ANN.BULL.RILP(1991)No.25. 63-73

# A NOTE ON PHONETIC COUNTS IN SPOKEN JAPANESE

Pierre A. Halle

## Introduction

By "phonetic counts" in a given language, we mean comprehensive information on the frequencies of occurrence of the various phonetic types in that language, when used by normal adult speakers. Phonetic counts certainly differ according to the corpus from which they are derived. Large corpuses of spontaneous speech are most desirable if the frequencies of use of phones in a spoken language are wanted. Least desirable are corpuses of written material. Finally, the most detailed information one can hope include contextual counts as well as non-contextual counts.

Phonetic counts are a helpful documentation. From an engineer's point of view, they may serve to improve a speech recognition system (Itahashi, Suzuki & Kido, 1971; Jelinek, 1976). More generally, they permit to compare different languages, or dialects, or to compare the speech of individual speakers (possibly with a speech pathology condition) to the "normal" adult speech in a given language. If these individual speakers are infants, the outcome of the comparison may be informative on the early development of speech acquisition.

Phonetic counts are available for various languages, in particular American English (Mines, Hanson & Shoup, 1978), British English (Fry, 1947), French (for an extensive review, see Wioland, 1972), and many others like Dutch, German, Italian, Standard Chinese, Cantonese, etc. The state of affairs however, is far less satisfiying in Japanese.

One can find some information in "Zusetsu Nihongo" (1982) where frequencies of occurrence of vowels are reported from Onishi (1937) and Kramsky (1966): There is some discrepancy because Onishi counted only short vowels while Kramsky counted both short and long vowels, but the two studies agree in that /a/ and /o/ are the most frequent vowels, /e/ and /u/ the least frequent, and /i/ stands in between. In his fourth "Study in Colloquial Japanese", Bloch (1950) presented counts of phonemes that also confirm this ordering of vowel frequencies. However, his counts are based on little and may be not quite representative data (his phonemic transcription of a short piece of written material: a version of Momotarou's story). We will come back later to Bloch's counts.

Apart from these scarce data, there exist many extensive lexicographic studies focusing on written Japanese, where counts of word types, kanjis, and kanas can be found. Outstanding works of that kind have been continuously reported by the Kokuritsu Kokugo Kenkyuusho (abbreviated as "Kokken"). But most of these works are based on written material, not speech, and present counts of kanas from which counts of phonetic types cannot be derived. As an example, let us consider a research from the Kokken by

Nakano (1973), focusing on "phoneme sequences". Apparently, one goal of this study was to compare the distributions of phonemic types in Japanese words from different origins: Native Japanese, kanji vocabulary, and more recent loanwords. The authors have used a considerable amount of written material as their source data (about one million words), and present their results in the form of counts of kanas. From these counts, counts of phonetic types cannot be recovered, even approximately. First of all, comes the problem of the A-series kanas (single vowels): The I kana may stand for both /i/ or the lengthening of [e]; the U kana for both /u/ or the lengthening of [u] or [o]... Next are the problems of the TSU kana (whether pronounced [tsu], or gemination mark), of the N kana ("haneru oto") whose phonetic realizations highly depend on the context (e.g. [m] before labials, [ŋ] before velars), of the HA and HE kanas, often pronounced [wa] and [e] instead of [ha] and [he]. And so forth. One consequence is that counts of vowels cannot be recovered from the counts of kanas. From an otherwise interesting tabulation of the observed 2-vowel sequences, it is obvious that counts of /i/ and /u/ are grossly overevaluated (pages 114-115: Tables 10-13).

However, there is at least one work dealing with speech material, also from the Kokken, by Sawaki and colleagues (1980), that permits an approximate recovery of counts of phonetic types. Again, results are largely presented in the form of counts of kanas, but kanas are used here as a means to transcribe sounds. The counts are detailed according to the right context and include counts of some useful symbols: A vowel-lengthening symbol ('-') is used where needed instead of the I or U kanas; main or secondary phrase boundaries are signalled by the usual symbols ('/' or '//'). In short, most of the problems mentionned above can be solved. Approximate phonetic counts can be derived from the counts of kanas, including contextual counts of vowels, according to the preceding consonant. We will now present the phonetic counts that we have been able to derive from Sawaki's investigation.

## Vowel types

In the case of vowels, we cannot take allophonic variations into consideration. Rather than phonetic types, we are thus counting phonemic types: The five vowels /i/, /e/, /a/, /o/, and /u/ of the Japanese vocalic system.

Since kanas are counted for each possible right context, and since vowel-lengthening is distinguished from A-series kanas, it is possible to infer counts of short and long vowels. We applied the following principles: Counts of the vowel V are derived from the counts of CV and V kanas. Kanas whose right context is the vowel-lengthening symbol '-', indicate a long vowel V. All others indicate a short vowel V. Hence the counts of the five vowels, shown in Table 1.

Table 1. Counts of vowel phonemes from Sawaki's data. Frequencies of occurence are indicated within vowels alone, and within all phones.

| VOWEL | Count of short vow. | Count of long vow. | Percent of long vow. | Total | Frequency within vowels | Frequency (overall) |
|---|---|---|---|---|---|---|
| /i/ | 1110 | 51 | 4.4% | 1161 | 18.6 | 9.9 |
| /e/ | 776 | 135 | 14.8% | 911 | 14.6 | 7.7 |
| /a/ | 1876 | 20 | 1.1% | 1896 | 30.4 | 16.1 |
| /o/ | 1089 | 288 | 20.9% | 1896 | 22.1 | 11.7 |
| /u/ | 735 | 154 | 17.3% | 889 | 14.2 | 7.6 |
| Total | 5542 | 648 | 10.5% | 6244 | | |

## Consonant types

We are facing here a more difficult problem, because *phonetic* consonant types cannot be inferred from the kana writing in a straightforward manner. Sometimes, they cannot at all: The consonant in the sounds transcribed with a GA-series kana may be whether [g] or [ŋ] (and possibly [ɣ] in the case of a lax articulation). An over-simplifiying rule tells us that it should be [g] in initial position, and [ŋ] in intervocalic position. This rule however, does not apply in many cases: [g], not [ŋ], is pronounced in intervocalic position in "expressive words" like /geragera/, /guzuguzu/, etc. (see Kinda-Ichi & Maes, 1978). Another difficulty is the phonetic value of /N/ (the "haneru oto"). However, the usual pronounciation of /N/ can be inferred in most cases, since its right context has been recorded for every occurrence.

We summarize below the principles we used to recover phonetic types from the kana transcription.

GA-series: [g] in initial position; [ŋ] in intervocalic position or after the haneru oto /N/.

SA-series: [s] before all vowels but /i/; [ʃ] before /i/ or before a "you'on"

("ya", "yu", or "yo").

ZA-series: [z] before all vowels but /i/; [ʒ] or [dʒ] before /i/ or you'on.: [dʒ] in initial position, [ʒ] in intervocalic position.

TA-series: [t] before /a/, /o/, and /e/; [ʦ] before /u/; [ʧ] before /i/ or you'on. The TE kana followed by a smaller I kana is retranscribed [ti].

DA-series: [d] before /a/, /o/, and /e/; [z] before /u/ (although [$^{d}$z] is a frequent realization). Same as for ZA-kanas before /i/ or you'on. The DE kana followed by a smaller I kana did not occur in the corpus.

NA-series: [n] before all vowels but /i/; [ɲ] before /i/ or you'on.

HA-series: [h] (laryngeal fricative) before /a/, /e/, and /o/; [ɸ] (bilabial fricative) before /u/; [ç] (palatal fricative) before /i/ or before a you'on.

YA-series: Always the palatal semi-vowel [j]. The you'on are ignored when appended to the SHI, CHI, JI, or HI kanas. They are counted as [j] elsewhere. For example, KYA (KI + ya) is retranscribed [kja], where one [k], one [j], and one [a] are counted.

RA-series: Always [ɾ] (although it may sometimes be realized [l], [ɾ], or some other variant of [l]).

WA-series: Bilabial semi-vowel [w] before /a/; velar [w] before /o/ (but [wo] occurred only twice).

Haneru oto /N/: [m] before a labial consonant; [ŋ] before a velar consonant; [n] before a dental or palatal consonant, excepted before fricatives; [ŋ] in final position, i.e. before a main or secondary phrase boundary; nasal vowel before a fricative [s] or [ʃ], or before a vowel. These nasal vowels can be transcribed from [ĩ] to [ũ], according to the vocalic context. Their counts are not considered.

Others: The counts of "tsumaru oto" according to the right context permit to recover the counts of geminated non-nasal consonants. Likewise, the counts of /N/ followed by a kana from the NA- or MA-series permit to recover the counts of geminated nasal consonants: For example /N/ + NO is retranscribed [nno] where one geminated [n] and one [o] are counted. Sawaki also gives counts of the glottal stop [ʔ]. Finally, counts of a given sound in initial position are derived from the counts of '/' or '//' symbols followed by that sound.

The detailed counts per consonant type are given in Table 2. They are grouped by place and mode of articulation in Table 3. From a total number of 11,864 phones, there are 5,558 consonants, 6,244 "regular" vowels, and 62 nasal vowel realizations of /N/. Hence, 52.6% of vowels, and 46.8% of consonants (disregarding nasal vowel realizations of /N/).

### Consonant-Vowel sequences

An other interesting tabulation of the data consists in contextual counts. From Sawaki's data, it is possible to derive the distribution of vowels according to the preceding consonant. The distribution of vowels according to the place of articulation of the preceding consonant is shown in Table 4. The distribution according to the mode of articulation of the preceding consonant is given in Table 5. These counts do not include vowels that occur in initial position or after another vowel.

### Comparison with Bloch's data and Discussion

Recovering phonetic types from Bloch's data also causes some difficulty, since he presents counts by phonemic types. His phonemic types however, are close to phonetics. We first have to remove his "pitch phonemes" which are superimposed on segmental phonemes, and pauses (/#/) which he counted as phonemes. In the remaining phonemic categories, there is still some uncertainty as to how dispatch occurrences of the /N/ phoneme into appropriate phonetic types (since contextual information does not appear in Bloch's data). Likewise, his phoneme /h/ stands for either [ɸ] or [h] (he posed a phoneme /x/ for[ç]), /n/ for either [n] or[ɲ].We have used the proportions of these phonetic types from Sawaki's data to derive comparable counts from Bloch's data. Hence the tabulation of phonetic types in Table 6.

The vowel counts recovered from Bloch's data and from Sawaki's data are very similar, and so are the counts of consonants when grouped by place of articulation. In detail however, individual phonetic types have rather discrepant frequencies of occurrence. For example, [t] is the most frequent consonant type in Bloch's data, whereas [k] is the most frequent consonant type in Sawaki's data. The proportion of vowels is also lower in Sawaki's data (52.6% versus 53.6%), more in conformance with Onishi's data (49.9% as computed from "Zusetsu Nihongo"). Given the material used by Bloch and Sawaki, it is reasonable to rely more on the spontaneous speech data, which is also more abundant (about 12,000 phones compared to 2,000 in Bloch's material). In his article, Bloch himself stated that his data was "too small a sample to yield significant statistical information". That is why only the counts corresponding to the main phonetic types (the vowels), or to some grouping of phonetic types (consonants by place of articulation), seem to be reliable in Bloch's data.

Table 2. Detailed counts of consonants from Sawaki's data: For each consonant type, total count and count of geminated occurences.

b) *Labial consonants*

| CONSONANT | [p] | [b] | [m] | [ɸ] | [w] |
|---|---|---|---|---|---|
| total count | 7 | 123 | 364 | 26 | 192 |
| (geminated) | (1) | - | (10) | - | - |

c) *Dental consonants*

| CONSONANT | [t] | [d] | [n] | [s] | [z] | [ts] | [ɾ] |
|---|---|---|---|---|---|---|---|
| total count | 556 | 391 | 620 | 504 | 23 | 115 | 417 |
| (geminated) | (75) | - | (33) | (2) | - | (2) | - |

d) *Palatal consonants*

| CONSONANT | [tʃ] | [dʒ] | [ʃ] | [ʒ] | [ɲ] | [ç] | [j] |
|---|---|---|---|---|---|---|---|
| total count | 66 | 60 | 229 | 50 | 143 | 38 | 187 |
| (geminated) | (8) | - | - | - | - | - | - |

e) *Velar consonants*

| CONSONANT | [k] | [g] | [ŋ] | [w] | [ʔ] | [h] |
|---|---|---|---|---|---|---|
| total count | 823 | 25 | 261 | 2 | 53 | 141 |
| (geminated) | (11) | - | - | - | - | - |

Table 3. Consonants grouped by place and mode, from Sawaki's data. The "palatal" category covers alveo-palatals to palatals. The "velar" category covers velars to laryngeals. Frequencies of occurence are indicated within consonants alone, and within all phones.

| | PLACE | | | | | | |
|---|---|---|---|---|---|---|---|
| MODE | labial | dental | palatal | velar | subtotal | Frequency within cons. | Frequency (overall) |
| Plosive | 131 | 1022 | - | 912 | 2065 | 37.2 | 17.4 |
| Affricate | - | 140 | 134 | - | 274 | 5.0 | 2.3 |
| Fricative | 26 | 506 | 317 | 141 | 990 | 17.8 | 8.3 |
| Liquid | - | 417 | - | - | 417 | 7.5 | 3.5 |
| Nasal | 374 | 653 | 143 | 261 | 1431 | 25.7 | 12.1 |
| Glide | 192 | - | 187 | 2 | 381 | 6.9 | 3.2 |
| Subtotal | 723 | 2738 | 781 | 1263 | 5558 | | |
| % in cons. | 13.0 | 49.3 | 14.1 | 23.7 | | | |
| % overall | 6.1 | 23.1 | 6.6 | 11.1 | | | |

Table 4. Distribution of vowels according to the preceding consonant: place of articulation.

| RIGHT CONSONANT | VOWEL-TYPE /i/ | /e/ | /a/ | /o/ | /u/ | Total |
|---|---|---|---|---|---|---|
| Labial | 54 | 32 | 466 | 117 | 40 | 709 |
| Dental | 78 | 679 | 624 | 743 | 516 | 2640 |
| Palatal | 442 | - | 62 | 144 | 133 | 781 |
| Velar | 105 | 86 | 565 | 245 | 136 | 1137 |
| Total | 679 | 797 | 1717 | 1249 | 825 | 5267 |

Table 5. Distribution of vowels according to the preceding consonant: mode of articulation.

| RIGHT CONSONANT | VOWEL-TYPE /i/ | /e/ | /a/ | /o/ | /u/ | Total |
|---|---|---|---|---|---|---|
| Plosive | 79 | 436 | 782 | 555 | 139 | 1991 |
| Affricate | 86 | - | 25 | 15 | 125 | 251 |
| Fricative | 213 | 85 | 187 | 211 | 317 | 1013 |
| Liquid | 77 | 97 | 77 | 17 | 139 | 407 |
| Nasal | 224 | 179 | 435 | 374 | 12 | 1224 |
| Glide | - | - | 211 | 77 | 93 | 381 |
| Total | 679 | 797 | 1717 | 1249 | 825 | 5267 |

Table 6. Detailed counts of phonetic types from Bloch's data. Frequencies of occurence (in %) are indicated within vowels or consonants alone, and within all phones.

a) *Vowels*

| VOWEL | /i/ | /e/ | /a/ | /o/ | /u/ |
|---|---|---|---|---|---|
| within vowels | 18.8 | 11.9 | 34.1 | 27.6 | 7.6 |
| overall | 10.1 | 6.4 | 18.3 | 14.8 | 4.1 |

b) *Labial consonants*

| CONSONANT | [p] | [b] | [m] | [ɸ] | [w] | total |
|---|---|---|---|---|---|---|
| within cons. | - | 2.8 | 10.8 | 0.4 | 2.8 | 16.9 |
| overall | - | 1.3 | 5.0 | 0.2 | 1.3 | 7.8 |

c) *Dental consonants*

| CONSONANT | [t] | [d] | [n] | [s] | [z] | [ts] | [ɾ] | total |
|---|---|---|---|---|---|---|---|---|
| within cons. | 18.4 | 3.9 | 9.9 | 7.5 | 0.7 | 0.7 | 7.9 | 49.0 |
| overall | 8.5 | 1.8 | 4.6 | 3.5 | 0.3 | 0.3 | 3.6 | 22.5 |

d) *Palatal consonants*

| CONSONANT | [tʃ] | [dʒ] | [ʃ] | [ʒ] | [ɲ] | [ç] | [j] | total |
|---|---|---|---|---|---|---|---|---|
| within cons. | 1.6 | 1.8 | 5.6 | - | 2.1 | 0.3 | 2.1 | 13.5 |
| overall | 0.7 | 0.8 | 2.6 | - | 1.0 | 0.1 | 1.0 | 6.2 |

e) *Velar consonants*

| CONSONANT | [k] | [g] | [ŋ] | [ʔ] | [h] | total |
|---|---|---|---|---|---|---|
| within cons. | 12.4 | 0.4 | 5.3 | 0.4 | 2.4 | 22.7 |
| overall | 5.7 | 0.2 | 2.4 | 0.2 | 1.1 | 9.6 |

To conclude this short review, we still feel that a more extensive and systematic investigation should be undertaken. The two main pitfalls in previous studies must be avoided: The use of written material, and the use of kana transcription instead of phonetic transcription.

**Acknowlwdgment**

This study was partly supported by a grant of the JSPS. We also express our thanks to Dr. Misono and Dr. Imaizumi for their help in scanning the Japanese literature.

**References**

Bloch, B. (1950). Studies in colloquial Japanese IV: Phonemics. Language, 26, 86-125.

Fry, D.B. (1947). The frequency of occurrence of speech sounds in Southern English. Archives Neerlandaises de Phonetique Experimentale. 103-106.

Itahashi, S., Suzuki, H. & Kido, K. (1971). Discrimination of some consonants in words with the aid of dictionary. Transactions of the IECE of Japan, 54-C, (1), 10-17. In Japanese.

Jelinek, F. (1976). Continuous speech recognition by statistical methods. Proc. IEEE, 64, 532-556.

Kinda-Ichi, H. & Maes, H. (1978). Phonologie du Japonais Standard. Travaux de Linguistique Japonaise (Universite de Paris 7). Volume V (60 pages). In French.

Kramsky, G. (1966). The frequency of occurrence of vowel phonemes in languages possessing vowel systems of identical structure. Prague Studies in Mathematical Linguistics, 1.

Mines, M.A., Hanson, B.F. & Shoup, J.E. (1978). Frequency of occurrence of phonemes in conversational English. Language and Speech, 21, (3), 221-245.

Nakano, H. (1973). Gendai nihongo no onsorenzoku no jittai (Phoneme sequences in modern Japanese). In Denshikeisanki ni yoru kokugo kenkyuu V (Computerized research on the National language V), Research Report no 49 of the National Language Research Institute, pp. 94-120.

Onishi, M. (1937). Go'on hindo yori mitaru jukka kokugo hatsu'on kitei (The basis of pronounciation in ten languages, from the viewpoint of speech sound frequencies). Onsei no kenkyuu (Acoustic Research), 6.

Sawaki, M. (1980). Characteristics in mora-sequence in Japanese discourse. In A Re-

search in the Evaluation of Basic Linguistic Ability for Japanese Language Teaching: On Spoken Japanese in the Educated People (II), Research Report of the Kokuritsu Kokugo Kenkyuusho (The National Language Research Institute), 22-32. In Japanese.

Wioland, F. (1972). Estimation de la frequence des phonèmes en français parle. Travaux de l'Institut de Phonetique, 4, 177- 204. In French.

Zusetsu Nihongo (A Picture of Japanese) (1982). O. Hayashi (ed.). Kadokawa Press. pp. 314-315.