ‘jaw’ (5) [ẽdãga] ‘branch-SUBJ’  
 (6) [tamattãgomi] ‘accumulated rubbish’

Why is this?

Also, in the Tokyo dialect, the following two examples suggest that the degree of affrication (lenition) is stronger in (7) than in (8) although this difference may vary depending on the speed of articulation and on the speaker.

- (7) [konna(d)ʒikan] ‘such a time as this’  
 (8) [onna(d)ʒika] ‘the same?’

Why is this?

The difference between these pronunciations must be explained phonologically. Writing these forms (omitting the accent symbols) as follows<sup>19</sup> does not provide a phonological explanation.

- (1) /'aida/ (2) /haide/ (3) /takai dai/  
 (7) /konna zikan/ (8) /'onnazika/

However, what if we postulate an American-style “juncture phoneme” /+/? This gives the following.

- (2) /haide/ (3) /takai+dai/ (7) /konna+zikan/ (8) /'onnazika/

If we do this, the two words in (3), when pronounced in isolation, cannot be transcribed as follows:

- (9) /takai/ (10) /dai/

This is because the phone corresponding to /d/ in (10) is strongly plosive, and the phone corresponding to the /d/ in (2) tends to be nasalised and is weakly plosive. Moreover, the phone corresponding to intervocalic /d/ in (3) is more plosive than that of (2) and less plosive than that of (10). In other words, we can explain that in (3) the voiced consonant weakens intervocalically as the voiced plosive is in the same environment as in (10). However, if we say that the voiced plosives in (2) and (10) are in the same environment, we cannot explain the existence of the voiced plosive in (3). Thus it becomes clear that these forms should be transcribed as follows.

- (2) /haide/ (3) /takai+dai/ (10) /+dai/

Doing this, because the initial portions of all of the other forms are in the same environment, we can see that the forms should be transcribed as below in order to be consistent.

- (2) /+haide/                      (3) /+takai+dai/  
 (7) /+konna+zikan/              (8) /+onnazika/

It is thought that this will provide an explanation. However, when postulating this kind of “juncture phoneme”, it also becomes necessary to postulate a phonological element to handle accentuation.

Instead of doing this, I propose to postulate a “prosodeme” which is placed on top of phonemes or syllable sequences. I believe that this prosodeme serves to unify the sequence of syllables it sits on top of, but its function does not apply at the juncture of two prosodemes. Consequently, there may be slightly differing phonetic phenomena which appear depending on the position within the prosodeme (including at its beginning or end). In phonological transcription, a sequence of phonemes possessing one prosodeme may be written without a break, and two prosodemes may be written with a break at their juncture. (This is not to say that this is the only possible transcription.) The forms discussed above can thus be transcribed as follows. (I have already discussed a similar transcription, but here it is reappearing with a phonologically new meaning.)

- (1) /ˈaɪda/      (2) /ˈhaide/      (3) /ˈtaˈkai \_dai/  
 (7) /konna zikan/      (8) /ˈonnazika/

In addition to the above-mentioned functions, prosodemes in Tokyo, Kyoto and other dialects have their own phonological characteristics. In the Tokyo dialect, for example, there is an “unaccented prosodeme” as in /ˈusi/ ‘cow’, /karada/ ‘body’, etc., and there are “accent prosodemes” as in /neˈko/ ‘cat’, /ˈinu/ ‘dog’, /iˈnoci/ ‘life’, /kokoˈro/ ‘mind’. Accented prosodemes can be further classified according to the location of the accent kernel.

One word is not limited to taking one prosodeme. For example, in the Tokyo dialect in examples like /neˈkoŋa/ ‘cat-SUBJ’ and /ˈinuŋa/ ‘dog-SUBJ’, word combinations (so-called *bunsetsu*, or breath groups)<sup>20</sup> may have one prosodeme, but on the other hand in my Kameyama dialect (Mie prefecture), examples such as /\_daiˈsaN/ ‘number three’ and /\_oNˈhahaŋimi/ ‘your mother’ should be classed as single words from their grammatical function, but have two prosodemes.<sup>21</sup>

**7.2.2** So-called “one-pattern accent” distributed in dialects of the southern Tōhoku region and elsewhere is a phenomenon which deserves attention. For example, in the Sendai dialect we have the following.

[hana]	'flower; nose'
[hanāda]	'flower/nose-COP'
[konohana]	'this flower/nose'
[konohanada]	'this flower/nose-COP'
[konohanajaso:da]	'this flower/nose-SUBJ so-COP'

Looking at the pitch contours on the part of the word corresponding to [hana] 'flower; nose', they do not appear to be at all constant, and the whole of each utterance is very monotonous. This is a situation where it would seem that we can safely say there is no accent. However, these dialects have a tendency to voice intervocalic plosives, and this tendency is stronger within words or word combinations (i.e. breath groups) than in word-initial position. Nasalisation of the vowel which immediately precedes /b/, /d/, /g/, /z/ also seems to be stronger within words (or word combinations) than in word-initial position. These differences in pronunciation must be explained phonologically.

On the other hand, in the Yamatohama dialect (Amami Ōshima) referred to above, accent patterns do not contrast phonologically, yet I believe it is appropriate to say the dialect has prosodemes. If that is the case, it is possible that there is a pronunciation intermediate between the Yamatohama dialect-type pronunciation and the Sendai dialect-type pronunciation, and it also may be that it is not possible to draw a clear distinction between the two, so with regard to the Sendai dialect too, I believe it is better to postulate a prosodeme. Not only that, but only once we adopt this position are we able to explain both the above differences in pronunciation and also aspects related to the prosodeme. Although I have not carried out a detailed observation of the pronunciation of the Tōhoku and other dialects which have the "one-pattern accent system", I believe that probably, in the case of all dialects, it is necessary to postulate prosodemes.

**7.2.3** If my position is correct, although the same term "prosodeme" is used, comparing those of the Tokyo dialect and that of the Sendai dialect, their content is very different. Tokyo dialect prosodemes contrast with each other according to their accentual features and they form a system. On the other hand, the Sendai dialect prosodeme does not contrast phonologically, and it has almost no accentual features. In other words, there is only one type of prosodeme on words of 1, 2, 3 or more moras' length, and there is nothing which could be called a constant "accent peak". The prosodeme of the afore-mentioned Yamatohama dialect also does not contrast phonologically, but we can probably say that because its corresponding "pattern" has a characteristic pitch contour, it does not lack accentual features.

About the only point the prosodemes of these various dialects have in common is that there is a difference between the initial position and the non-initial position with regard to phones corresponding to the same phoneme.

For the Tokyo and certain other dialects, one characteristic of the prosodeme where there is not a kernel on the initial mora is that “the initial mora is usually pronounced low”. Some scholars explain that this feature is just as crucial to the “accent pattern” as the “accent kernel”, but that this is incorrect should be clear from the above explanation. In other words, this feature is no more than one possible mark of the beginning of a prosodeme.<sup>22</sup> As already stated, from the viewpoint of contrasts between prosodemes, this feature is non-distinctive, and together with other features gives rise to one prosodeme.

### 7.3 The same can be said of the so-called stress accent.

The accentuation of languages such as Czech, Finnish, Polish, and probably French is, using the terminology of my book, “phonologically non-significant”. The reason for this is that, in the first two languages, the initial syllable of a stress group (i.e. a sequence of phonemes corresponding to a word or word combination) is always pronounced with increased intensity, in Polish the penultimate syllable and in French the ultimate syllable are always pronounced with (a degree of) intensity, and there is no phonological contrast due to different accentuation. However, in these languages the accent peak (i.e. intensity) not only marks the boundary between words (or word combinations),<sup>23</sup> but because there seems to be a tendency to group a sequence of phonemes corresponding to a word (combination) into one grouping using other phonetic features, in order to explain these facts, it is most appropriate to assume that words and word combinations in these languages have one prosodeme.

Concerning the “phonologically significant” accent such as that observed in English and Russian, it is usually written that “stress”<sup>24</sup> (Russian *ударение*) is added to various syllables within the word. However, I believe that assuming a prosodeme is added over the word or word combination as a whole is more convenient to explain various phonetic phenomena.

For example, regarding English, rather than saying “*better* has stress (or the accent peak) on the first syllable and *become* (or *the book*) has it on the second”, I would say “*better* has a prosodeme with a kernel on the first syllable and *become* (or *the book*) has a prosodeme with a kernel on the second”.<sup>25</sup> Treating stress as a relative notion, there are scholars who do not mark stress on monosyllabic words such as *come* and *go*, but there are also scholars who do mark these words for stress. Following my treatment, the description would be “*come* and *go* have a prosodeme with a kernel on that syllable”, so there is no room for this kind of inconsistency in marking stress.

### 7.4

Thinking in this way, whereas it might have been necessary to say that “there are languages (or dialects) with accents and those without accents”, now we can not only say “all languages (dialects) have prosodemes”, we can even say that in both pitch accent languages and stress accent languages, the function of prosodemes is the same. I believe this is a very favourable result.

All languages have prosodemes in addition to segmental phonemes as elements which form the shape of linguistic forms (words etc.). Languages like Chinese also have tonemes in addition to segmental phonemes and prosodemes, but I propose the existence of the “syllabeme” which forms phonological

syllables, so I should probably say “all languages have segmental phonemes, syllabemes and prosodemes”.<sup>26</sup> There is probably no need to point out that these elements by themselves do not convey meaning (expressing this more accurately, these elements in isolation do not, in principle, express semantic features). All of these elements combine to form the shape of linguistic forms, and that shape expresses semantic features. As a cover term for segmental phonemes, syllabemes and prosodemes, I use the term phoneme. It goes without saying that phonemes are handled in the phonology and not in the morphology or morphophonology.

There are languages or dialects where prosodemes contrast and so, as a whole, they form various systems. There are also languages or dialects where the prosodemes do not contrast and so form only the simplest of systems where they simply form a list. In both cases, I will refer to all phenomena related to prosodemes in these languages (dialects) as “accent”.

Because prosodemes are phonological elements, we cannot say that they are “accent with no phonologically contrasting patterns” or, in other words, a so-called “one-pattern accent system”. Therefore I will do away with the terms “phonologically significant accent” and “phonologically non-significant accent”, and replace them with “distinctive accent” and “non-distinctive accent”.

Finally, I would like to draw attention to the distinction in usage of terms. As mentioned above, “prosodeme” and “accent kernel” are phonological notions, but I would like to limit the terms “accent pattern” and “accent peak” to

expressing phonetic notions.<sup>27</sup> For example, the example [k̄otob̄a] ‘word’ from the Shimo-Kitayama dialect (§6) has “accent peaks” on its first and third syllables, but the “accent kernel” is only on the third mora and is not on the first.

In the same dialect, [k̄odomo] ‘child’ has a different “pattern” from the Tokyo dialect [kodomo], but they have the same “prosodeme” (3 moras and no kernel).

Also, I believe one can say that Tokyo dialect [komaŋire] ‘finely chopped’ and [ko:baŋ] ‘police box’ have different patterns, but their patterns correspond to the same prosodeme.

---

### Notes

<sup>1</sup> [§2.1] What I have stated very simply here actually has a deep significance and relates to a basic issue in phonetics and phonology.

In chapter 3 of my book *Onseigaku* [Phonetics] (Iwanami Zensho, 1951), I attempted to objectively explain phones purely phonetically, but in the case of [sn] a “faucal plosive” is audible, but I was forced to state that it is difficult to recognise this faucal plosive as belonging to either [s] or [n] (supplementary note 1, p.62). The sequence of phones [sn] can correspond to the sequence of

the phonemes /sn/, but in another language it may correspond to the phoneme sequence /stn/.

Speaking in extremely precise terms, one could even say that one would be mistaken to attempt to define the “phone” objectively without reference to the segmental phoneme. In the case of the syllable, this contradiction is of an even greater magnitude. I believe this is why there is such variation in scholars’ phonetic definitions of the syllable. The syllable has a certain phonological structure in each language, but I believe that trying to define the syllable objectively while ignoring this fact leads to various difficulties.

The difficulty in the phonetic classification of “voiceless vowels” can be traced to the same reason. The reason why the [j] in Tokyo dialect [k̟ita] ‘north’ is called a vowel is because it is a “phone” which appears in the location corresponding to the vowel phoneme /i/.

In a talk entitled “Phonemic structure of Japanese” which I gave at a luncheon conference at the 1950 Linguistics Institute held at the University of Michigan (Aug. 9<sup>th</sup>), I criticised the procedure used most commonly in the U.S.A. for “phonemic analysis” (i.e. the method of grouping together several phones to construct a phoneme), and stated words to the effect that “this is, after all is said and done, unconsciously relying on the phoneme when determining the phones, and then relying on the phones to determine the phoneme, which is unacceptable”, so what I have stated in this paper may seem very conciliatory compared to what I stated then. However, those who read what I have written above will be fully aware that I have not made any fundamental concessions.

<sup>2</sup> [§2.1] In my talk I called this the “operational principle of environment”, but I have decided to call it the “operational principle of environment assimilation” because leaving out the word “assimilation” could lead to misunderstandings.

<sup>3</sup> [§2.2] After a question from Haruhiko Kindaichi at my talk, I came around to thinking that the following case was possible. What about the example where, in a certain dialect, there are two types of bimoraic noun — those without an accent kernel and those with an accent kernel — and that the latter are further divided into the following two types:

- (a) the vowel in the second mora is /i/ or /u/ and the accent peak is on the first mora
- (b) the vowel in the second mora is another vowel phoneme and the accent peak is on the second mora.

When answering his question I explained that, in such a case, whether (a) and (b) are phonologically the same pattern or different patterns can be determined by the “operational principle of the system”. I shall expand on this here.

If this dialect’s accent system is as follows (accent kernel marked with <sup>ˈ</sup>):

1 mora words: /○/ /○<sup>ˈ</sup>/

3 mora words: /○○○/ /○○○<sup>ˈ</sup>/ /○○<sup>ˈ</sup>○/ /○<sup>ˈ</sup>○○/

4 mora words: /○○○○/ /○○○○<sup>ˈ</sup>/ /○○○○<sup>ˈ</sup>○/ /○○<sup>ˈ</sup>○○/ /○<sup>ˈ</sup>○○○○/,

we can interpret the (b) pattern and (a) pattern given above as corresponding to /○○<sup>ˈ</sup>/ and /○<sup>ˈ</sup>○/ respectively. In this situation, it would be quite easy for the

dialect to acquire a word like /ne<sup>1</sup>ko/ through influence from another dialect. This is because the reason why there are no words of the shape /CV<sup>1</sup>Co(/a/e)/ is not that it is not permitted by the phonological system, but that it is simply a “gap” in this particular dialect’s lexicon.

On the other hand, if this dialect’s accent system has only two accentual types on forms with the same number of moras, for example

- (i) pattern with no accent kernel
- (ii) pattern with an accent kernel,

the (a) and (b) above would both phonologically belong to the (ii) type. In that case, if a speaker of this dialect were to try to mimic another dialect speaker’s

pronunciation of [ $\bar{n}eko$ ], he would probably pronounce it as [ $n\bar{e}ko$ ].

The accent of the Shuri dialect of Ryukyuan (as pronounced by Seibin Shimabukuro) resembles somewhat the latter case. For words (or word combinations) of the same length, there are only the following two types of accent patterns. (In long compounds there is a third pattern where the final mora or two moras are pronounced low.)

- (A) begins high, and drops to low after the first or second mora. This will be transcribed phonologically as / $\circ\circ$ /, / $\circ\circ\circ$ /, / $\circ\circ\circ$ /, ....
- (B) begins low, and becomes high on the final mora. This will be transcribed phonologically as / $\wedge\circ\circ$ /, / $\wedge\circ\circ\circ$ /, / $\wedge\circ\circ\circ$ /, ....

(Although I use “low”, in this dialect it is “low” relative to the “high” tone but is higher than the “low” tone of the Kyoto and Kameyama dialects.)

The location of the fall from “high” to “low” in the A pattern (represented here with  $\bar{\text{˘}}$  for convenience’ sake) is determined by the shape of the word (combination) and whether another word (combination) follows or not. For example:

- [CV<sup>1</sup>CV] : [ $\bar{\text{˘}}ku^1nu$ ] ‘this’ [ $\bar{\text{˘}}ku^1u$ ] ‘bird cage’
- [CV<sup>1</sup>VCV] : [ $\bar{\text{˘}}tu^1uku$ ] ‘distantly’
- [CVCV<sup>1</sup>CV] : [ $\text{?}aka^1ku$ ] ‘red-ADV’

Although [ $ku^1nu$ ] has this pitch contour when used in isolation or when an A-pattern word follows, when a B-pattern word follows then the following pitch contour is used. (A rise from “low” to “high” is marked with  $\bar{\text{˘}}$ .)

- [ $\bar{\text{˘}}ku^1nu^{\bar{\text{˘}}ki^1i}$ ] / $\wedge$ kunu $\wedge$ kii/ ‘this hair’
- [ $\bar{\text{˘}}kunu^1kii$ ] / $\wedge$ kunu $\wedge$ kii/ ‘this tree’

[ $\bar{\text{˘}}tu^1uku$ ] ‘distantly’ is pronounced with this pitch contour in isolation, but when other words follow the pitch contour is as below.

- [ $\bar{\text{˘}}tuu^1kunajuN$ ] / $\wedge$ tuuku  $\wedge$ na $\wedge$ juN/ ‘become distant’

The B pattern is virtually level in isolation, but when another B-pattern word (combination) follows it becomes as below.

[waa] /wa'a/ 'my'

[waa<sup>r</sup>ki<sup>i</sup>] /wa'a \kii/ 'my hair'

[wa<sup>r</sup>a<sup>i</sup>kii] /wa'a.kii/ 'my tree'

Three-mora examples are as follows.

[takaku] /takaku/ 'high-ADV'

[taka<sup>r</sup>ku<sup>r</sup>tubu<sup>1N</sup>] /takaku \tubu<sup>1N</sup>/ 'fly high'

[taka<sup>r</sup>ku<sup>1</sup>naju<sup>1N</sup>] /takaku naju<sup>1N</sup>/ 'become high'

In other words, in this case, when an A-pattern word (combination) follows, the last syllable is raised somewhat.

When this situation arises, the difference in the locus of the fall in pitch (after the first or second mora) is non-distinctive. Furthermore, looking at words (or word combinations) of three moras or longer, it seems that in the A pattern the pitch falls from the second mora to the third mora, and depending on various phonological environments this fall can move back or forward in the word by one mora, so I suggest that this dialect's B pattern has no accent kernel and the A pattern has an accent kernel on the second mora.

<sup>4</sup> [§2.2] I have repeated this orally on many occasions, so I was surprised when Haruhiko Kindaichi said he was hearing it for the first time. (I believe the first time I stated this opinion was at the conference referred to in note 1.) For this reason I think it would be worth providing a little more explanation.

What I call phonological syllables and phonetic syllables are not the same. For

example, in the Tokyo dialect, when the word [h̄aʃi] 'chopsticks' is pronounced short and rapidly as [haʃ] with no dip then rise in sonority, it is phonetically one syllable. However, the word form which corresponds to this pronunciation is interpreted as /ha'si/, i.e. as phonologically disyllabic /CVCV/. Why? Because there are pronunciations corresponding to this form which are phonetically disyllabic [ˌha\_ʃi] and also [haʃ] (where the [ʃ] is not a "phonème décroissant" (*Phonetics* p.178) and cannot be said to be phonetically a syllable), and a completely monosyllabic pronunciation is not common.

On the other hand, the Tokyo Japanese words [ko<sup>1</sup>] '(crab or tortoise) shell', [ko<sup>1N</sup>] 'navy blue' etc. are /ko<sup>1</sup>o/ (i.e. /CVV/) and /ko<sup>1</sup>N/ (i.e. /CVN/), and so have a phonologically monosyllabic structure, and so their pronunciations are almost always phonetically one syllable.

As is clear from the above, the mora, a phonological unit, does not always correspond to the phonological syllable, and I must emphasise the fact that in the Tokyo dialect it is necessary to recognise the syllable as a phonological unit in addition to the mora. In other words, it is not the case that because we recognise the phonological syllable we do not need the mora. (Incidentally, it goes without saying that the distinction between the two kinds of syllable discussed above differs from Grammont's "syllabe phonétique" and "syllabe phonologique". See *Phonetics* p.179.)<sup>d)</sup>

Regarding long vowels in Tokyo and other dialects, there have been conflicting theories that they are monosyllabic and that they are disyllabic, but the above view I believe is a theory which can resolve the conflict between the two theories

and is not itself a third contrasting theory. Also, I am not claiming that in all languages long vowels are sequences of two identical vowel phonemes. For example, both German *satt* [zat] 'satisfied' and *Saat* [za:t] 'sowing' are always pronounced as monosyllables and so should not be interpreted as a contrast between /zat/ and /zaat/, but should rather be /zAt/ and /zat/, suggesting that in German, in addition to the vowel phonemes /A, I, U,.../ there are also the vowel phonemes /a, i, u,.../ which have a different length feature (i.e. they are longer). I believe that one reason why Japanese long vowels can be pronounced as two syllables is because they have the structure /aa, ii, uu,.../.

In the Tokyo dialect the phonological syllable has the following structure. (/C/ represents a consonant phoneme and /V/ represents a vowel phoneme. However, in front of /a, o, u/ the phoneme group /Cj/ may appear in place of /C/.)

/CV/

/CVV/ (where /VV/ is a sequence of identical vowels)

/CV<sub>N</sub>/

/CVʔ/ (because /ʔ/ expresses tenseness of the larynx etc., a different symbol may be substituted)

(In addition to the above, I have in the past argued that there is also the syllable /CVi/, but because there are many Tokyoites who pronounce this as /CV'i/ (i.e. as two syllables), I wish to investigate this issue further.) Therefore /ka'o/ 'face', /'a'o/ 'blue', /'u'e/ 'above' all have the structure /CVCV/, composed of two syllables, so their corresponding pronunciations are usually disyllabic [kao], [ao], [ue].

There are some dialects where the mora and the phonological syllable coincide. The Ryukyuan dialect of Shuri (as pronounced by Seibin Shimabukuro and Shunchō Higa) is such an example. In this dialect the mora or phonetic syllable has the structures /CV/, /CN/ and /Q/. (However, in certain environments the phoneme group /Cj/ or /Cw/ may appear in place of /C/, and the /C/ which precedes /N/ is restricted to the phonemes /' /, /ʔ/ and /h/.) Consequently, what is often pronounced as long vowels are /VV/, and in careful pronunciations are disyllabic.

[kaa] /ka'a/ 'skin'

[kii] /ki'i/ 'tree'

[kuu] /ku'u/ 'powder'

[ʔee] /ʔe'e/ 'indigo'

[oo] /'o'o/ 'king'

This is very different from Tokyo and other dialects' long vowels which are /VV/.

In the Shuri dialect, the moraic nasal following a vowel is also /'N/, which differs from the Tokyo and other dialects' moraic nasal. It has a marked tendency to form a syllable by itself, and it is long. There is a tendency for there to be a lowering in stress between it and the preceding vowel, so it feels little like the Tokyo dialect's moraic nasal. The following are Shuri dialect examples.

---

[ʔa<sub>N</sub>] / ʔa'N/ 'have'

[<sup>-</sup>wu<sub>N</sub>] / <sup>-</sup>u'N/ 'be'

[kaga<sub>N</sub>] / kaga'N/ 'mirror'

Also, in the case of the Tokyo and other dialects' moraic nasal, it is normal for it to assimilate in place of articulation to a following consonant, but in the Shuri dialect there is a strong tendency to try to maintain the articulation of [N]. For example:

Tokyo : [amma] /'anma/ 'massage'

Shuri : [ʔam<sub>N</sub>ma] /ʔa'Nma/ 'massage'

Comparing these examples, the Tokyo /N/ becomes the labial nasal, but in contrast to this, in the Shuri example, the /'N/ is pronounced as a bilabial stop but there also appears to be the stop [N] being articulated simultaneously. Listening carefully one can hear both timbres.

Not only this, but whereas in the Tokyo dialect the moraic nasal cannot stand as the first mora of a free form, it can in the Shuri dialect. (The symbol for syllable nucleus is omitted.)

[ŋ·kaʃi] / /Nkasi/ 'long ago'

[<sup>-</sup>n·na] / <sup>-</sup>Nna/ 'everyone'

[m·ba] / /Nba/ 'no way!'

[<sup>-</sup>n·da] / <sup>-</sup>Nda/ 'let me see!'

[n:da] / /N'Nda/? / /NNda/? 'let's look'

[ʔn·na] / ʔNna/ 'faeces'

[ʔm·ma] / ʔNma/ 'horse'

[<sup>-</sup>ʔm·ba] / ʔNma/ 'soy milk skin'

[N<sub>N</sub>:] / hNN/ 'yes (response to wife)'

In this way, we can consider the Tokyo dialect moraic nasal to be phonologically a syllable non-nucleus, whereas the Shuri one is phonologically a syllable nucleus, so I have transcribed the former with a small /N/ and the latter with a large /N/.

The moraic obstruent in the Tokyo dialect is characterized by being tightly connected to the preceding vowel, being a *phonème décroissant* and involving tensing of the laryngeal area. In contrast, the Shuri dialect moraic obstruent, although also involving tensing of the laryngeal area, differs in being a *phonème croissant-décroissant*, and having a tendency to try to form a syllable by itself, and it is long.

[ʔaʃ·sa] / ʔaQsa/ 'that much'

Not only that, but whereas Tokyo dialect's moraic obstruent cannot stand as the first mora in a free form, it can in the Shuri dialect. (The symbol for syllable nucleus is omitted.)

[t.t̥ʃu] /ʌQcu/ 'person'

[k.kwa] /ʌQkwa/ 'child'

These [t̥:], [k̥:] are not released.

In the Shuri dialect there is a contrast between / ' / and the laryngeal plosive phoneme /ʔ/ (see below), so to distinguish the moraic obstruent from this I use /Q/.

[ʔii] /ʔi'i/ 'stomach'

[(j)ii] /ʔi'i/ 'picture'

[ʔwi\_i] /ʔwi'i/ 'above'

[wii] /wi'i/ 'nephew'

As explained above, because in the Shuri dialect the mora and the phonological syllable coincide completely, as far as this dialect is concerned, either one of these terms can be done without. On the other hand, in the Tokyo dialect, because the mora and phonological syllable do not coincide, it is necessary to retain both terms to distinguish the entities.

<sup>5</sup> [§2.2] Taking just the description in this section as one example, it should be clear that I am not simply a mentalist who emphasises subjectivism. Where I differ from being what I call "mechanistic mechanical" is that I did not shy away from adding the last sentence in this section, but that I recognised that it was necessary.

<sup>6</sup> [§3.1] Very strictly speaking, this expression is incorrect. Rather than using the word "phoneme" I should have used "the majority (or all) of the various phones which correspond to the phoneme". For example, the phone which corresponds to the phoneme /m/ in the Tokyo and other dialects is a bilabial or labiodental nasal but the majority of instances of these are voiced.

<sup>7</sup> [§3.2] This is a very rough description. When the accent kernel is on the final mora of a form, often the form has a pitch contour indistinguishable from the unaccented or flat pattern, but it differs from the flat pattern in that there is a pressure to lower the following form. However, when the following form has a "high" tone on its initial mora, there is a tendency for the accent kernel itself to be realised with a falling pitch and this tendency does not exist with the final mora of flat pattern forms. Please refer to Shin Kawakami's paper "'*haná takashi*' and '*hana takashi*'" ['The flower is expensive' and 'the nose is big'] in *Onsei-gakkai Kaihō* [The Bulletin of the Phonetic Society of Japan] 82: 6-9 (1953).

<sup>8</sup> [§3.2] Note that this new theory differs from what I wrote in *Phonetics* p.193. In that book I followed the rules of the International Phonetic Alphabet and placed the tone symbols in front of the phonetic letters making up the syllable, but Japanese accent is fixed according to the words, not the syllable, so to make the transcription of Japanese easy to follow, I have placed / ˈ / on the symbols representing the mora with the accent kernel.

<sup>9</sup> [§3.2] Russian has a “stress accent” but the word pitch contours which accompany it are patterns which are fixed by social convention (stressed syllables are not necessarily pronounced high). These word pitch contour patterns are phonologically non-distinctive for Russian accent, but if one uses random pitch contours when mimicking Russian pronunciation, even if one stresses the correct syllables, it will be judged as being pronounced with a foreign accent. See my paper “Roshiya-go-no tango-no onchō-ni tsuite” [On the intonation of words in Russian] in *Onsei-no Kenkyū* [Study of Sounds] 6: 91-93 (1937).

<sup>10</sup> [§3.2] Haruhiko Kindaichi stated words to the effect that the “high-beginning pattern” in these dialects actually begins at normal height, and the “low-beginning pattern” is actively depressed below that level.

- (a) [konokodomo] ‘this child’  
       [konosuzumē] ‘this sparrow’
- (b) [ānokodomo] ‘that child’  
       [ānosuzumē] ‘that sparrow’

The (a) examples above would seem to support this view, but in (b) it must be said that (the first mora of) [kodomo] has been actively raised. The pitch contour pattern of [kodomo] is high-beginning, and that of [suzumē] is low-beginning, and we must recognise that this is a distinctive feature. However, *kono* ‘this’ and *kodomo* ‘child’ are both high-beginning, and *aru* ‘a certain’ and *suzume* ‘sparrow’ are both low-beginning, and all four words lack accent kernels, but in spite of this ‘this child’ is [konokodomo], and ‘a certain sparrow’ does not become low-level but instead becomes [arū suzumē]. However, this is due to the fact that the pattern “that begins low and has no accent kernel” has the non-distinctive feature that, so long as a high-beginning pattern does not immediately follow, the last mora is pronounced high, and as [ānosuzumē] ‘that sparrow’ and [naitasuzumē] ‘sparrow that sang’ demonstrate, when it continues on from a low-ending pattern, it continues on at the same pitch and does not further drop in pitch.

<sup>11</sup> [§4] Miyata in the same paper published in *Onsei-no Kenkyū* [Study of Sounds] 2 says that in the Tokyo dialect word [ataraʃi:] ‘new’, the feature “rise in pitch from the first syllable to the second syllable” can be thought to be “unimportant as an accentual phenomenon”, and his primary reason for this is that “even if we pronounce this word avoiding a rise in this location as in the provisional pronunciations [ataraʃi:], [ataraʃi:] and [ataraʃi:], we do not feel

that they are fundamentally different patterns". What I stated in the main text applies equally to this point.

<sup>12</sup> [§4] I have explained that distinctive features and non-distinctive features should be distinguished and in phonology one should especially pay attention to the former, but I do not go so far as to say that non-distinctive features should generally be completely ignored. This point is important. On the occasion of my talk I was asked a question which led me to suspect that this point was misunderstood, so I draw attention to it here.

Also, the question of whether, if my theory is correct, all observations of accent made until now which are based on intended-pronunciation phonology are completely worthless is related to the misunderstanding referred to in the preceding paragraph. It is precisely because we have past observations and descriptions that we have been able to build our theory on these, and in order to consider the validity of this theory, we must further refine these kinds of observations.

Another point which should be mentioned is that the fact that the [ʃi] of the Tokyo and other dialects is interpreted phonologically as /si/ does not mean that in the process of aiming to pronounce [si], the [s] assimilates to [i] and coincidentally we end up with [ʃi]. Only by recognising that Tokyoites, even when pronouncing very carefully, aim at [ʃi] do we recognise that, of the various phonetic features [ʃ] has, palatalization is phonologically non-distinctive. And of course this does not mean that it is acceptable to pronounce all instances of Tokyo dialects /si/ as [si].

In the same way, interpreting Tokyo and other dialects' [atamaŋa] 'head-SUBJ' phonologically as /'atamaŋa/ does of course not mean that this word may always be pronounced as [atamaŋa].

<sup>13</sup> [§5] *Language* (Journal of the Linguistic Society of America) 26(1): 86-125 (1950), section 3.

<sup>14</sup> [§5] There are cases where an accent pattern does not express meaning but sentence intonation does. See *Phonetics* p.197, and §7.4 of this paper.

<sup>15</sup> [§5] I carried out this investigation at Kominoue in Totsukawa village, Yoshino county, Nara prefecture on 3<sup>rd</sup> Sept. 1932. Mr Uenishi was at the time 18 years old and a student at Bunbukan Totsukawa Middle School.

<sup>16</sup> [§6] Writing this does not mean that, when describing historical accent change, we should only concern ourselves with phonological changes and need not describe phonetic changes. I am claiming that we need to describe historical accent changes paying attention to the distinction between the two, and by doing this I believe we can hope to further deepen our understanding of the historical accent changes themselves.

Haruhiko Kindaichi asked how I view the sound change [Φ] > [h] which took place from the Muromachi period onwards (1392 – ) in Japanese. He probably asked because in Dr Arisaka's style of phonology this is a phonological change, but certain people have viewed it as a phonetic change. To this question I responded that I view it as a phonological change because whereas /Φ/ is like /b/ in having a bilabial feature, when it became /h/ it came to have a laryngeal

feature like / ' /. Moreover, /b/ and /' / have coexisted throughout all periods as phonemes different from either of them. Below I will expand on this.

The segmental phonemes of a language/dialect are not isolated and unrelated to each other; as a whole they form a phoneme system where two (or more) segmental phonemes share features. Any change which changes the system will be a phonological change. Changes in distinctive features are, accordingly, phonological changes. The same can be said of accent.

It is thought that in the (Standard) Japanese at the end of the Muromachi period (i.e. late 16<sup>th</sup> century) there was the following phonemic contrast in voicing. (/ ' / is a voiced laryngeal phoneme which is observed in the modern Tokyo and Kyoto dialects, among others. Refer to the end of this note.)

/b d g z ' /

/Φ t k s /

In addition to these there was also [p], but I think this was still a phone which occupied an opening in the phoneme system. This is because, according to the 1603 Nagasaki edition of the *Vocabulario da Lingoa de Iapam*, there were only the following small number of words where this phone stood in word-initial position.

Pappato 'briskly'

Pararito *or* fararito 'dropping in small light pieces'

Pararito vchicuzzusu 'knock down easily'

Patto 'suddenly'

Chiqemuriga patto tatta 'a spay of blood spurted up'

Paxxito 'smack (sound)'

Paxxito atatta 'hit with a loud smack'

Pinpin 'with vigour'

Vmaga pinpinto fanuru 'A horse prances vigorously'

Pixxito 'crack (sound)'

Ponpon *or* poponto 'in rapid fire'

Poppoto 'puffing (of smoke or steam)'

Fonouoga poppoto tachinoboru 'flames rise up gently'

The above words are all annotated with the abbreviation *Adu.* (adverb). It is noteworthy that all of these words are mimetic vocabulary which were usually pronounced emphatically. In all languages these kinds of words frequently appear with phones which lie outside the phoneme system. In the above Japanese words we can consider /Φ/ to have strengthened to become [p]. (Note also that the emphatic form *Pararito* has a non-emphatic form *fararito* /Φararito/.) I thus consider it invalid to use the [p] in these words as evidence that /Φ/ was pushed out from the voiceless phoneme position contrasting with /b/ and /p/ subsequently moved into it.

In the examples of foreign words in Prof. Shinkichi Hashimoto's book *Bunroku Gannen Amakusa-ban Kirishitan-kyōgi-no Kenkyū* [Research on the 1592

Amakusa edition of the *Doctrina Christan*] (1928), there are a little over twenty words which begin with *p*, including *Pã* (Jp. *pan*) 'bread', *Paciencia* (Jp. *pashiensha*) 'patience', and *Padre* (Jp. *pãdere*) 'father, priest'. Of all of these words only *pan* 'bread' is used in the modern colloquial language, so it may be assumed that these words were used only among Christians. Prof. Hashimoto notes that in the Japanese script version of this *Dochiriina Kirishitan* [*Doctrina Christan*] (date of publication unknown) the words beginning with *p* are all written with syllabic script characters which represent *h*. Thus, even if Christians with the phoneme / $\Phi$ / pronounced these words using [p], they were using the [p] which appears in the above mimetic words, and as they became more familiar with these foreign language words, they probably came to pronounce not a few of these words with / $\Phi$ / in place of [p].

In a period when / $\Phi$ / still retained a strong labial articulation, it would have contrasted with the voiced consonant /b/, so these two phonemes have in common the feature "mellow", and this differed from the "strident" of /z/ and /s/. Thinking in this way, we can understand why the phoneme / $\Phi$ / was a bilabial fricative and not the "strident" labiodental fricative [f].

Nevertheless, when this bilabial fricative lost its bilabial articulation and became /h/ (the pronunciation of modern Tokyo and Kyoto dialects) the following voicedness contrast arose on the phonemes. (/t, d/ in front of /i, u/ changed to /c, z/.)

/b d g z ' /

/p t k c s h /

The new phoneme /p/ appears in onomatopoeia, mimetic words and high frequency words such as *pan* 'bread', *pen* 'pen', *pin* 'pin' and *ponpu* 'pump'.

According to Prof. Hideyo Arisaka's research (*Kokugo On'in-shi-no Kenkyū* [Research on the Phonological History of Japanese] p.215ff (1944)), in the Kyoto dialect of the Kanbun period (1661-1673), it can be assumed that / $\Phi$ / had already changed to *h* or was extremely close to *h*, but in the Nagasaki area it was labial until around the 18<sup>th</sup> century. If *pan* 'bread' is a loanword from Portuguese, it can be considered to have first set down its roots as a loanword in an influential dialect (such as the Kyoto dialect) where / $\Phi$ / had already changed to /h/, rather than in a dialect of the Nagasaki area. The reason for thinking like this is that if the Nagasaki dialect retained / $\Phi$ / until such a late period, as the word gained currency among the general populace, it would have become / $\Phi$ aN/ and then later *han*.

Incidentally, I would like to add the following. In the *Bunroku Gannen Amakusa-ban Kirishitan-kyōgi* (1592), the Japanese moras beginning with *s*, *z*, *t*, and *d* are Romanised as follows.<sup>e)</sup>

<i>sa</i> <sa>	<i>si</i> <xi>	<i>su</i> <su>	<i>se</i> <xe>	<i>so</i> <so>
<i>za</i> <za>	<i>zi</i> <ji>	<i>zu</i> <zu>	<i>ze</i> <je>	<i>zo</i> <zo>
<i>ta</i> <ta>	<i>ti</i> <chi>	<i>tu</i> <tçu>	<i>te</i> <te>	<i>to</i> <to>
<i>da</i> <da>	<i>di</i> <gi>	<i>du</i> <zzu>	<i>de</i> <de>	<i>do</i> <do>

Prof. Hashimoto proposes that these spellings represented the following pronunciations (*Kokugo On'in-no Kenkyū* [Research on Japanese Phonology] p.88,90 (1950)).

---

[sa	ʃi	su	ʃe	so]
[za	ʒi	zu	ʒe	zo]
[ta	tʃi	tsu	te	to]
[da	dʒi	dzu	de	do]

Dr Arisaka also claimed that the Kansai dialect *chi* and *tsu* were already pronounced as [tʃi] and [tsu] in the late 16<sup>th</sup> century, and so agrees with Prof. Hashimoto's theory (*Kokugo On'in-shi-no Kenkyū* [Research on the Phonological History of Japanese] p.157 (1944)).

Either way, if the consonants in *zi* and *di*, *zu* and *du* were clearly distinguished as phonemes, using the operational principle of the system the moras listed above must be interpreted as corresponding to the following phoneme sequences.

/sa	si	su	se	so/
/za	zi	zu	ze	zo/
/ta	ti	tu	te	to/
/da	di	du	de	do/

Analysing it in this way, the consonants in the syllables /ti, tu, di, du/ were not "strident" affricates, and even if they were tending toward affrication, they would probably have been strongly plosive and the fricative part would have been a weak and relatively short "mellow" sound. (I explained the terms "mellow" and "strident" in §2.1. See section 11 of my paper in *Kyōiku Gijutsu*.)

For example, the pronunciation of *chi* should be represented as something like [tʃi], and its consonant [tʃ] can probably be explained as a voiceless aspirated /t/ in the environment of a following /i/ having been palatalised to this following vowel. The *chi* of the modern-day Kōchi dialect is close to this in pronunciation. (See my paper "Kōchi-hōgen-no hatsuon-ni tsuite" [The pronunciation of the Kochi dialect] in *Onsei-no Kenkyū* [Bulletin of the Phonetic Society of Japan] 23: 6-7, 3 (1931).)

The pronunciation of *tsu* could be represented as [t<sup>s</sup>u] or the like, and I suggest its pronunciation can be explained as a voiceless aspirated /t/ assimilating to a following /u/.

Kōchi dialect *tsu* can be represented as [tu] (the consonant is somewhat aspirated). In my 1931 paper mentioned above, it is recorded that when a certain acquaintance of mine from Kōchi city pronounces this syllable with emphasis, the pronunciation approaches [tsu], but in normal speech the pronunciation is close to [tu]. Actually, this person had been living in Tokyo for about six years before I recorded this, and I later discovered that because his [tu] pronunciation was always laughed at in Tokyo, he tried his best to hide his original pronunciation by copying the Tokyo [tsu]. I have spoken about this to Dr Arisaka and others, but I believe this may be the first time I have published this in print form. I record this here to prevent any misunderstanding.

As can be seen from Kōchi dialect examples, because the back vowel *u* does not assimilate /t/ changing it to the affricate /c/, we can assume that the affrication of the consonant in /tu/ began in dialects where *u* was a central vowel or close to that, and in the course of its pronunciation the tongue blade

approaches the front teeth and the alveoli and this pronunciation then spread to dialects where *u* was not pronounced like this.

When the consonants in the moras /ti, tu; di, du/ changed from “mellow” to “strident”, they became /ci, cu; zi, zu/ with different consonant phonemes from the consonant phonemes /t, d/ in /ta, te, to; da, de, do/, and the /zi, zu/ formed in this way probably merged with the pre-existing /zi, zu/ in the /za, zi, zu, ze, zo/ series.

In the *Bunroku Gannen Amakusa-ban Kirishitan-kyōgi* (1592), the following transcription is used.

<i>a</i> <a>	<i>i</i> <i,y,j>	<i>u</i> <v,u>	<i>e</i> <ye>	<i>o</i> <vo,uo>
<i>ya</i> <ya>		<i>yu</i> <yu>		<i>yo</i> <yo>
<i>wa</i> <va,ua>				

Based on this Prof. Hashimoto concluded that the vowels *a*, *i*, *u* “were approximately the same as the modern pronunciation” but <ye> and <vo> (<uo>) probably transcribed pronunciations close to *ye* and *wo* respectively (*Bunroku Gannen Amakusa-ban Kirishitan-kyōgi-no Kenkyū* [Research on the 1592 Amakusa edition of the *Doctrina Christan*] p.34). However, referring to his table of CyV and CwV syllables (pp.8,9), we observe that there were only the sequences /Cja, Cju, Cjo, kwa/ in the book, and there were no /Cje/ or /kwo/. This is because the long vowels transcribed as <ō> and <ô> were the long vowels [ɔ:] and [o:] respectively, and because these are interpreted as /ao/ and /oo/ respectively, <quō>, <qvō> (i.e. [kwɔ:]) would be interpreted as /kwao/, and <quô> [kwɔ:] would be /kwoo/ which did not exist. The reason for interpreting [ɔ:] as /ao/ is as follows. Postulating that this long vowel corresponded to the special phoneme /ɔ/ would be a major violation of the operational principle of the system, so it is appropriate to postulate either /au/ or /ao/ (which differ from /a'u/ and /a'o/). According to Prof. Hashimoto, this long vowel was produced by the change *au* > *ao* > *ō* (*Kokugo On'in-no Kenkyū* [Research on Japanese Phonology] p.89) so I believe /ao/ should be chosen.

With this it becomes clear that the sounds represented by the spellings above should be interpreted phonologically as follows.

<i>a</i>	<i>i,y,j</i>	<i>v,u</i>	<i>ye</i>	<i>vo,uo</i>	<i>ya</i>	<i>yu</i>	<i>yo</i>	<i>va,ua</i>
/a/	/i/	/u/	/e/	/o/	/ja/	/ju/	/jo/	/wa/

Thus, although the pronunciation of what was written as <ye> and is considered to correspond to /e/ tended to sound like [je], its [j] was probably shorter and weaker than the [j] in [ja], [ju], [jo] which correspond to /ja, ju, jo/. A similar phenomenon can be observed in the Ryukyuan dialect of Shuri and the Kyushu dialects among others. Similar to this, although the pronunciation of what was written as <vo>, <uo> and is considered to correspond to /o/ tended to sound like [wo], its [w] was probably shorter and weaker than the [w] in [wa] which corresponds to /wa/.

From this the following syllables which contain long vowels will of course be interpreted phonologically as follows.

<vō>	[ <sup>w</sup> ɔ:]	/ao/
<vô> <uô>	[ <sup>w</sup> o:]	/oo/

Please note that, parallel to [jɛ] being the pronunciation corresponding to /'e/, the pronunciation corresponding to /se, ze/ will be close to [ʃɛ], [zɛ].

<sup>17</sup> [§7.1] In my book I wrote *on'inrontekini* 'phonologically' (i.e. of the theory of phonology), but this should be corrected to *on'intekini* 'phonologically' (i.e. of the phonemes).

<sup>18</sup> [§7.2.1] This is based on the pronunciation of Minoru Ueta of Susaki-chō, Takaoka county, Kōchi prefecture and others. See p.19 of my paper "Bunsetsu to akusento (1)" [Breath group (syntagma) and accent] in *Hōgen-to Minzoku* 3: 9-18, Nihon Hōgen Kenkyūjo (ed.) (1949).

<sup>19</sup> [§7.2.1] We must clearly view the nasalisation on the vowels as belonging to the phoneme /d/ which follows the vowels. This is because that if we were not to do this, we will no longer be able to explain the existence of vowel nasalisation phonetically, as we cannot postulate nasal vowel phonemes in addition to the oral vowel phonemes. We should also view /g/ as similarly being a phoneme which accompanies transitional nasalisation. That being said, however, Tosa dialect /b/ does not trigger similar nasalisation. But this does not mean that this phoneme stands isolated in the system. This is because /z/ also does not cause nasalization. In other words, /d/, /g/ and /b/, /z/ each share a feature which the other group does not have. In this way these phonemes in this dialect by themselves form a system.

In contrast to this, /b/, /d/, /g/, /z/ in the Tōhoku dialects all have the feature of being accompanied by transitional nasalisation, and this is a point of difference from the Tosa dialect. And because of this we can interpret Tōhoku dialect word-initial [˜g] and phrase-medial [ŋ] as corresponding to the same phoneme /g/. (This does not violate the operational principle of environment assimilation.) However, both the operational principle of the system and the operational principle of environment assimilation prevent us from viewing Tokyo dialect [g] and [ŋ] as belong to the same phoneme, and we have to view them as corresponding to the separate phonemes /g/ and /ŋ/. These two Tokyo dialect phonemes are probably the result of a split into two phonemes of a /g/ like that of the Tosa dialect due to a phonological change, and examples where the two contrast phonologically are extremely rare. The fact that everywhere within the historical boundaries of Tokyo city the number of individuals who use only /g/ and do not use /ŋ/ is increasing is probably an indication that this vestigial /ŋ/ is dying out.

<sup>20</sup> [§7.2.1] One reason why I have avoided using the term *bunsetsu* 'breath group' is because this term contains the morpheme *bun* which means 'sentence', and in my terminology the "sentence" is the primary abstract linguistic unit and the "word" is a secondary abstract linguistic unit. The so-called *bunsetsu* 'breath group' belongs to the same level of forms as the "word", so rather than calling it a *bunsetsu* (lit. 'sentence - division'), I will call it a 'word - combination'.

<sup>21</sup> [§7.2.1] These various points deserve further consideration, but I shall omit it from this paper.

<sup>22</sup> [§7.2.2] Like the high-beginning prosodemes of the Kyoto and other dialects, there are prosodemes which lack this feature.

<sup>23</sup> [§7.3] At my talk at the University of Michigan, Prof. Jakobson reported that he had been able to confirm the following fact experimentally. He built a device which played sounds with a “strong-weak-weak-strong-weak-weak...” rhythm with an equal interval between each sound, and when this was played to a Czech, it was heard as “strong-weak-weak / strong-weak-weak...”, when it was played to a Pole, it was heard as “weak-strong-weak / weak-strong-weak...”, and when it was played to a French person, it was heard as “weak-weak-strong/ weak-weak-strong...”. What was physically the same stimulus was being interpreted differently by speakers of different languages, and this is due to differences in the accentuation of those languages, and we must assume that a system which interprets in such a manner lies in the brain.

<sup>24</sup> [§7.3] Bloomfield calls this a “secondary phoneme”. There are some who call it a “prosodeme”, but it is clear that that usage differs from what I call the “prosodeme”.

<sup>25</sup> [§7.3] I will not discuss the issue of secondary stress here.

<sup>26</sup> [§7.4] There is room to reconsider the postulation of the existence of the “mora”, peculiar to languages like Japanese and Mongolian, from this viewpoint, but I will leave deep analysis to another occasion.

<sup>27</sup> [§7.4] At the social gathering after the Society of Japanese Linguistics Research Symposium on 11<sup>th</sup> Oct. 1953, someone said to me words to the effect that “that which has one accent peak is recognised as one word (i.e. compound)”, and to this Teruo Hirayama responded with a question along the lines of “how should we analyse dialects where, even in careful pronunciation

there are words like [āsagāo] ‘morning glory’ which have two accent peaks?”,

and he asked my opinion. My reply was: “In examples such as [āsagāo], the accent kernel is on [ga], and the “accent peak” on [a] is phonologically non-distinctive. The reason we can say this is that, in such dialects, when the accent kernel is on the third mora or later in a form, the first (or first and second) moras are regularly high, and this “high” can in some situations disappear. In general, the phonetic description of careful pronunciations and its phonological interpretation may differ. For example, Tokyo and other dialects’ *shi* is pronounced, even in careful pronunciation, as [ʃi] and not as [si], but it is interpreted phonologically as /si/.” The explanation may have been oversimplified and so difficult to understand, so in order to expand the above explanation I gave a talk entitled “The accent of Japanese from the phonological viewpoint” at Kokugakuin University on 14<sup>th</sup> November.

After responding to Mr Hirayama as above, I started to wonder why such a question arose. It is clear that at its root is the fact that phonetic description and phonological interpretation were not being distinguished, but I started to feel that another direct reason may lie in the term “one-pattern accent system”. It occurred to me that if only pitch contours which contrast phonologically, such as the accent of the Tokyo and Kyoto dialects, are called “accent”, and the so-called “one pattern accent system” is not called “accent”, the above misunderstanding may be prevented, and I proposed this in my talk.

Later reflection led me to think that “accent” should be defined as in this paper, and by restricting the terms “pattern” and “peak” to phonetic notions, and by

---

newly adopting the phonological notions of “prosodeme” and “accent kernel”, the same effect could be achieved.

Finally, on the occasion of my talk at Kokugakuin University, there were some statements which suggested misunderstandings, and I address these below.

Distinguishing phonetics and phonology does not mean giving up on a holistic approach to the sounds of language, and it certainly does not mean that if we do phonology we do not need phonetics. Until now I have been claiming that because a phonological approach was unconscious and moreover inconsistent, it is necessary to consciously make a clear distinction between a phonetic approach and a phonological approach, and further push forward with the phonological approach.

Phonological analysis is only possible if built on detailed and accurate phonetic observations, and phonetic observations can only become more accurate if they are supported by phonological analysis. Phonetics and phonology are, in a manner of speaking, two wheels on the same vehicle — they exist separately while at the same time depending on each other.

There is the suspicion that, when phonetics and phonology are separated, we may actually lose contact with the reality of the sounds of language, but this too misses the mark. As our approach becomes more analytical, our research becomes more scientific, and observing and inquiring analytically does not mean letting go of comprehensive observation and inquiry. On the contrary, the former is the path to the latter. I believe it goes without saying that observation and inquiry where phonetics and phonology are differentiated is more scientific than inquiry where phonetics and phonology are not differentiated.

### **Supplementary note**

The above is an amended version of a talk I gave at Kokugakuin University on 14<sup>th</sup> Nov. 1953 to which I have added section 5 and the notes, and section 7 was rewritten.

I express my heartfelt appreciation to Prof. Kyōsuke Kindaichi who recommended I give the talk, and to the staff and students of Kokugakuin University. I am also grateful to Teruo Hirayama whose question I refer to in note 27 provided the impetus for writing this paper. I also express thanks to all those who asked me questions after my talk. Many of the notes appended to this paper would not have been written were it not for those questions. I am become more keenly aware that questions and debate are important for the advancement and spread of scholarship.

(First revision: 29 November 1953

Second revision: 7 February 1954)

---

### **Translator’s notes**

- a) Here I abbreviate low, mid and high to L, M and H respectively.
- b) What is called the Tosa dialect here and elsewhere is the same dialect as what is called the Kōchi dialect in note 16.
- c) Yamatohama is the same as the Yamatoma given in Hattori’s paper “Prosodeme, syllable structure and laryngeal phonemes” (*Bulletin of the Summer Institute of Linguistics, Vol. 1: Studies in Descriptive and Applied*

---

*Linguistics*, pp. 1-27 (International Christian University, Tokyo, 1961). The Yamatohama dialect word for the locality is *Yamat<sup>h</sup>oma*.

- d) Grammont, adopts Saussure's terminology where *phonétique* corresponds to what we call "phonology", and *phonologie* is what we call "phonetics".
- e) I use angled brackets ( < > ) to enclose the orthography used by the Portuguese missionaries in this book.

---

**ORIGINAL PAPER**

HATTORI Shirō, 1954, On'in-ron-kara mita kokugo-no akusento, *Kokugo Kenkyū* 2, Kokugakuin Daigaku Kokugo Kenkyū-kai, pp. 2-50.

Translated by Wayne Lawrence (The University of Auckland)

Proofed by John H. Haig (University of Hawai'i at Mānoa)