Metaphonological Abilities of Japanese Children

CARMELA SPAGNOLETTI, JOSÉ MORAIS, JESUS ALEGRIA and MARC DOMINICY[1]

Université libre de Bruxelles, Brussels, Belgium

ABSTRACT: Linguists generally conceive of speech as a sequence of elementary sound-units or "phones". On the other hand, cognitive psychologists have established the existence of a close link between segmental awareness (SA) and alphabetic literacy: the best speech segmenters are also the best readers; conversely, the lack of SA leads to lagging in the acquisition of the written code. Studies carried out with illiterate adults and preliterate children suggest that SA does not arise outside the context of learning to read and write. Furthermore, cross-cultural studies have so far supported the hypothesis that SA is promoted by alphabetic literacy only, and not by logographic or syllabic literacy. The present study is inspired by Mann's (1986) research and aims at studying the phonological (i.e. syllabic and phonemic) analysis abilities of Japanese first-graders who are learning to read in a syllabic writing system. The relatively poor performance of subjects tends to confirm the starting hypothesis that SA is closely linked to alphabetic literacy. Moreover, the subjects apparently resorted to spelling-based strategies to accomplish the tasks presented. This comes to support the general idea that the written characters used to transcribe a language tend to exert a strong influence on the conscious perceptual representation of the phonology of this language.

KEYWORDS: Speech analysis, phonological awareness, conscious perceptual representation of speech, literacy, Japanese children.

Several experimental studies have led to the conclusion that segmental awareness (SA), i.e. the ability to conceive of speech as a sequence of phones, is closely linked to literacy in an alphabetic system. The development of SA increases sharply once the reading acquisition process begins; conversely, the lack of SA leads to lagging in the acquisition of reading and writing.

The first piece of evidence in favor of this hypothesis was provided by Liberman et al. (1977): in a task consisting in tapping on the table as many knocks as there are phonemes in a word uttered by the experimenter, only 17% of kindergarteners reached the criterion of six consecutive correct responses whereas 70% of first graders attained the same criterion. Morais et al. (1979) carried out a study of the same vein with Portuguese illiterate adults who had been unable to benefit from school attendance in their childhood for purely socio-economic reasons. For a deletion task, in which the subject was required to repeat an utterance produced by the experimenter but without uttering the first phone, the mean percentage of correct answers was 17%. On the other hand, ex-illiterates, i.e. people who

Reading and Writing: An Interdisciplinary Journal **2**: 221—244, 1989.

had acquired literacy later in their life by attending adult-classes, were tested too and scored 73% for the same task. The inability of preliterates and illiterates in general to describe speech as a sequence of phones and to accomplish tasks which require explicit analysis of utterances at the segmental level, is consistent with the hypothesis that SA is a characteristic of literate populations.

Furthermore, cross-cultural studies have shown that it is not literacy in general which promotes SA but, more specifically, alphabetic literacy. For example, Read, Zhang, Nie and Ding (1986) have tested two groups of Chinese literates: one group was composed of people who had learned the logographic system only, while the subjects of the other group had learned the logographic system and the Roman alphabet. The procedure was exactly the same as the one used by Morais et al. (1979) with Portuguese subjects, and the results were strikingly similar: non-alphabetic and alphabetic Chinese groups' mean scores were 21% and 83% correct answers, respectively. Hence the conclusion that the alphabet triggers the emergence of SA while the logogram does not. A very probable reason for this divergence lies in the different ways in which the logographic and alphabetic characters map onto the phonological units. In the alphabet, the information is encoded at the phonemic level: morpheme x must be pronounced as phoneme 1 followed by phoneme 2 followed by phoneme 3, etc. The logogram does not go beyond morphemic decomposition; given that each Chinese morpheme has the dimension of the syllable, the information conveyed is simply that morpheme x must be pronounced as syllable y. Thus there is no need a priori for logographic literates to engage into segmental analysis.

Languages like Japanese have a third type of writing system, called a "syllabary". A syllabary tells us that morpheme x is pronounced as syllable 1 followed by syllable 2, etc. Given the current interpretation of the results reported before, we expect that Japanese literates who only use the syllabary will not exhibit SA. The first relevant experimental work on Japanese (Mann 1986) was precisely designed to verify this prediction. In the following, we will present a critical assessment of Mann's results and conclusion. Afterwards we will describe and comment two experiments of ours which were carried out in order to evaluate the impact of some factors not taken into consideration by Mann.

In order to clarify the whole discussion, it is essential to give a brief description of the Japanese writing system. This system includes three sets of traditional characters, two of which are syllabaries (Hiragana and Katakana) and a third one consists of the Chinese logographs (Kanji). Besides, the syllabic characters have been transliterated into Roman characters in order to facilitate international relationships: this gave birth to a fourth, alphabetic, set called Romaji which is rarely used. These various sets are not used indifferently. Kanji is used for the transcription of lexical words or roots, the inflected portion being written with "kana" (i.e. Hiragana or Katakana) characters. Hiragana is used for traditional Japanese words whereas Katakana is mainly restricted to words of foreign origin whose pronounciation has been adapted to the Japanese phonetic system. Onomatopeias are written with Katakana as well. These particular usages may suggest that Japanese people make use of Katakana as a phonetic transcription system. There is a one-to-one correspondence between Hiragana characters and Katakana ones.

The way children learn the characters deserves our attention since it is likely to account for some experimental results. As a matter of fact, Japanese children learn Hiragana by means of a matrix (see appendix) organized according to the principle that all the elements which belong to the same row share the same vowel. Similarly, all the characters appearing in the same column share the same consonant (at the phonemic level) except the first right hand column which consists of isolated vowels. All syllables of the matrix are of the CV type except for the isolated vowel characters and the isolated /n/ character (which constitutes in itself a rhythmic unit traditionaly called a "mora"). In school textbooks, colors have been added in order to facilitate the memorization: all the characters which include the same vowel share the same color. Great importance is given to that grid since dictionaries are organized according to it.

The Japanese orthography presents some additional properties which are of great interest for our concern. One of them is the fact that no different characters are assigned to the members of a voiced vs unvoiced consonantal pair: diacritic signs are used instead of specific characters. Two strokes added to the right hand top corner of the character indicate the voice feature (e.g. き /ki/ > ぎ /gi/). There exists a second sign, the small circle ○, whose sole function is to produce the voiceless labial plosive /p/ on the basis of the kanas listed in the /h/ column. So the basic sign は /ha/ yields ば /ba/ and ぱ /pa/. Another peculiarity is the existence of digraphs created to transcribe syllables which do not appear in the matrix. Note that the space (virtually a square) occupied by a digraph is the same as that required for a single character: the first sign is written full-sized and the second one is reduced and confined to the right hand bottom corner of the virtual square. Some digraphs represent diphonemic syllables, others triphonemic ones. The mechanism which governs the production of those signs supposes a manipulation at the phonemic level which is worth describing: the initial part of each digraph is the transcription of a syllable ending with the vowel /i/ and the final portion is the transcription of a syllable beginning with the semi-vowel /j/. The vowel /i/ can thus be interpreted as a link. In diphonemic syllables both the vowel /i/ and the semi-vowel /j/ disappear phonetically (e.g. しゃ : /ʃi/ + /ja/ > /ʃa/). In triphonemic ones the /j/ is preserved (e.g. びゃ : /bi/ + /ja/ > /bja/). Note that all triphonemic digraphs are of the CjV form.

In her 1986 paper, Mann started from the general consideration that two components are merged into the notion of phonological awareness, and that they are differently related to reading experience. Awareness of syllables would be less dependent on familiarity with the written medium and would develop naturally through maturation of the cognitive system. By contrast, awareness of phonemes would strongly depend on the experience of learning to read and write in an alphabetic system and on the teaching method used as well (phonics or whole word). The Japanese writing system being syllabic, a study of Japanese children's speech analysis abilities should put to test this hypothesis. According to her expectations, those children should not show segmental awareness until they have reached the end of the fourth grade, when they are taught the Romaji alphabet. The tasks administered were the current syllable and phoneme counting and deletion tasks. First through sixth graders took part in the experiments. Very high scores for syllable counting and deletion revealed their awareness of syllables. On the contrary, as far as phonemes were concerned, the results remained very low from first to third grade. Moreover, performances on both syllable and phoneme counting correlated with reading skills. A comparison with Liberman's American first graders suggested that children who had received alphabetic training were better phoneme segmenters than those who had learned the Kana syllabary. The influence of literacy was further evidenced by the fact that the children gave an extra tap for each word containing a digraphic syllable. Moreover, they also resorted to this "spelling strategy" in the phoneme counting situation: some children reported they had counted the number of kanas needed to spell the word and added an additional tap to arrive at the supposed correct answer. These first results tend to support the starting hypothesis that awareness of phonemes strongly relies on the learning of an alphabet. Yet, other data suggest that knowledge of the alphabet might not be the sole factor favoring the emergence of SA. Fourth graders were tested prior to their instruction in Romaji and their results displayed a clear improvement in phoneme counting (their performance was, however, only slightly superior to that of American first graders), and no sharp improvement was detected from fourth to fifth grade despite the introduction of tuition in Romaji. An important change seems to have occurred between third and fourth grade but Mann fails to provide a clear explanation for that. More surprising are the results yielded by re-entering pupils, i.e. children who had spent some time abroad and consequently had learned the alphabet either in English or in German. Those children should have been aware of phonemes since they knew the alphabet; yet their scores at counting phonemes were not superior, and even slightly inferior, to the scores of their normal class peers. As far as deletion tests were concerned, the scores were also high from fourth grade on and a type-of-item effect was observed: the results were strikingly better for /k/ than for /ʃ/ deletion in the Japanese groups, whereas no marked difference was detected in the American homologues. Mann put forward the hypothesis of a "character substitution" strategy based on a manipulation of the matrix. As a matter of fact, the /k/ column is immediately adjacent to the column of isolated vowels. It can be imagined that the child spelt the word and then replaced the initial kana by the character that lies directly on its right hand side in the matrix. As a consequence, the elision of /k/ would be easier to perform than that of /ʃ/ because the column corresponding to this last phoneme is located farther from the vowel column. Moreover, except for the /ʃi/ character, all other syllables beginning with /ʃ/ are represented by a digraph whose second component is situated at the very end of the grid. However, interviews with the subjects revealed that they did not resort exclusively to this strategy. Some of them described, in more or less precise terms, what Mann called a "phonological" strategy which first reduplicates the vowel of the first syllable in the word and then deletes the initial consonant-vowel portion (i.e. ki-pi → ki-i-pi (reduplication) and then ki-i-pi → i-pi (deletion)).

In short, Mann concludes that Japanese children are strongly influenced by their reading experience which makes them rely on the written form more than on sounds. She also observes that there are differences of performance between Japanese and American children: American children are better at tasks which involve manipulation of phonemes, whereas Japanese children are better at manipulating syllables, this finding being consistent with the idea that experience with an alphabet facilitates awareness of phonemes while the knowledge of a syllabary promotes the awareness of syllables. Nevertheless, the manipulation of syllables turned out to be easier than the manipulation of phonemes in both groups, suggesting that the familiarity with a certain type of writing system is not the only factor relevant to phonological awareness. The most unexpected result is surely the fact that while Japanese first graders could manipulate syllables and not phonemes, the majority of Japanese children were able to manipulate both syllables and phonemes by the age of nine, whether or not they had been introduced to alphabetic writing. Mann suggests as a possible explanation that experience with kana promotes the awareness of both phonemes and syllables because it is a phonological orthography (compared, for instance, to the Chinese logographic system which is not).

Mann's (1986) explanatory framework neglects two significant properties of the Japanese syllabary. The first one is the existence, in the matrix, of a character ん representing the isolated nasal consonant, beside a complete series of characters representing NV syllables (i.e. /na/, /ni/, /nu/, /ne/, /no/). This character ん never appears before a vowel. It can only be found in word-final position, or before a consonant, or before a glide, and its phonetic realization varies according to the following segment. Kuroda (1979) gives a good phonetic description which is worth quoting here:

> "The phonetic realization of the nasalized consonantal mora may best be described as a nasalized continuous transition from the preceding segment to the following one. When the following segment is a consonant [fn: The preceding segment is necessarily a vowel], the nasalized mora consonant has the same point and manner of articulation as that consonant; for example, before *p, b,* or *m,* the nasalized mora consonant is *m,* but before *s* it becomes a nasalized (and voiced) *s,* which we shall denote by *s̃.* If the following segment is not a (true) consonant, the nasalized consonant can be phonetically described as a nasalized vowel of varying degree of closure and point of articulation. It is probably somewhat closer and nearer-to-mid than the vowel preceding it. In word-final position, the nasalized mora consonant is again a nearer-to-mid nasalized prolongation of the preceding vowel or, possibly, *n* with a very light closure." (pp. 201—202).

Moreover, when the character ん occurs before an /n/, it provokes the lengthening of this subsequent consonant. The fact that the same consonantal sound is represented differently when it belongs to a CV syllable and when it does not (in which case it can be the first segment of a geminate cluster) might call attention to the fact that syllables can actually be separated into phonemic constituents. The second peculiar characteristic of the Japanese syllabary is the way in which it marks the voice feature of consonants (see above): the kanas which stand for syllables that include a voiced consonant are actually the same as the kanas that contain the voiceless counterpart with this difference that diacritic signs are added (except for nasal consonants, which are spontaneously voiced). This technique apparently implies an analysis at the sub-syllabic level, and more precisely, at the sub-phonemic level since it isolates distinctive features.

Moreover, Mann (1986) makes no reference to the processes used by the Japanese when they have to assign a spelling to words of foreign origin. There exist a great number of English and French words which have been adapted to Japanese writing and pronounciation. Consonant clusters are systematically separated by a vowel whose value is generally chosen according to the adjacent consonants. Yet there is no well defined rule and different speakers may adopt different pronounciations. The version which eventually appears in dictionaries is the first documented chronologically. Since the number of different syllables is very limited compared to the great variation which appears in alphabetic languages, Japanese had to create transcriptions for new syllables. The problem could be solved either by coining new characters corresponding to the new syllables or by combining pre-existing kanas in an adequate manner. The decision was to adopt the conservative strategy. The important point here is that this adaptation technique leads the users to manipulate kana characters in such a way that they have to segment the word at the phonemic level. For example, the word "disco" which includes the syllable /di/, absent in the Japanese set, is written with two characters: ディ in which

デ stands for /de/, and ィ, which represents the vowel /i/. But in this case, /de/ + /i/ is pronounced /di/. The Japanese child may thus be led to conceive of the written form of /di/ as including the consonant /d/ in quite the same way as Western children learn to read the alphabetic sequence "di" after they have been taught letter-sound correspondences. They have to ignore the vocalic sound necessarily attached to the consonant.

Notice that in the katakana representation of /di/ the first character is written in full size whereas the second one is reduced. This is to indicate that we are dealing with one single mora even if two kanas are written. Here again, individual variations exists: for instance the syllable /di/ is sometimes written "ジ" i.e. as /dʒ i/. The speakers who opt for such a transcription apparently rely on phonetic clues, since /dʒ / is the palatal affricate corresponding to the dental plosive /d/. The variation displayed in the forms which are encountered indicates that phonetic adaptation is a living and active process.

An increasing familiarity with such katakana manipulations might be a possible cause for the improvement in phoneme segmentation performances displayed by Mann's (1986) fourth graders. Indeed, as has been said earlier, a special syllabary, the katakana, is used for writing words of foreign origin and onomatopeias; it is thus employed whenever one has to focus on sounds. This second syllabary is introduced at the end of first grade.

To sum up, it appears that the Japanese writing system, which at first sight seems to limit speech analysis up to the syllabic level, actually includes some clues, or presents some characteristics, which might account for segmental awareness.

In order to have a better idea of the SA of Japanese children, we conducted two experiments which will be discussed below. These experiments were originally inspired by Mann's (1986) work. The subjects were regular pupils of the "Japanese School of Brussels", a state school which applies the education program used in Japan except that all children receive a one-hour-a-week instruction in French language. Pupils are separated into different language classes according to their previous knowledge of French. Those who have never attended a Belgian French-speaking school are classified as "beginners", those who have spent some time in a Belgian kindergarten attend the "intermediate" classes. Note that such labels are rather arbitrary, since there is no sharp difference between those two categories: children of both categories are alphabetic illiterates. Moreover, even from the point of view of the mastery of spoken French there seems to be no great difference since shifts from one class to the other, for personal reasons, are not rare. The third category is labelled "advanced" and is composed of those children who have previously attended a Belgian primary school for at least two trimesters and consequently have been taught the alphabet.

The first experiment concerned only "beginners" and was carried out during the first trimester, so that familiarity with the written form was reduced to the minimum. It should be noted, however, that most children receive some kind of instruction of Hiragana at home, while their knowledge of French is very limited, if not absent, and their progress in French very slow. According to the French teacher, her pupils are able to recognize and write about ten letters by the end of the academic year. As for Japanese language, children have learned all the Hiragana characters by the end of the second trimester, and Katakana as well as Kanji characters were introduced progressively. Thus, it can be inferred that the subjects tested during the second testing session, which took place between the end of the second trimester and the beginning of the third one, had a fairly good knowledge of Hiragana. The tasks were those generally used in this kind of study, namely syllable and phoneme counting, initial syllable and initial phoneme deletion. The same tasks were used by Mann, so that a comparison will be possible. Besides, a third type of task was administered, namely the classification task which consists in deciding which two utterances, among a series of three, sound similar. The interest of this task is that it allows us to determine whether a resemblance due to a common vowel is treated differently from a resemblance due to a common consonant, given that isolated vowels are assigned a kana in the matrix whereas isolated consonants are not represented graphically (except for the special case of the nasal /n/). Moreover, this kind of task does not involve the production level, contrary to deletion tasks.

The majority of items were pseudo-words in order to avoid possible misinterpretations of the tasks induced by semantic considerations.

EXPERIMENT 1

METHOD

Subjects

Subjects were 10 first graders (mean age: 6;5 years), all attending the "beginners" French classes. This group will be labelled B(1), after "beginners experiment 1" so as to be distinguished from the B(2) group of experiment 2.

Tasks, Procedure and Material

The tasks were: syllable and phoneme counting, initial syllable and phoneme deletion, and free classification. The items were in great majority pseudo-words. They were pronounced by a Japanese native speaker whose voice had been tape-recorded. Instructions were given in Japanese.

Training trials including feed-back corrective information always preceded test trials for which no correction was provided.

A. *Counting*

In the counting task, the subjects had to tap on the table as many knocks as they perceived units in the word presented.

The units to be counted were first syllables and then phonemes. Syllables had all V or CV structures and each item contained either one, two or three units. As for phoneme counting, the test items included ten V structures, four VN (where N stands for the isolated nasal /n/), fourteen CV, eight CjV (where "j" stands for the semi-vowel /j/), and six VCV words. The number of units per item was either one, two or three. For the V and VN words the number of phonemes corresponds to the number of written characters, whereas for all the other types the number of phonemes is superior to that of kanas.

During the training period, items were presented in a progressive order, the first word containing one unit, the second two and the third three. The experimenter pretended she wanted to play a word-game and showed the child how to play. First, a pseudo-word containing one syllable was uttered, then the experimenter stroke one knock. Then the experimenter asked the child to do alike insisting that he should repeat the word before tapping. Then followed a two-word progressive series and a three-word one. In order to make sure that the child had understood the task the examples were repeated in a random order and the child was asked to tap on the table. If an error was made, a correction was provided. Immediately after that, a series of 42 test items followed. The same procedure was employed for the phoneme counting task also.

B. *Deletion*

The deletion task consisted in repeating an utterance provided by the experimenter but without uttering the initial syllable or the initial phoneme.

The items were two- or three-syllable long for the syllable deletion series and one- or two-syllable long for the phoneme deletion lists. The syllable to be deleted was the vowel /a/, and the phoneme was /k/. A series of fifteen training items and a series of twenty test items were designed for each condition.

The task was disguised as a puppet game, as in some previous studies of our laboratory (Content 1985). The experimenter pretended that the utterances belonged to the language spoken by the animals living in the forest. Then she told them a little story to introduce the test. The puppets represented two characters, a very old and wise man and a lion. The old man uttered the items and the lion had to delete the initial syllable. A series of 15 examples was presented. Then the child was invited to play the lion's part and find the answers. The same series was repeated and if the child made an error, the experimenter gave the correct answer. This was followed by 20 test items. The same procedure was adopted for phoneme deletion.

C. *Free classification*

Free classification consisted in finding which two utterances, in a series of three, sounded similar.

For this task, 20 triads of CV syllables were created, 4 triads served as training items, the remaining 16 were test items. The three syllables were chosen in a way which allowed two acceptable types of classification. Two syllables shared one phoneme, either a vowel or a consonant. The third syllable was built in such a way that it was very far apart from one of the two previous syllables but presented some "global" resemblance to the other. For example, the triad /ni/-/me/-/ki/ contains two syllables which share the vowel /i/; thus one possible response might be /ni/-/ki/. This type of classification is labelled P, after "phonemic". However, since the vowel /e/ is very similar to the vowel /i/, and since /n/ and /m/ only differ by their point of articulation (/n/ is dental whereas /m/ is bilabial), a grouping of /ni/ and /me/ is also acceptable. This kind of classification is labelled S, after "global similarity". On the contrary, a choice of /ki/ and /me/ is considered "anomalous", and consequently labelled A, because the initial segments are very different (/k/ is a voiceless velar plosive whereas /m/ is a voiced bilabial nasal). In our experiment, the subjects were asked to group together those syllables which were "nearly the same". The training triads contained two identical syllables in order to allow a spontaneous choice between the phonemic and the global resemblance criterion during the test.

Colored cards on which appeared three identical drawings of familiar animals, each drawing occupying each vertex of an equilateral triangle, were presented. The instruction was: "Here are three animals, those who have the same name or nearly the same name are friends. Now listen carefully and find out who is whose friend". Correction was provided during the training trials, but not during the test trials. The common segment was the consonant for one half of the material and the vowel for the other half. In the same way, the variation of phonological features was either one of point or one of manner of articulation.

RESULTS AND DISCUSSION

Preliminary Remarks

Because of the very nature of the subject sample the data have been analyzed with non-parametric statistical techniques.

As for the counting task, the individual raw scores have been modified so as to suppress among the correct responses those which had actually been produced by guessing. As a consequence, a 0% score means that the subject's answer is at chance level, a negative score means that the subject responded incorrectly by following an alternative strategy.

Counting Tasks

All the subjects were successful when the counted units were syllables whereas no subject attained the criterion of six consecutive correct answers in the phoneme situation (cf. Table 1). For the latter case however, the analysis of responses indicates that the children resorted to a strategy other than counting phonemic units. Since we supposed that a spelling-based strategy was at work, we computed separate scores for each type of phonological structure as well as the distribution of number of taps per structure (cf. Table 2). This new approach reveals that scores are much better for V and for VN items than for every other type ($\chi^2(1)$ is significant at $p < 0.001$ in each case except for the pair V-VN for which $p < 0.01$). These results support the hypothesis of a spelling-based strategy, since for both V and VN items the number of phonemes matches the number of kanas needed to spell the items, which is not the case for CV, CjV and VCV items. In the remainder of this paper, the former type of items will be qualified as "matched", the latter as "unmatched". Table 2 clearly illustrates the presence of this systematic bias: when CV items are presented, a great majority of subjects respond with one tap (in Japanese a CV syllable is generally written with one kana); similarly, they strike two

Table 1. Median number of correct responses per task for the B(1) group

SYLLABLE COUNTING	PHONEME COUNTING				
	Phonological structures				
	V	VN	CV	CjV	VCV
100.0	95.5	85.0	−33.5	−28.3	−45.0
SYLLABLE DELETION	PHONEME DELETION				
98.0	67.5				

FREE CLASSIFICATION					
COMMON VOWEL			COMMON CONSONANT		
P	S	A	P	S	A
3.5	2.6	0.7	3.5	1.7	2.5

taps for CjV and VCV items (both are represented with two kanas and contain three phonemes). The bias for CjV items is apparently less marked than for VCV ones ($\chi^2(1) = 5.89$; $p < 0.05$). This is most probably due to the fact that although CjV and VCV items are both written with two kanas, the former are one-syllable long, and the latter two-syllable long. This structural difference is further indicated in the written form by reducing the size of the second kana in CjV syllables. But it is possible that for a first grade child, this peculiarity is not totally clear yet. That is probably the reason why there are more two-tap than one-tap responses for CjV items (cf. Table 2). There is however a relatively substantial number of three-tap responses as well; still we would not claim that the subjects counted the number of phonemes. A more probable interpretation seems to be that since all CjV items contain the glide /j/, the children might well have learned to respond with three taps whenever this sound occurred.

Table 2. Number of occurrences of 1, 2, and 3 taps as a function of the phonological structure for the B(1) group

Number of Kanas	Number of Phonemes	Phonological Structures	Number of Taps observed		
			1	2	3
1	1	V	**96**	3	1
2	2	VN	5	**34**	1
1	2	CV	115	**22**	3
2	3	CjV	20	45	**14**
2	3	VCV	3	57	**2**

The correct responses are printed in bold characters.

Deletion Tasks

Here again the performance is much better for syllable than for phoneme deletion (Wilcoxon $T = 0$; $p < 0.05$). Six subjects out of ten reached 100% score at deleting the initial syllable. On the other hand, only four children attained that level of performance in the phoneme deletion condition.

However, scores for phoneme deletion are higher than expected, which is all the more surprising since the phoneme counting task turned out to be quite impossible to accomplish. As already suggested by Mann (1986), this apparent ability to succeed in phonemic segmentation by deleting the initial phoneme might simply reflect an ability to manipulate kana representations and not segmental awareness. Indeed, the correct answer may be arrived at by resorting to the basic matrix of kanas: the initial kana of the test item would be replaced by the kana which, in the matrix, stands on its right hand side.

With the aim of determining which strategies could have been used by the subjects, an analysis of errors was carried out. It appeared that in the most frequent type of errors (27%) the deletion of the initial phoneme involved a modification of the subsequent vowel. This could result from a spelling-based strategy which yields a correct shift of column but an incorrect shift of row. The next most frequent types of errors are syllable inversion (16%) and the voicing of the initial voiceless consonant (11%). This latter modification could also be graphically based since it simply consists in adding a diacritic sign to the corresponding kana.

In short, it appears so far that the underlying mental mechanism involved in the phoneme deletion task might not be conscious segmental analysis but some other, possibly spelling-based, strategy. Moreover, the present study shows a markedly better performance at deleting phonemes than at counting them. This is at variance with other results (Alegria, Morais and D'Alimonte, in preparation) which suggest that, for Western children, deletion tasks are more difficult to accomplish than counting ones.

Free Classification Task

The results displayed on Table 1 suggest that the subjects responded according to some strategy since there is only a minority of anomalous answers. Furthermore, this difference is found to be significant. We used the Student t test to determine whether the subjects resorted to a strategy and if so,which strategy (P or S) was the most frequently adopted. The use of a strategy was detected by the calculation of the mean value of $(T - 3A)$, where T stands for the total number of responses, and 3A stands for the total number of anomalous responses. Since the mean value of $(T - 3A)$ was significantly different from zero ($t(9) = 2.32$; $p < 0.05$), we can conclude that the subjects responded according to some strategy. The next step was to test the preference if any of the P or the S strategy. For this purpose we considered the ratio $P' = (P - A)/(T - 3A)$, where P stands for the total number of "Phonetic" responses, and the ratio $S' = (S - A)/(T - 3A)$, where S stands for the total number of "global similarity" responses. Since $(P' + S') = 1$, the preferred strategy can be determined if P′ or S′ is significantly superior to 0.50. We chose an interval of confidence around the mean value of P′ (or S′) for the significance level of $p < 0.05$. The results showed no preferred strategy even though Table 1 shows a majority of P responses. We obtained no significant difference of treatment according to the nature of the common segment, be it vocalic or consonantal (Wilcoxon T = 20). The classification according to a common vowel was thus not facilitated despite the fact that the Japanese writing system contains special characters to represent isolated vowels (and not isolated consonants). One should remember that in this task the subjects were left completely free as for the strategy to use; they were only asked to detect which two syllables, in a series of three, sounded alike.

It seems at first sight that the subjects possess some sensitivity to sound similarity since they detect either a common segment or a more global set of sub-phonemic similarities. However, note once again the possible use of a spelling-based strategy. Indeed, the classification according to a common segment may be achieved by resorting to the basic matrix of kanas. More precisely, the classification according to a common vowel might consist in putting together those syllables which appear in the same row, and which share the same color; similarly, the classification according to a common consonant might consist in putting together those syllables which are listed in the same column.

In short, the present results evidence a better ability to deal with syllables than with phonemes. For the phoneme counting task, the separation of scores according to the phonological structure of the items reveals a spelling-based strategy which yields high scores when the number of phonemes matches the number of kanas and low, not random scores in the opposite situation. A similar effect had already been found by Mann (1986). The results are surprisingly good for the phoneme deletion test. This does not necessarily reflect SA: indeed, the spelling-based strategy can also account for it, given that the correct answer can be arrived at by resorting to the matrix.The /k/ column being immediately adjacent to the column of isolated vowels, children might well have interpreted the task as the replacement of the initial syllable of the item by the kana which occupies its immediate right hand side slot in the matrix. Colour might also be a facilitating factor since each syllable sharing the same vowel also shares the same colour in the matrix. Finally, a spelling-based strategy could also account for the ability to classify according to a common phoneme, as has been explained above.

EXPERIMENT 2

A second experiment was planned for various reasons. By the time of the second testing period, all Hiragana characters had been learned. Interesting information could also be derived if subjects belonging to the "advanced" French classes were included in the test sample.

SUBJECTS

Two groups of children were tested. The six subjects who already took part in experiment 1 were labelled B(2) (the remaining B children had left Belgium in the meantime). The second group included thirteen pupils of the "advanced" French class and was labelled A. All those 19 children (mean age: 6;8 and 7;0 years for B(2) and A groups, respectively) were first graders and had received the same program as far as Japanese language is concerned.

MATERIAL

The tasks to accomplish were not completely identical to those of experiment 1. Given that syllable awareness was clearly established, syllable counting and deletion tasks were not included in experiment 2. The phoneme counting task was maintained since scores had been very low during the first testing period.

As for phoneme deletion, new materials were added in order to verify the influence, if any, of the "proximity" and the "color" facilitating factors. In order to test the "proximity" factor, a series of training and test items including /p/ as initial phoneme was designed. In the matrix, syllables beginning with /p/ are located very far from the vowel column and include a diacritic sign. In order to verify the "color-influence" hypothesis, a third series of items was constructed. The initial phoneme of this series is /ʃ/. Except for /ʃi/, such syllables are written with digraphs; consequently, they do not appear in the basic matrix which Japanese children use at school. Besides, if those syllables were to be assigned a particular color, a mixture would be necessary since the characters used to write them correspond to syllables which include two different vowels. For example, the syllable /ʃa/ is represented with the kana for /ʃi/, written in full size, and the kana representing /ja/, written in reduced size (see above).

The classification task was also modified. In experiment 1, the free classification tested the children's spontaneous choice. It appeared interesting to examine what would happen in a constrained situation in which a specific choice would be suggested during training trials.

Reading tests were added so as to be sure that the subjects could master the Japanese writing system. The subjects were asked to read clearly but rapidly a series of fourteen "short" words (two syllables long) and a series of fourteen "long" words (three and four syllables long) written with Hiragana characters. The childen who had learnt the alphabet, namely the A group, were required to read additional equivalent lists of French words.

PROCEDURE

The procedure was exactly the same as in the experiment 1 except for the constrained classification task. This time a correction was provided when the child did not give the "phonemic" response during training trials. For the phoneme deletion task, as well as for the classification according to a common vowel or a common consonant, the order of presentation of the different series was balanced across subjects. The sequence of tasks was always counting — deletion — Japanese reading — French reading (for the A group) — classification.

RESULTS AND DISCUSSION

Reading Tests

As expected, Table 3 shows that the subjects' level of reading ability was fairly good. Even though A subjects seem to be slower and make more mistakes, the difference does not reach the significance level (Mann-Whitney U = 24 and 22.5 for "short" and "long" words, respectively). We can also observe that the subjects belonging to the A group could read the alphabet but they read very slowly and made much more mistakes than in Japanese.

Phoneme Counting

Table 4 shows that for both B(2) and A groups the spelling bias is obvious again, scores for counting V items being significantly much higher than those obtained for VN, CV, CjV and VCV ones ($\chi^2(1)$ reached the significance level at $p < 0.001$ for each pair except for the V-VN pair in the A group, in which $p < 0.01$). Interestingly enough, the correct response for VN items is less easily obtained than for V ones. This

Table 3. Mean reading speeds (RS), in seconds, and mean numbers of words read incorrectly (IW) per list of "short" and "long" words, for the B(2) and A groups

GROUPS	JAPANESE				FRENCH			
	Short words		Long words		Short words		Long words	
	RS	IW	RS	IW	RS	IW	RS	IW
B(2)	10.9 (3.6)	0.2 (0.4)	15.6 (6.5)	0.3 (0.5)				
A	14.1 (6.1)	0.3 (0.6)	23.3 (10.9)	1.0 (1.1)	44.2 (18.3)	6.2 (2.4)	104.4 (44.4)	11.4 (1.4)

NB. The standard deviations are in Parentheses.

Table 4. Median number of correct responses per task for the B(1)′, B(2) and A groups

GROUPS	COUNTING PHONEMES					PHONEME DELETION		
	V	VN	CV	CjV	VCV	/k/	/p/	/ʃ/
B(1)′	97.5	81.3	−34.6	−41.5	−45.8	67.5		
B(2)	95.0	50.1	−7.6	34.0	−25.0	88.3	86.7	92.5
A	95.3	16.4	9.8	−21.9	−38.5	95.0	75.0	70.0

GROUPS	CONSTRAINED CLASSIFICATION					
	Common vowel			Common consonant		
	P	S	A	P	S	A
B(2)	3.5	2.0	1.0	5.5	0.8	1.0
A	3.5	2.5	1.5	5.0	0.7	1.5

difference did not attain the significance level during experiment 1 though it was already present. The explanation for this difference of treatment most probably resides in the somewhat special status of the isolated nasal consonant. Indeed, as was said earlier, VN combinations can be pronounced either as nasalized vowel plus nasal consonant or as nasalized vowel alone. The subjects probably responded according to their own pronounciation. Yet if we look at the distribution of the numbers of taps (1, 2, or 3) per type of item (cf. Tables 5 and 6), it appears that all the errors produced by B(2) subjects consist in responding with three taps and that those produced by A subjects are equally distributed between one and three taps. Both facts tend to falsify the hypothesis formulated above. It rather seems that the children did not have a clear idea of how to treat such items.

Table 5. Number of occurrences of 1, 2, and 3 taps as a function of the phonological structure for the B(2) group

Number of Kanas	Number of Phonemes	Phonological Structures	Number of Taps		
			1	2	3
1	1	V	**58**	2	0
2	2	VN	0	**16**	8
1	2	CV	41	**38**	5
2	3	CjV	0	21	**27**
2	3	VCV	0	22	**14**

The correct responses are printed in bold characters.

Table 6. Number of occurrences of 1, 2, and 3 taps as a function of the phonological structure for the A group

Number of Kanas	Number of Phonemes	Phonological Structures	Number of Taps		
			1	2	3
1	1	V	**109**	20	1
2	2	VN	10	**32**	10
1	2	CV	89	**78**	15
2	3	CjV	16	55	**33**
2	3	VCV	3	52	**23**

The correct responses are printed in bold characters.

The comparison between B(2) and the remaining B(1) children shows an improvement of performance for CV, CjV and VCV structure items which is significant in each case ($\chi^2(1) = 21.26, 25.65, 12.88$, respectively; $p < 0.001$). Nevertheless scores are still very poor. No difference was observed for V and VN structure items for which the performance was very good, almost perfect in the former case.

In our opinion, this slight improvement is not to be attributed to some developing SA but to a task learning ability, given that B(2) children had already taken part in this test. The scores are much too low for "unmatched" items to support the former idea, especially as the analysis of errors reveals once again the use of a spelling-based strategy: in the majority of incorrect responses, the subjects resort to the corresponding number of kanas. As far as group A is concerned, the expected manifestation of segmental analysis abilities was not observed. Their scores were significantly better than B(1)'s scores for the "unmatched" items ($\chi^2(1) = 23.49$; $p < 0.001$ for CV; $\chi^2(1) = 10.43$; $p < 0.01$ for CjV, and $\chi^2(1) = 9.65$; $p < 0.01$ for VCV) but once again, their level of performance is very low. As for the comparison with B(2) subjects, there appears to be no significant difference except for CjV items ($\chi^2(1) = 7.27$; $p < 0.01$), for which scores were relatively high in the B(2) group. The most probable explanation for this seems to be the fact that all CjV items contain the glide /j/. The children might well have learned to respond with three taps whenever this sound occurred. Such a strategy yields the correct response for this particular type of time but has nothing to do with segmental awareness.

In short, Japanese first graders do not manage to count phonemes even though they have acquired a fairly good knowledge of syllabic writing as has been verified by means of reading tests. Moreover, the analysis of errors clearly shows the existence of a spelling-based strategy which yields correct responses for "matched" items and incorrect ones for "unmatched" items. A most surprising result is the fact that A subjects, although they had learnt the alphabet, did not perform better than the others. This point will be discussed below (see our general discussion).

Phoneme Deletion

B(2) subjects are apparently better performers than A ones but this difference is not statistically significant (Mann-Whitney U test, U = 25). They seem to have improved from experiment 1 as well. Unfortunately the small number of subjects made statistical analysis impossible: the number of matched pairs different from zero was too small to allow the application of the Wilcoxon test.

The most interesting aspect of this deletion test resides in its being composed of three sets of items of presumably different degrees of complexity (i.e. items beginning with /k/, /p/ and /ʃ/, respectively). We had hypothesized that if the subjects' strategy actually consisted in resorting to the matrix of kanas and then shifting to the right, the task would turn out to be more difficult to perform when the initial phoneme was part of the kanas listed very far from the vowel column (the phoneme chosen to verify this prediction was /p/). The other possibly intervening factor was color (all the kanas which share the same vowel also share the same color in the subjects' textbook). Thus the use of digraphs would suppress or at least reduce the influence, if any, of this facilitating factor. Moreover, digraphs do not appear in the basic matrix; consequently, if children actually resorted to the matrix, this strategy would be less efficient for such items (the phoneme chosen to verify this prediction was /ʃ/).

The statistical analysis carried out on the results (cf. Table 4) shows no significant difference according to group. A subjects are not better segmenters than B(2) ones despite their knowledge of the alphabet. The two groups will be merged for further analyses.

According to our expectations, scores should be better for /k/ deletion than for /p/ and /ʃ/ deletions, the latter being the most difficult to perform. These predictions are only partially confirmed: the performance for /ʃ/ deletion turns out to be significantly worse than for /k/ deletion (Wilcoxon T = 10.5; $p < 0.05$) but surprisingly enough, does not differ significantly from /p/ deletion (T = 26). This may reasonably be accounted for by the small size of the sample, which makes the statistical analysis difficult. Performances for /k/ and /p/ deletions do not differ significantly either (T = 22). This result suggests that "distance" is not a relevant factor. It could mean that the subjects were able to shift to the first column from any column, near or distant; or, simply, that they did not resort to the matrix of kanas but to some other strategy. The inferior, yet still good, performance for /ʃ/ deletion could suggest that when the initial mora does not appear in the basic matrix, the task is more difficult to accomplish. Whatever the exact strategy used, the analysis of errors reveals the use of a spelling-based strategy. Indeed a very frequent type of error consists in deleting the initial /ʃ/ sound but with the addition of a /j/ phoneme before the subsequent vowel (either correct or modified). This corresponds to 19.7% and 48.6% of the total number of errors for A and B(2) groups respectively. This kind of error is produced if the subject has interpreted the task as consisting in the suppression of the first kana from the initial digraph. For example, the initial mora /ʃa/ is written with the /ʃi/ kana plus the /ja/ kana, so that the deletion of /ʃi/ yields /ja/. Note that this type of error did not occur for the two other series of items.

Constrained Classification

Median values displayed on Table 4 suggest no difference in performance according to group. The statistical analysis (Mann-Whitney U test, $U = 29$ for common vowel and $U = 35$ for common consonant) shows no significant difference between B(2) and A subjects. We then decided to merge both groups for further analyses.

Contrary to what happened in the free classification of experiment 1, the subjects were trained to give as response the two syllables which share a common phoneme, either a consonant or a vowel. According to the median values (cf. Table 4), there is an apparent preference for a phoneme-based strategy. The statistical method used in experiment 1 has been used here in order to determine whether the subjects resorted to a strategy and if so, which strategy (P or S) was the most frequently adopted. Since the mean value of $(T - 3A)$ is significantly different from zero (Student $t(17) = 2.83$; $p < 0.02$ and $t(17) = 5.99$; $p < 0.01$ for common vowel and common consonant respectively), it can be fairly inferred that the subjects were not responding haphazardly. P answers are significantly more frequent than S ones for items sharing a common consonant ($p < 0.05$). As for the items which share a common vowel, the superiority of P responses does not reach the significance level.

Thus it appears that Japanese first graders are able to accomplish classification according to a common consonant when required to do so. This could reflect some sensitivity to the phonemic constituents of speech, or, once again, a faculty to find out a task specific strategy. Indeed, here too the correct response can be arrived at by resorting to the basic matrix of kanas: the classification according to a common consonant consists in putting together those syllables whose kanas are listed in the same column of the matrix. Similarly, the classification according to a common vowel consists in putting together those syllables which appear in the same row of the matrix. However, in the common vowel condition the preference for the P strategy is not marked, i.e. does not reach the significance level. This may be due either to sample characteristics or to the fact that, after all, two syllables containing different but close vowels are felt to be similar in sound quite as much as two syllables containing the same vowel are.

Intertask Correlational Analysis

It has been previously observed that scores were much better for phoneme deletion than for phoneme counting and that good performances at deleting the initial phoneme and classifying according to a common phoneme could be explained by the possible use of the basic matrix of kanas. Consequently, the performances for these two tasks should be correlated whereas none of them should be correlated to the performances for the phoneme counting task with "unmatched items". The Spearman rank correlation coefficients yielded by the analysis point to that conclusion: phoneme deletion correlates with classification according to a common phoneme ($r_s = 0.80$; $p < 0.01$); phoneme deletion and classification according to common phoneme taken together correlate with phoneme counting for V items ($r_s = 0.43$; $p < 0.05$); but they do not correlate with VN items ($r_s = -0.066$), nor with CV, CjV and VCV items taken together ($r_s = -0.24$).

GENERAL DISCUSSION

The results of both experiments 1 and 2 show a fairly good ability of Japanese first graders to accomplish segmentation tasks except for counting phonemes. Mann (1986) had arrived at the same conclusion and suggested that this was due to the syllabic nature of the Japanese writing system. This hypothesis implies that the awareness of phonemic units is an obligatory consequence of the awareness of syllabic units; yet there is so far no evidence that this is so. In our opinion, Mann underestimates, though she considers it, the importance of the orthographic information and the role of the basic matrix of kanas. The spelling bias appears rather clearly when a close analysis of responses is carried out. The existence of spelling-based strategies is established by several pieces of evidence. First of all, the distribution of scores by phonological structures shows that much better scores are obtained when the number of phonemes matches the number of kanas than when there is no such correspondence. Besides, the analysis of errors shows that the majority of incorrect responses correspond to the number of written units. Secondly, better scores were obtained for initial phoneme deletion than for phoneme counting while the reverse pattern had been found with Belgian children who were being taught to read with the whole-word method (Alegria, Morais and D'Alimonte, in preparation: 32.9% and 5.6% for phoneme counting and phoneme deletion, respectively). This tends to support the idea that the

Japanese subjects resorted to some special strategy. Thirdly, the subjects obtained better scores for the /k/ deletion task than for the /ʃ/ deletion task. Furthermore, the analysis of errors confirms that the subjects resorted to a spelling-based interpretation of the task. Finally, the intertask correlational results show that the phoneme deletion task is highly correlational results show that the phoneme deletion task is highly correlated to the classification according to a common phoneme ($r_s = 0.80$; tasks were only slightly correlated (Kolinsky 1984, $r = 0.28$; $p < 0.05$). The use of the matrix in both tasks explains, in all likelihood, the high correlation observed for Japanese children.

What is suggested by our study, as well as by Mann's (1986) results, is that Japanese first graders can manage to accomplish speech segmentation tasks by resorting to spelling-based strategies which do not require phonemic awareness. This supports, once again, the idea that the written characters used to represent a language tend to dominate the conscious perceptual representation of the phonology of this language. When subjects are questioned about the phonological constituents of speech without any explicit reference to written representations, they spontaneously resort to their knowledge of the writing system.

Another interesting aspect of the present results is that subjects who had learned the French alphabet (which was assessed by French reading tests), i.e. subjects belonging to the A group, were no better performers. The majority of those children made a lot of errors and read very slowly (cf. Table 3) but they apparently were aware of the sound-letter correspondence principle. Mann (1986) made the same observation about her re-entering fourth, fifth and sixth graders. One possible explanation could be that the reading tests always took place at the end of the testing session, just before the classification task, i.e. in a context where all the subject's attention was directed to the Japanese language (the instructions were given in Japanese, the items were Japanese pseudo-words, etc). It is conceivable, then, that the subjects responded according to their knowledge of the Japanese language and did not think of using their experience of another linguistic system. Thus it seems that the conscious perceptual representation of speech not only depends on the nature of the writing system, but could also be language specific. In order to verify such a hypothesis, we would have to design an experimental paradigm which would include items of both languages. For example, the material could be built up so as to present pseudo-words of both languages; obviously, the subjects should be informed of that peculiarity.

NOTE

[1] This study would not have been possible without the precious help of Miss T. Yamamoto for the construction of the experimental items. We are also much indebted to

Mrs. A. M. Van Craen for her good advice and for having introduced one of us (Carmela Spagnoletti) to the Japanese School of Brussels. We also wish to thank the Director of the Japanese School of Brussels for his permission to carry out the experimental work with his pupils, and to all the professors for having invited the parents to cooperate and provided us with many useful information. We are especially grateful to the children who participated to the experimental sessions as well as to their parents.

The present work has been partially supported by the Belgian "Fonds de la Recherche fondamentale collective" (convention n° 2.4562.86) and the Ministry of scientific Policy (Action de Recherche concertée: "Processus cognitifs dans la lecture").

Requests for reprints should be addressed to: José Morais, Laboratoire de Psychologie expérimentale, Av. Ad. Buyl, 117, B—1050 Bruxelles, Belgium.

REFERENCES

Alegria, J., Morais, J. and D'Alimonte, G. 1988. The development of speech analysis abilities and reading acquisition in a whole-word setting. In preparation.

Content, A. 1985. *L'analyse de la parole chez l'enfant.* Thèse de doctorat. Université libre de Bruxelles.

Kolinsky, R. 1984. *Différences individuelles dans le développement de l'habileté d'analyse phonétique explicite: recherche de facteurs explicatifs.* Mémoire de licence. Université libre de Bruxelles.

Kuroda, S.-Y. 1979. *Generative grammatical studies in the Japanese language.* New York and London: Garland.

Liberman, I. Y., Shankweiler, D., Liberman, A. M., Fowler, C. and Fisher, W. F. 1977. *In* A. S. Reber and D. L. Scarborough (eds), *Toward a psychology of reading* Hillsdale: Lawrence Erlbaum.

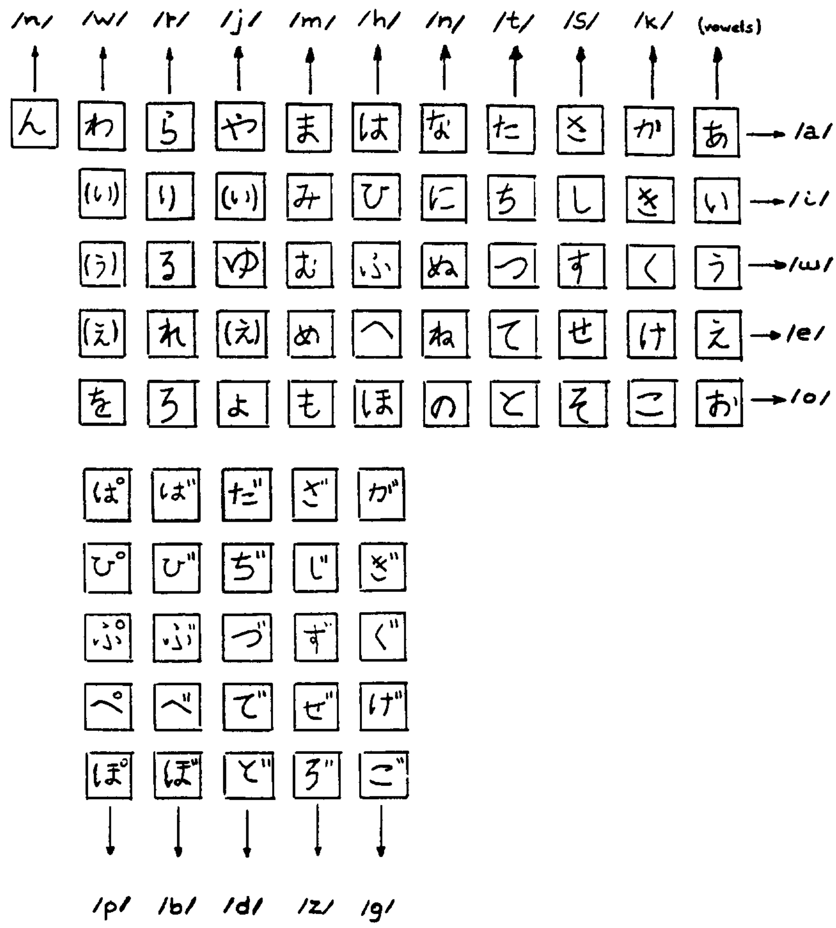
Mann, V. A. 1986. Phonological awareness: the role of reading experience. *Cognition, 24,* 65—92.

Morais, J., Cary, L., Alegria, J. and Bertelson, P. 1979. Does awareness of speech as a sequence of phones arise spontaneously? *Cognition, 7,* 323—331.

Read, C. A., Zhang, Y., Nie, H. and Ding, B. 1986. The ability to manipulate speech sounds depends on knowing alphabetic reading. *Cognition, 24,* 31—44.

APPENDIX

The Basic Matrix of Hiragana Characters



NB. 1. The phonemic transcriptions have been added by the authors of this paper.

2. The kana し is pronounced [ʃi] and not [si]
The kana つ is pronounced [tsui] and not [tui]
The kana ち is pronounced [tʃi] and not [ti]