How Arabs Read Roman Letters

Mick Randall, *West Sussex Institute of Higher Education and*
Paul Meara, *Birkbeck College, London University*

This article describes three experimental studies in which native Arabic speakers were asked to perform a simple search task involving Arabic or Roman letters. The results show that Arabic speakers react to Roman letters in the same way as they react to Arabic letters. However, this is very different from the way native English speakers react to Roman letters. The implications of this difference for teaching reading in English to Arabic speakers are discussed.

INTRODUCTION

Reading in a foreign language has been the object of greatly increased research activity in recent years. For obvious reasons, much of this activity has centred on the relatively higher-order skills of discourse organisation and the interpretation of continuous text. Yet for many of the world's learners, the problem is not so much that of understanding at the text level as processing at the word level, and this problem is one that must be particularly acute for learners whose L1 is written in a script which is radically different from the Roman script used to write English. Surprisingly, although there is a large amount of work on the psycholinguistics of the English writing system, the question of how readers whose L1 does not use the Roman script approach the task of decoding writing has received very little attention.

In this paper we will be looking at the way native speakers of Arabic process strings of Arabic letters, and comparing this with the way that they process strings of Roman letters. There is, of course, plenty of anecdotal evidence that native Arabic speakers experience severe difficulties in processing Roman script, although reliable experimental studies are much harder to come by (cp. Haynes 1981). Even where experimental evidence is available, it is generally concerned with establishing that Arabs do indeed experience difficulties, and we are not aware of any work which has attempted to explore what the nature of these difficulties might be.

In this paper we are taking as our starting point a series of studies by David Green (Hammond and Green 1982; Green, Hammond and Supramamian 1983) who investigated the way that native speakers of English process arrays of letters and nonalphabetic shapes. This work showed that native English speakers process arrays of letters and digits differently from the way in which they process strings of

---

Mick RANDALL is a Senior Lecturer in TESOL at the West Sussex Institute of Higher Education. He has worked extensively with Arabic speakers, first in Kuwait as an English teacher, then in the Sultanate of Oman as an Inspector of English, and recently as a teacher trainer in England. Paul MEARA is Head of the Department of Applied Linguistics at Birkbeck College, London University (London WC1 H0PD). He is best known for his work on vocabulary acquisition in an L2 and has published extensively in this field. He would really like to be a world-famous musician.

shapes or letter-like forms which are unfamiliar. Green's task is basically very simple: on each experimental trial, his subjects were presented with a single letter or shape as a target, and after a brief delay they were shown an array of five similar targets. The subjects' task is simply to state whether the target occurs in the array or not, and the time it takes for this response to be produced is measured. For nonalphabetic shapes, the results are very straightforward. Targets are recognised fastest if they lie in the middle of the array, and much more slowly if they lie at the ends of the array. This pattern of responding gives rise to a characteristic *search function* which is U-shaped.

The results for letter targets and digit targets are rather different. With these stimuli, native English speakers produce the fastest response to targets on the left of the string. There are also slightly faster responses to targets in the middle of the string and to targets in the rightmost end of the string. Thus, the characteristic search function that emerges from this material can best be described as a tilted M-shape.

Both these characteristics of the letter search function can be explained in terms of what we know about the psycholinguistics of English. English words are written from left to right, and this obviously accounts for the left-to-right slope. The beginnings and ends of English words seem to play a prominent part in word storage and word recognition (Bruner and O'Dowd 1958; Horowitz, White and Atwood 1968), and this may account for the fast reaction times at the beginnings and ends of the arrays. However, the actual reason for the shape of the search function is not really important here. The thing to note is that native English speakers seem to develop special ways of processing arrays of letters. The most likely explanation for this phenomenon is that it has something to do with the way English native speakers approach the task of reading.

The purpose of the experiments described in this paper is to investigate the way that native Arabic speakers search strings of letters, and to see if there is any degree of similarity between the search strategies of speakers using the Roman alphabet and those using Arabic script. It will also look at the way that native Arabic speakers search strings of Roman characters to see if they adopt similar or different search strategies when working in an L2, and (in particular) the way that Arabic speakers change their search strategies as their familiarity with the L2 increases.

There are a number of reasons why we might expect native speakers of Arabic to develop search functions which are systematically different from the search functions that are characteristic of native speakers of English. Firstly, Arabic script is written from right to left, rather than from left to right. At the very least, we would expect this to produce right-to-left search strategy in its readers. Secondly, Arabic is an example of what Sampson (1986) calls a *consonantal language*. In normal Arabic script the vowels are not marked and the reader has to supply the correct vowelling from an understanding of the context in which the word appears. This fact might well affect the way that words are stored in the mental lexicon with a resulting difference in search strategies, although the precise nature of this effect is difficult to specify in advance. Thirdly, Arabic words are usually quite short, the vast majority of the words being derived from a triliteral root by the addition of prefixes and suffixes and occasionally infixes. The majority of the words in Arabic are less than six characters long and even the words which consist of five or six characters would usually consist of a three-character root, plus a prefix or a suffix or both. Again, this may also produce different search strategies in native Arabic speakers: affix-stripping in Arabic (i.e. searching for the word's root) may be a more general process than it is in English. Finally, Arabic characters are much less redundant than Roman letters: in some cases, only the presence or absence of a single dot above or below a standard character shape distinguishes between letters. This could well lead to the adoption of quite different search strategies in native Arabic speakers who are looking for small differences between letters.

All these differences suggest that we might expect characteristic search patterns for Arabic to be different from what we find in English. The experiments reported in the next section investigate this claim formally.

THE EXPERIMENTS – GENERAL OVERVIEW

Three experiments are reported in this section. All the experiments were run on similar lines, using a BBC microcomputer to present visual stimuli via a visual display unit (VDU) and to record responses. Reaction times (RTs) for correct "yes" responses were collected, stored, averaged and printed out at the end of the experiment along with the total number of correct responses and the total number of errors for each individual. Only in experiment 1b were individual trial-by-trial RTs collected.

Procedure

Each experiment was conducted in an Amcom econet studio. The subjects were all given instructions in English in which they were told that they would be shown a target letter followed by a set of five letters, and that their task was to say as quickly as possible whether the target letter appeared in the set of five letters or not. They were to respond by pressing one key for "yes" and another key for "no". In all experiments except for 1b the subjects responded "yes" with the right hand and "no" with the left. In 1b, half responded "yes" with the right hand and half responded "yes" with the left hand. The experimental trials were divided into four blocks of 100 trials, with results being collected at the end of each block. Subjects were allowed a pause after each 50 trials. Before beginning the experimental trials, each subject had 20 practice trials.

Design

The target letter was present in the string on a randomly selected half of the trials, and absent on the other half. When the symbol did appear in the string, it was located equally often at each of the five array positions.

Materials

In experiments 1a and 1b a set of 25 Arabic letters were used as stimuli. In experiment 2, a set of 25 Roman letters were used. The Roman letters consisted of a set of 25 upper case letters generated by the computer on a matrix of dots 8 rows deep by 8 columns wide. The Arabic letters were generated on the same matrix. At a viewing distance of 18 inches, the target array subtended a visual angle of 3 degrees.

Analysis

The basic design provides a set of five RTs for each subject for each type of letter, corresponding to the five possible target positions. These five data points can fall into four simple patterns. They can either represent a straight line (a linear pattern); they can form a U-shaped curve (a quadratic pattern); they can form an S-shaped or Z-shaped curve (a cubic pattern); or they can form an M-shaped or W-shaped curve (a quartic pattern). These patterns can also be combined to form other patterns. For example, the normal pattern for English native speakers scanning English upper case letters is an upward-sloping M-shape, a combination of a linear pattern with a quartic pattern. Analysis of variance and a technique called orthogonal decomposition enables one to establish which patterns are present in a particular set of data.

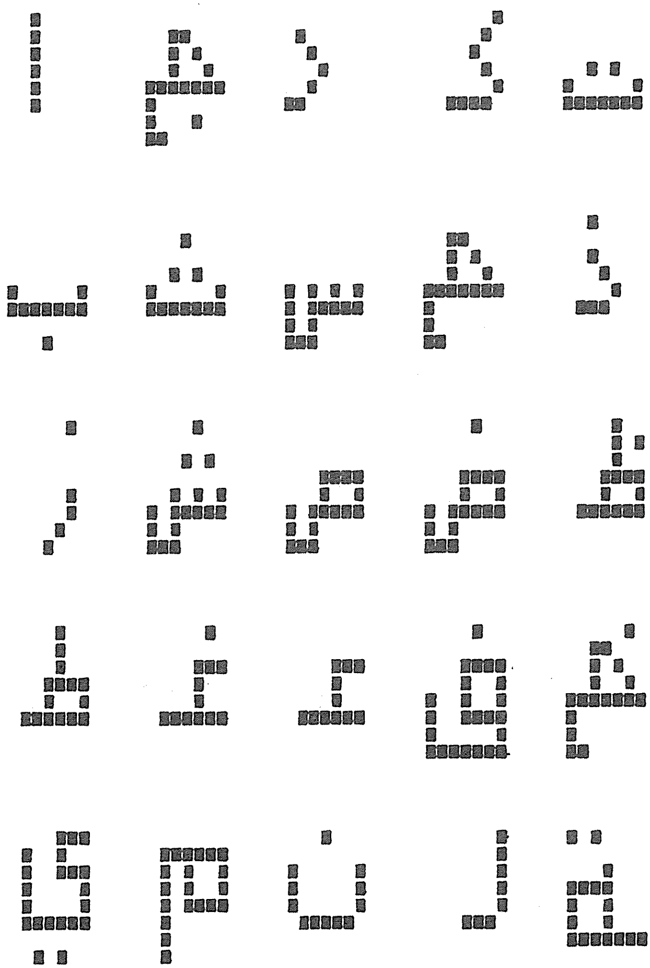
THE EXPERIMENTS

Experiment 1a

Method

Subjects: 21 Saudi Arabian trainee teachers were used in this experiment. All were primary school teachers in Saudi Arabia and all were following a two-year course at the West Sussex Institute of Higher Education, Bognor Regis to train as middle school English language teachers.

Materials: The Arabic character set used in this experiment and in Experiment 1b was as set out on p.137. Devising a character set for Arabic presents a number of problems. Arabic has no upper and lower case letters such as English has. Each letter can however take one of two or (more rarely) three forms depending on its immediate environment within a word. Additionally, there are a separate set of letters which are used if the letter is to be used by itself (e.g. in dictionary headings, paragraph headings, etc.,) and the set of 25 characters used in this experiment was based on this set of independent forms.



A rather different problem concerns the complexity of the characters. As we discussed above, Arabic characters are considerably less redundant than Roman letters, and in order to achieve the necessary degree of clarity within the 8 by 8 matrix available on the BBC microcomputer, certain characters had to be enlarged and were thus out of scale with the majority of the other characters. The effect of this distortion is difficult to assess. However, only characters which were identified with 100% accuracy by a separate group of native Arabic speakers were used as stimuli in this experiment.

Results and Discussion

The analysis reported below is based on the averaged RTs for each individual over 400 trials. Generally error rates were low (less than 7%) and there was no evident trade-off between speed and accuracy.

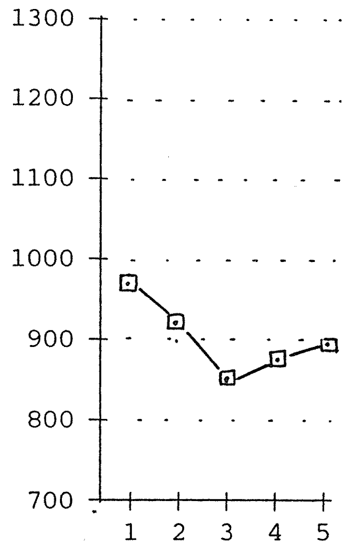
Table 1 and Figure 1 show the mean RTs to correctly identified targets as a function of target position. These results were submitted to an Analysis of Variance in which the main effect was Target position. This effect was highly significant ($F(4,80)=10.8$, $p<.001$). Detailed analysis of this effect revealed that it was largely made up of two subcomponents. The pattern of results is basically U-shaped ($F(4,80)=17.45$, $p<.0005$), but superimposed upon this is a right-to-left linear component ($F(4,80)=14.82$, $p<.001$). Taken together, these two components, which are both clearly visible in Figure 1, account for 87.5% of the total position variance. There was also a small quartic component which accounted for 11% of the variance and which was weakly significant ($F(4,80)=5.25$, $p<0.05$).

Thus it would seem from this group of subjects that reading direction is an important factor in the way that native Arabic speakers search arrays, just as it is in the case of speakers whose L1 uses the Roman alphabet (RAs). However, Arabic native speakers would appear to differ radically from RAs in other details of the scanning process. The U-shaped curve produced by this group of subjects is the typical curve produced by RAs when they search arrays of shapes.

Table 1 (Experiment 1a): Mean reaction times (msecs) to stimuli in five target positions

	target position				
	1	2	3	4	5
mean rt	968	927	851	877	892
st dev	199	206	186	220	181

Fig. 1: Experiment 1a



Experiment 1b

Method

One of the factors that is noticeable from the previous work carried out by psychologists on visual searches and scanning is that the subjects for their experiments are nearly all undergraduate students and thus are all relatively fluent readers. It could well be that the effects that have been observed in other experiments could only be applicable to fluent readers and would not necessarily hold for less fluent readers, though work with children carried out by Green et al (Green, Hammond and Supramamian 1983) suggests that the effects seen in English native-speaker scanning in fact start from a very early age. The subjects in Experiment 1a were generally not academically high achievers, and had generally slow reading speeds; it was felt that the effects observed could be due to factors of academic and reading sophistication. Experiment 1b was undertaken to test whether the results obtained in Experiment 1a were also found in native Arabic speakers with a more sophisticated academic background.

Subjects: 8 Algerian graduate students who were studying English at Chichester College of Higher Education prior to taking up places on post-graduate courses in England. All had followed Arabic-medium studies throughout their schooling and their first degree studies.

Materials: The set of 25 characters was as for Experiment 1a.

Design: The design followed that of Experiment 1a, except that half of the subjects responded "yes" with the right hand (the dominant hand) and half responded with the left hand.

Results and Discussion

As individual trial-by-trial RTs (reaction times) were collected it was possible in this experiment to correct the mean data for any exceptionally long response latencies from individual subjects. The mean RTs for correct responses were corrected by rejecting any RT which deviated more than 2.5 standard deviations from the group mean.

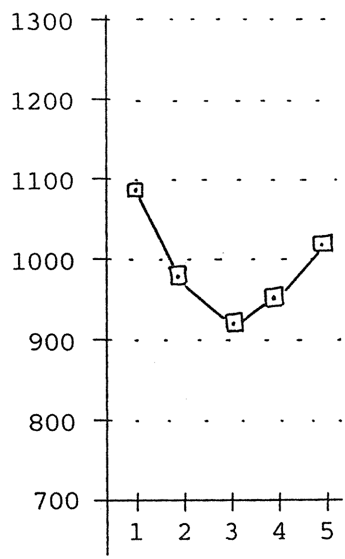
Table 2 and Figure 2 show the arithmetic mean RTs for correct responses plotted as a function of target position. As in the previous experiment, error rates were low (less than 2%) and there was no apparent trade-off between speed and accuracy.

These results were submitted to an analysis of variance in which the main effect was target position. The analysis showed that there was a highly significant position effect ($F(28,4)=25.21$, $p<.001$). This effect was largely accounted for by an extremely significant U-shaped component ($F(28,4)=478.70$, $p<.001$), which accounted for 80% of the variance, and a significant linear component ($F(28,4)=18.89$, $p<.005$) which accounted for 19% of the variance. Essentially, this result is the same as the previous one: it seems to indicate that academic sophistication was not a major factor in the effects observed. On the contrary, both groups of Arabic speakers tested appear to scan arrays of Arabic letters in a right-to-left direction as would be expected from the reading direction of Arabic, and both groups produce a very strong U-shaped search function similar to the function produced by RAs when they search arrays of non-letter shapes.

Table 2 (Experiment 1b): Mean reaction times (msecs) to stimuli in five target positions

	target position				
	1	2	3	4	5
mean rt	1090	978	921	945	1014
st dev	104	133	134	116	138

Fig. 2: Experiment 1b



Experiment 2

If Arabic native speakers scan arrays of Arabic characters in such a different way to RAs, it is interesting to ask whether Arabic speakers transfer these strategies to the reading of English, or whether they adopt different strategies when scanning arrays of English characters. It is also interesting to ask whether Arabs learning English alter their scanning strategies as they become more fluent speakers/readers of English. In this experiment, therefore, we report a longitudinal study of a group of native Arabic speakers over a period of a year of intensive English language tuition.

Method

Subjects: 12 Saudi Arabian trainee teachers from the West Sussex Institute of Higher Education who were following a language access year of English tuition amounting to a total of 900 hours. Subjects had previously followed a one-year course of English Tuition in Riyadh. The subjects are a sub-group of those included in Experiment 1a. The Arabic data was collected on a separate occasion after completion of this longitudinal study.

Materials: A set of 25 upper-case English letters (see above for details).

Design: Data was collected at three points throughout the year (November, February and June). Prior to each data collection, the subjects' reading speeds in

English were estimated by administering three different passages from the introductory section of 'Faster Effective Reading'.

Results and Discussion

Table 3 and Figure 3 show mean reaction times to correctly identified stimuli by target position. As before, error rates across the three tests were generally low (less than 4%); as in the other experiments, there was no evident trade-off between speed and accuracy.

These data were submitted to an analysis of variance in which the main effects were target position and Test Time. The analysis showed that there was a significant general decrease in RT with each successive test ($F(22,2)=19.96$, $p<.001$). This could be an experimental effect in that the subjects were becoming more familiar with the experiment and the way of responding, or this could be due to the fact that their growing familiarity with English characters improved the speed at which they were able to identify English characters in an array. The fact that the three tests were carried out over a year with at least three months between each test would tend to favour the hypothesis that most of the effect is due to increasing familiarity with the English script.

The analysis also revealed a significant target position effect ($F(44,4)=7.01$, $p<.001$), which failed to interact in any way with the test time variable. In all cases, this effect is composed almost entirely of a highly significant U-shaped component, which accounted for between 75% and 96% of the variance. This was the only significant component on all the tests, except for a cubic component on Test 1 which accounted for 19% of the variance on that test. It is interesting to note that although on each test the rightmost letters were recognised faster than the leftmost letters, on no test did this produce a significant linear component.

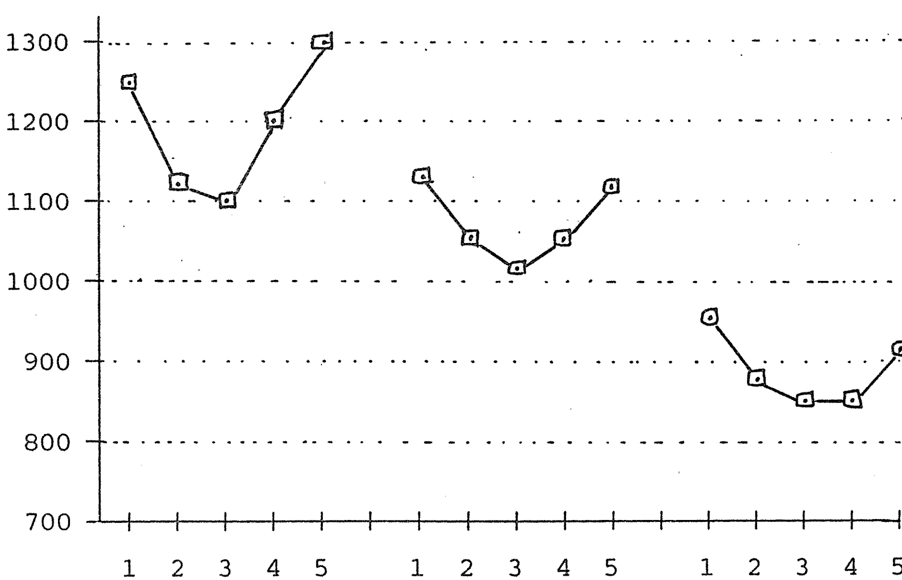
The estimates of the subjects' reading speeds throughout the year shows a systematic improvement from an average of 52 w.p.m. with 27% comprehension on the first test, to an average of 85 w.p.m. with 73% comprehension on the third test. While these reading speeds are extremely slow by native speaker standards, or even by the standard of foreign students who are familiar with the Roman Alphabet, they nevertheless represent a significant improvement by the group over the year.

Thus, it would seem clear from this experiment, that this group of Arab learners did produce some changes in their search strategy when presented with English characters. The linear right-to-left component which was a significant feature in both the other experiments is no longer a significant feature when they scan arrays of English characters. However, the most stable and significant feature of both the scanning of Arabic characters and Roman letters is the characteristic U-shaped component. This group of learners did produce faster reaction times as they became more familiar with English and as their reading speeds increased, but there was no sign in the data that they were adopting the characteristic M-shaped curve produced by RAs.

Table 3 (Experiment 2): Mean reaction times (msecs) to stimuli in five target positions on three test occasions

	target position				
	1	2	3	4	5
mean rt t1	1244	1118	1102	1200	1299
mean rt t2	1131	1056	1014	1051	1112
mean rt t3	963	883	868	865	913

Fig. 3: Experiment 2



CONCLUSION

In these experiments, we have shown that native-speaking Arabic readers produce search functions which are radically different from the search functions of readers whose script uses the Roman Alphabet. Some basic similarities are to be found, in that reading direction in both writing systems has an effect on scanning processes, but in other respects the processes are quite different: the processes used by Arabic readers are more akin to the processes used by RAs when searching arrays of shapes. It also appears that this search strategy is an enduring one which is transferred to tasks involving searching arrays of Roman letters.

The reasons why Arabic speakers adopt search strategies which are so different from what we find in RAs are beyond the scope of this paper and will need more investigation. One possible explanation could be the radical differences between Arabic and Roman characters, with the lack of redundancy in the former forcing the reader into a search for detail which is unnecessary when dealing with the Roman script. This attention to detail was noticed in early cross-cultural work on the Rorschach Ink blot test with Moroccans (Bleuler and Bleuler 1935), and it could be that the linguistic and the cultural phenomena are interlinked. Alternatively, the difference could be due to the nature of the Arabic writing system, rather than the forms of the individual letters that implement it: the fact that the orthography is a consonantal system rather than an alphabetic one may induce different search patterns. Finally, it is possible that the lexical structure of Arabic affects search of written strings: most words are based on tri-consonantal root forms, which are shared by large numbers of forms belonging to a single semantic area, and this structure is very different from the simple Root and Affix structure found in most European languages.

However, what is of importance for those interested in teaching reading to non-native speakers of English is that the strategy of an Arabic learner faced with an English word is very different from that of the native English speaker. The direction of script and the resultant difference in eye movements have long been recognised in the field, and many pedagogic materials contain some materials such as cartoons which aim to help students cope with the different processes involved in learning to read the English script. This is only part of the story, however. Presumably RAs do not use U-shaped search strategies because there is some good psycholinguistic reason why these strategies do not work for English. If Arabs continue to rely on inefficient strategies, as these experiments indicate, they will always be faced with some difficulties on the level of word processing which native speakers of English do not experience. Clearly, Arab learners will need considerable training if they are to be taught more efficient word-handling strategies. The experiments reported here suggest that Arab learners do not naturally evolve strategies more closely related to those used by native speakers of English as their proficiency in English increases.

REFERENCES

Bleuler M. and R. Bleuler (1935) Rorschach ink-blot tests and social psychology. *Character and Personality*, 4, 99–114.

Bruner J. and D. O'Dowd (1958) A note on the informativeness of parts of words. *Language and Speech*, 1, 98–101.

Hammond E. and D. Green (1982) The detection of targets in letter and non-letter arrays. *Canadian Journal of Psychology*, 36, 67–82.

Green D., E. Hammond and S. Supramamian (1983) Letters and shapes: developmental changes in search strategies. *British Journal of Psychology*, 74, 11–16.

Haynes M. (1981) Breaking the alphabet barrier. Mimeo: Michigan State University.

Horowitz L., M. White and D. Atwood (1968) Word fragments as aids to recall: the organisation of a word. *Journal of Experimental Psychology*, 76, 219–226.

Sampson G. (1985) *Writing Systems*. London: Hutchinson.