DISTINCTIVE FEATURES OF KANNADA ALPHABET

G. PURUSHOTHAMA *
A. JAGADISH
P. J. KUMAR

ABSTRACT

Kannada alphabet letters were partially exposed to be predicted by 31 subjects. The confusions found were systematic. A list of Distinctive Features was drawn based on a confusion Matrix. In a sorting task of various CV and CCV letters it was found that even people unfamiliar to, could perceive the minimal pairs of the vowel allographs. The allographs can also be reduced to various distinctive features.

Reading is the process of extracting information from printed text. The meaningful units of print are words. It is still not clearly known as to how these units are processed to derive information. The views on this vary very much. At one extreme it is thought that letters are processed one by one and then words are perceived (Gough 1972.) At the other extreme the thinking is that not only the words are perceived as wholes but that the words can be expected even before they are actually seen. (Smith 1971). Both views have been supported, though not fully.

A similar problem seems to exist with the perception of. letters too. A simpler view is that there may be prototypes of letters in mind that are matched to the printed ones in the process of identification. But it has also been possible to explain the identification of letters on the basis of feature recognition (Spoehr and Lehmuhkle 1984.) Similar to the description of phonemes based on distinctive features, features of English letters have been described (Gibson 1969). There is reason to believe that the letters are perceived not as whole units but from parts thereof. Huey (1908) observed that a passage could have been read better when the bottom halves of the letters were removed than when the top halves of letters were removed. Massaro, et al. (1980) found that the legibility of English letters differ depending on the nature of their form.

* AIISH, Mysore-6

It is necessarry to understand the preceving of letters before the nature of reading is understood. The intent of this study was to find the features of Kannada letters that contribute to their recognition. A brief description of Kannada letters is in order. Kannada has a phonetically regular syllabic script. Whereas the English letters represent phonemes, Kannad letters represent syllables. Kannada has an alphabet of 50 letters. The following are the letters classified for the present purpose.

Kannada Vowels :

Short Vowels : ಅ ಇ ಉ ಋ ಎ ಒ

Long Vowels : ಆ ಈ ಊ ೠ ಏ ಓ

Dipthongs : ಐ ಔ

It may be observed that many of he Iong vowels are related to short vowels in their form. All the consonants in the alphabet are written with the vowel [a] Whereas most of them are written with ;he along graph — for the vowel [a] others have a zero allograph.

Consonants with the allograph :

ಕ ಖ ಗ ಚ ಛ ಝ ಠ ಡ ಢ
ತ ಥ ದ ಧ ನ ಪ ಫ ಭ ಮ
ಯ ರ ಶ ವ ಷ ಸ ಹ ಳ

Consonants with the Zero allograph:

ಅ ಙ ಜ ಞ ಟ ಣ ಲ ಬ

It may be noted that unlike the English, letters Kannada letters have round envelope. More than half the number of the letters have an allograph ligatured at their top. Thus it was interesting to see as to which part of the letters are important for their identification.

EXPERIMENT I :

Subject : There were 31 subjects of age range 17-30 years. The subjects had practiced reading Kannada for 12 years in school and were adept in reading Kannada print.

Procedure : One way of finding the features is to find the confusions in reading the letters. Confusions can be brought about by the degradation of the print or by inappropriate exposure. As the invent was also to find the parts right, left, top and bottom of the letters contributing to the identification the task of

reading with partial exposure was used. Letters were printed in the centre on cards of size 13 cm by 16 cm. The subjects were instructed that they would see parts of letters and were required to predict and name them. The letters were shown through an opening in a screen so that only half of a letter was exhibited at a time. Each letter was exhibited until a response was made and then removed. The responses were recorded. Fortyeight letters of the alphabet was presented in a random order. All the halves of letters ; right, left, top and bottom were shown to each subject.

Results : Different halves of letters exposed ; right, left, top and bottom, were treated as four attempts at reading a particular letter. Thus each letter had partial exposures for 124 times. It can be seen from the confusion matrix in Table 1 that letters have been confused to those which look similar and also to those with common features. For example the letter ಅ [a] has been confused to ಛ [ra] and ಚ [ta] as well as ಆ [a:] which looks similar. It can be seen that it is the curve at the bottom part of these letters that has caused the confusions. The matrix presents such widespread confusions. The letters having similar features in the same part have been confused. In Table 2 it can be observed that the confusions of the different parts of letters have been systematic. The confusions of different halves of letters are given separately in columns. In each row the numbers in front of letters present the number of confusions (maximum 31) of letters printed next to them. It was observed that the bottom parts of the letters were most confused. Least number of confusions were observed in the top halves of letters. This observation was inspite of the repeated occurrence of the feature '◡' on a large number of letters. The features on the left half and right half of the letters were also confused.

The legibility of Kannada letters is also rank ordered in the Table 2. The number preceding each letter in the first column is the number of correct identifica. tion of letters out of 124 partial exposures. For example the letter ಈ [i :] was identified 119 times correctly and no part of it was confused with any other letter of alphabet. Even though the subjects were informed that the task was of identifying letters of the alphabet, sometimes the letters were misidentified as letters other than those of the alphabet i.e., blended with allographs of other letters.

On the basis of confusions observed a list of distinctive features of Kannada alphabet was made (Table 3). The list of confused letters based on different features is given in Table 4.

Discussion : The bottom halves of letters are more confusable than their top halves. It is interesting to note that, like in English as Huey (1908) had

Table 2. Confusions of different letter parts

NO. OF CORRECT IDENTIFICATION OUT OF 124 PARTIAL EXPOSURES	No. OF CONFUSIONS TO DIFFERENT SIDES of LETTERS			
	RIGHT	LEFT	TOP	BOTTOM
119 ಈ				
110 ಗ			4ರ 2ಕ	
109 ನ				
108 ಳ				
107 ಝ		2ಯ		
106 ಣ	2ಉ 2ಊ 2ಪ			
104 ಉ				
102 ಟ				3ಎ 2ಛ
100 ಇ		4ಇ 2ಐ 2ಶ		7ಇ
99 ಊ	4ಋೂ	11ಉ		
98 ಹ				4ದ 3ಬ 2ಓೂ
97 ಕ				13ಳ
92 ಜ				9ಒ 4ಉ ಓ 3ಐ
91 ತ	9ದ 2ಕ			9ಲ 3ಗ
89 ಷ		14ಪ 8ಫ 2ಫ		3ಪ 2ಫ
89 ೂ		15ಋ		
87 ಠ		2ವ	5ಥ 4ದ 2ಯ	4ರತ
87 ಋ				
85 ಸ	12ದ 2ವ		6ಮ 3ಶ	10ಸ
85 ಐ	2ಐ			12ಒ 7ಜ 2ಔ
83 ರ	5ದ 2ಫ		3ಗ 2ಕ	4ತ 2ಕಳ
82 ಕ	21ರ			3ಳ 2ರ
81 ಫ		21ಫ		7ಫ 4ಬ 3ಒ
81 ಔ	2ಐ ಟ			26ಜ
80 ಖ	2ಮಯ	10ಖೂ 6ಋ		
79 ಡ		26ಧ 2ಧ		7ವ 2ಧ
78 ಚ	16ಕ 2ತ			6ದಲ 3ಜ ಓ 2ಒ
77 ಥ	4ಡ	5ಥ 2ಡಧ		14ಥ 3ಫಭರ
75 ದ	4ವ 3ಜ 2ಧನ	4ಧ 2ರ	5ದ 2ಜ	4ಒ 3ಜ 2ಓಬ
75 ಲ	3ಉ	6ಅ		8ರ 6ತ 2ಅ
74 ಅ		12ಅ		14ತ 5ರ 2ಛ
73 ಎ	7ದ 4ಏ 3ವ			9ವ 4ಏಐ 2ಧ
72 ಧ	2ವ 2ಛ		18ದ 2ಥಅ	2ಛಥಭರ
72 ಬ	16ಎ 3ವ 2ಯ	5ಛ		14ದ
72 ಐ	6ಹ			9ವ 6ಏ 5ಎ 2ಪಟ
71 ಮ	11ಯ 3ಋ 2ಖ	14ವ		
68 ಅ		19ಅ 2ಡ		15ರ 5ತ 3ಲ
67 ಓ	21ಒ 2ಜ			6ಜಒ 5ಜ
65 ಇ		9ಇ 5ಜ 2ಕಖ		26ಇ
64 ಯ	10ಮ			8ಮ
63 ಒ	12ಓ 3ಇ 9ಎದ 3ವ 2ಬ	11ಬ 4ಛ 2ಜ		8ಜ 5ಒ 2ಐ ಳ 10ವ 3ಖ ಎ 2ಬದಟಪ
61 ಏಫ	3ಥ 2ಫ 11ಛ 3ಧ			18ಛ 3ಫ
60 ಛ	19ದ 3ಥ 2ಫ	4ಒ 2ಬ		6ಧ 4ಥ 2ಬ ಕಛ

Table 3a. Distinctive features of Kannada alphabet

L/F	◡	(	)	⌐	O	|	—	ω	∽	~	⌒	ಎ	3	ə	·	ε		
ಅ	•	•	•	•	•													
ಆ	•	•	•		•		•											
ಇ			•						•	•								
ಈ	•	•	•	•	•	•	•				•							
ಉ	•	•	•		•						•							•
ಊ	•	•	•		•						•							•
ಋ			•	•				•						•				
ೠ			•	•				•						•				
ಎ			•		•			•			•							
ಏ			•		•	•		•			•							
ಐ			•		•			•	•									
ಒ		•			•			•										
ಓ		•			•	•		•										
ಔ		•		•	•	•	•	•										•
ಅಂ	•	•	•	•	•													•
ಅಃ	•	•	•	•	•													•

Table 3b. Distinctive features of Kannada alphabet (Continued)

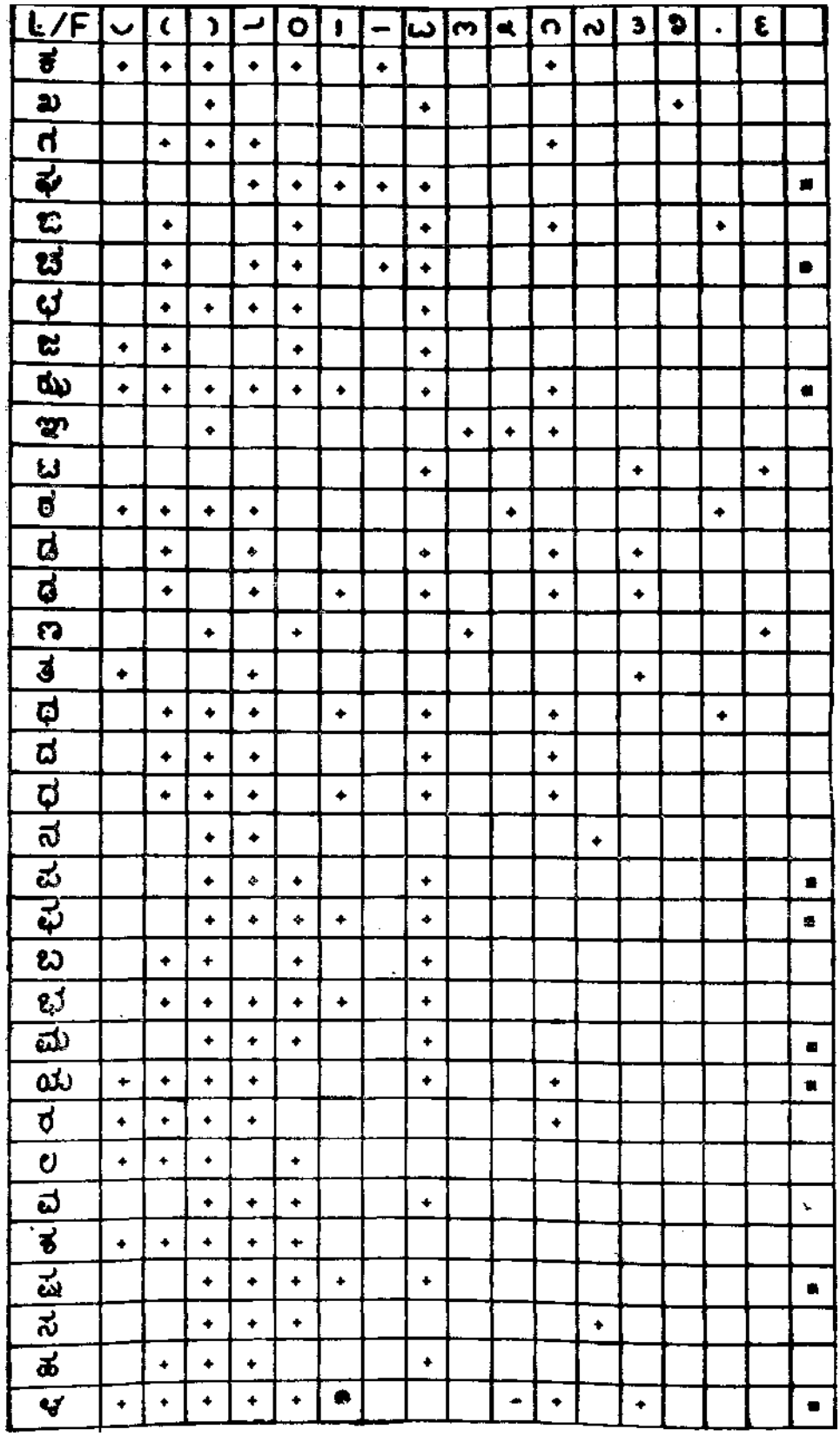

Table 4. List of confused letters based on the distinctive features

೮	ಲ ಠ ರ ತ ಕ ಳ ಅ ಅ ಶ
೩	ದ ಬ ಹ ಜ ಒ ಓ ಉ ಷ ಫ ಘ ಛ ಧ ಥ ಭ ಢ ವ ಐ ಏ ಙ ಔ ಟ ಪ
೭	ಝ ಯ ಡ ಧ ಢ ಅ ಅ
೨	ಣ ಉ ಊ ಪ ಖಾ ದ ವ ಫ ರ ಯ ಮ ವ ನ ಎ ಏ ಖ ಬ ಹ
೫	ನ ಸ
೪	ಡ ಶ ತ ಚ ಕ ಢ
೯	ಞ ಇ ಐ
೪	ಛ ಡ ಥ ಫ
೧	ರ ಗ ಕ ದ ಡ ಠ ಥ
೪	ಇ ಞ

noted, when the lower halves are covered Kannada letters can be read easily than when their upper halves are covered. The top halves seem to have important features which are less confusing. The confusability of right and left halves seem equal. As mentioned earlier, the Kannada letters have round envelopes and the bottom halves are formed of either one curve or two curves with an occasional conditional marker. Similarly, curves on right and left sides have also caused confusions. In addition to the smaller and bigger curves there are many other features which are noted including the position and direction of curves and gaps. The task of partial exposure was useful in determining the features in the respective parts of the letters.

Kannada script is syllabic. The consonant letters of alphabet are modified by ligaturing the Vowel Allographs to form syllables. Thus reading Kannada involves not only perceiving the alphabet letters but also other allographs of them Ligatering rules for the allographs are fairly regular.

The following is an example of consonant ರ [ra] with various vowels.

ರ ರಾ ರಿ ರೀ ರು ರೂ ರೆ ರೇ ರೊ ರೋ ರೌ

And the following are examples of different geminated consonants.

ಕ್ಕ ಡ್ಡ ದ್ದ ತ್ತ

It was needed to see if the vowel allograph could be perceived independent of the consonant, based on their forms. As the script is phonetically regular it might be expected that even the subjects not exposed to Kannada will be able to find the differences precisely among allographs and set them into minimal pairs. Another experiment was carried out for this purpose.

EXPERIMENT II

Subjects : There were eight subjects of age range 17-32 years. The subjects were American nationals who were not exposed to Kannada script any time. The subjects volunteered to undergo the experiment.

Materials : Various ligatured consonants were printed on cards 7 cm. by 12 cm. which were handy for sorting. There were three consonant letters (ಗ, ಮ, ದ. with various vowels (CVs) and 15 geminated consonants written on the centre of cards. Altogether there were 48 letters used in the experiment.

Procedure : **A** task of sorting was used for classifying the letters based on their forms. The set of cards of randomly ordered letters was given to each subject and the subject was asked to group the similar looking letters in to as many groups as they found appropriate.

Generally, the subjects sorted them into atleast six groups. The groups mainly included the letter combination of the three consonants. The subjects were asked to further sort two of the bigger groups into as many groups as possible.

Results : The outcome of sorting tasks were interesting. The subjects sorted the different CV letters (ಗ, ಮ, ದ, with various vowels) into individual groups. The geminated letters were put into other three or more groups. The grouping of the geminated letters were dependent on the features similar among them. The following are some of the pairs.

ನ್ನ ಜ್ಜ ಣ್ಣ ಲ್ಲ ಪ್ಪ
ಸ್ಸ ಖ್ಖ ಟ್ಟ ಕ್ಕ ವ್ವ

In the sec.nd round of sorting two of the CV letter groups were given for sorting. The subjects consistently grouped the CV letters which can be considered as minimal pairs together which differed by single features. There were few exceptions which included pairing of CCV with a similar looking CV. The following is an example of the grouping.

a	b	c	d	e	f
ಮೀ	ಮಿ	ಮು	ಮಾ	ಮೋ†	ಪ್ಪ*
ಮೇ	ಮ	ಮೂ	ಮೊ	ಮೌ	ಮ್ಮ

a	b	c	d	e	f
ದು	ದೆ	ದೀ	ದೂ	ದ	ದೋ*
ದಾ	ದೌ	ದೇ	ದೊ	ದಿ	ದ್ದ

It was clear from the sorting tasks that one can perceive the allographs in minimal pairs very well even without knowing the orthography. It can be seen from the examples of sorting that they could sort the pairs of letters with only one feature difference. The only exceptions have been CV and CCV pairs which were probably inevitable.

Discussion : The sorting tasks revealed that perception of the Kannada vowel allographs has been possible depending on the distinctive features of the allographs It can be noted that all the vowel allographs can be reduced, to the different features shown in Table 3.

General Discussion : Owing to their form, the bottom halves of Kannada letters are more confusing. Even English letters when printed in lower case letters

---

• Different CV pairs

t Pairs differing by more than one feature.

generally having round envelopes, have more information in their top halves. Depending on the confusions that have been observed among different halves of Kannada letters the distinctive features have been delineated. It should be noted that some Kannada letters have features repeating within them (Table 3).

Vowel allographs of Kannada letters were also seen as combination of distinctive features. It was observed that the same list of features can serve for the allographs too. The sorting task showed that even people unfamiliar with Kannada script could arrange the minimal pairs. The list of distinctive features of Kannada alphabet can be thought as a useful meterial in the related works to come on Kannada reading.

Acknowledgement : We thank our Kannada friends and the non-Kannada American friends who volunteered to be the subjects for the study.

BIBLIOGRAPHY

Gough, P. B. (1972). One second of Reading. In Kavanagh and Mattingly (ed) Language by Ear and by Eye. Cambridge, MIT Press.

Smith, F. (1978). Understanding Reading, New York : Holt Rinehart and Winston.

Spoehr, K. T. and Lehmuhkle, S. W. (1982) Visual Information Processing, Sanfrancisco, W. H. Freeman and Company.

Gibson, E. J. (1969), Principles of perceptual learning and development. New York ; Appleton Century Crofts.

Huey, E. B. (1908). The Psychology and Pedagogy of Reading, New York: Macmillan.

Massaro, D. W, Taylor, G. A. Venezky, R. L. Jastrzembski, J. J. and Lucas, P. A. (1980). Letter and word preception. Advances in Psychology. New York: North Holland Publishing Co.