Illiteracy: A cause for biased cognitive development

ALEXANDRA REIS and ALEXANDRE CASTRO-CALDAS
Language Research Laboratory, Centro de Estudos Egas Moniz, Hospital de Santa Maria, Lisbon, Portugal
(Received July 15, 1996; Revised December 3, 1996; Accepted March 17, 1997)

Abstract
Learning to read and write generates new rules within the language processing systems. These new rules significantly change the manner in which some operations are performed. This finding was studied, by comparing the performance of literate and illiterate persons in several tasks. It was found that illiterate individuals (1) had difficulties in repeating pseudowords, (2) were worse at memorizing pairs of phonologically related words compared to pairs of semantically related words, and (3) were unable to generate words according to a formal criterion. Illiterate persons use strategies that are good for semantic processing, but inadequate for phonological analysis, while literate individuals are able to use several parallel running strategies. (JINS, 1997, 3, 444–450.)

Keywords: Illiteracy, Language, Cognitive development

INTRODUCTION
In the basic neurosciences, experimental animal models have been developed in which the substitution or deletion of a single gene at the beginning of embryonic formation results in changes in the adult biological configuration of the animal. We borrowed this idea and imported it into the field of cognitive neuroscience. The adaptation can be done as follows: (1) the building up of a cognitive function depends, at least partially, on the acquisition of specific basic skills; (2) if one of the basic skills is not acquired, in the proper moment of the developmental process, the final function will be distorted; and (3) the performance of participants in tasks related to that function will depend on the fact of having learned or not learned the specific basic skill. The model we selected to test this hypothesis is the natural model of illiteracy, and its influence on oral language and cognition in general.

First, there must be a careful selection of the case material. Second, the methodology should be planned according to a theoretical hypothesis. In relation to the case material, a target population is needed that can provide a good control. It is very easy to confound illiteracy with other cultural aspects and thus the findings, comparing groups, will reflect several differential aspects of culture rather than literacy as a single piece of knowledge (or basic skill).

Some authors (Lecours et al., 1987a, 1987b, 1988) and even some of us (Damásio et al., 1976a, 1976b) published interesting findings comparing groups for which cultural aspects other than literacy were not considered. In non-brain-damaged samples, differences were apparent in multiple neuropsychological tests, as for example visuospatial and memory abilities (Ardila et al., 1989), and language and praxic abilities (Rosselli et al., 1990). Differences in the oral naming, matching, and repetition of words by aphasics and their controls were found by Lecours et al. (1988). However, these differences can not be attributed only to literacy as a single factor but may be related to a more general cultural deficit. Naming and matching tests in the study of Lecours et al. (1987a, 1987b, 1988) were based on two-dimensional representation of objects and actions. As we have shown, illiterate and semiliterate individuals without brain lesions have difficulty decoding two-dimensional drawings, and even photographs (Reis et al., 1994). This has to be understood following the same methodology that we are now using to study language (Castro-Caldas et al., 1996).

Our work is based on the study of a particular fishing community in the south of Portugal. Fifty years ago, schools in this country were scarce, and it was often difficult for children to attend classes. Most of the older girls in each family were unable to acquire a formal education, because they were needed at home to care for younger siblings and deal with some of the basic activities of daily living. The younger children had fewer responsibilities at home, so they were sent to school, but the family unit grew together in the same cultural environment. It is very common, nowadays,
to find sisters with ages ranging from 50 to 70, living close by and leading similar lives, the older being illiterate and the younger literate. They have equal lifestyles and modes of conversation.

Our first approach to this community was to select a homogeneous population based on intelligence and cultural aspects (assessed with the appropriate instruments) that could be separated into two subgroups simply by the fact of knowing the phonemic value of a set of graphemes. All of our studies are thus based on the comparison of these two subgroups, otherwise matched for both intelligence and cultural environment.

We can now describe the theoretical hypothesis that is to be tested in these carefully selected populations. The first steps of language acquisition were similar for all of our sample. They develop normal phonological skills and semantics. The difference starts at the ages of 6 to 7 years, when they are ready to start learning to read and write. Learning to match graphemes and phonemes is learning an operation in which units of auditory verbal information heard in temporal sequence are matched to units of visual verbal information, which is spatially arranged. This type of treatment of auditory verbal information modulates a strategy in which a visual–graphic meaning is given to units that are smaller than words, and thus independent of their semantic representation. Children need to be aware of phonology in order to segment the continuum of the oral language. After being isolated as significant units, these segments are introduced sequentially in a working memory system (Baddeley, 1995) that is present from the early stages of language acquisition, and to which a new content of visual experience is added. If we, as normal adult readers, are asked to spell a word, we evoke a visual image of its written form. The awareness of phonology also allows us to play with written symbols (which can be transcribed to sounds) to form plausible words, independently of semantics. Therefore, learning to read and write introduces into the system qualitatively new strategies for dealing with oral language; that is, conscious phonological processing, visual formal lexical representation, and all the associations that these strategies allow.

Illiterate individuals do not develop these strategies and, therefore, they do not stimulate the neural substrate to subserve them. Taking into account a possible general theory of brain function, we would expect new strategies added to previously existing ones to be processed in parallel (Mesulam, 1990).

The following hypothesis were explored in this investigation:

1. The absence of a conscious processing of phonology was demonstrated in the work of Morais et al. (1979). These authors showed that illiterate persons had difficulty dealing with explicit phonological processing. On the other hand, Castro (1992), showed that, for some aspects of implicit processing, illiterate individuals were similar to literate ones. Repetition of pseudo-words is a task that requires phonological processing, probably involving both implicit and explicit mechanisms, forcing a phonological route as proposed by some authors (Caplan, 1992; Martin & Saffran, 1992).

2. The absence of a reference system related to the form of the words constrains the strategies of behavior mainly to a framework of semantic content of the information. One would expect that, confronted with problems that can be solved both by form and by semantic content, illiterate individuals would clearly rely on the latter.

3. The combination of the two former aspects must have an important impact on the way the lexical experience is stored by illiterate individuals in comparison to literate ones. Therefore, tasks of verbal fluency based on formal aspects and based on semantic ones ought to reflect such a difference.

EXPERIMENT 1–REPETITION OF PSEUDO-WORDS

Materials and Method

A list of 24 highly frequent words (Nascimento et al., 1987) was selected for repetition. Fifty percent of the words were disyllabic and the remainder were composed of words with one syllable (three words), three syllables (seven words), four syllables (one word) and five syllables (one word). Frequency, rather than word length, was chosen as the selection criterion for the words used in our task. A list of 24 pseudo-words was designed on the basis of the first one, changing the consonants and maintaining the length and the vocalic structure of the original word. Words and pseudo-words were randomly mixed in a single list and read by the examiner, who asked the participant to repeat exactly what she heard. Scoring was based on the total number of correct responses on repetition, and the quality of the error was also noted.

Research Participants

Thirty female participants from the same social and cultural environment volunteered to take part in this investigation: Twenty were illiterate (mean age = 60.9 years), and 10 were literate (mean age = 58.1 years, mean years of schooling = 9.7). Participants were considered illiterate when they had never attended school for social reasons, when they were unable to identify any kind of written material (except in some cases, their name), when they had never had a job requiring familiarity with written material, and when they were born of illiterate parents. Participants were considered literate when they showed acceptable performance on tests of reading and writing. Their intelligence and cultural background were assessed by means of a questionnaire in which pragmatic aspects of daily living were considered. Both groups were judged as similar concerning these aspects.
Results

A two-way ANOVA was performed (Table 1 and Figure 1) for Group (literate vs. illiterate) × Word Type (words vs. pseudo-words). Results showed a literacy group effect $[F(1,56) = 19.74, p < .0001]$; a word type effect $[F(1,56) = 20.49, p < .0001]$, and an interaction between the two factors $[F(1,56) = 16.22, p = .0002]$. A post hoc analysis (Scheffé test) was done to verify the factors that accounted for the significant effects. Results revealed the following significant interactions: Illiterate participants repeating words versus repeating pseudo-words ($p < .0001$); illiterate participants repeating pseudo-words versus literate ones repeating pseudo-words ($p < .0001$); illiterate participants repeating pseudo-words versus literate ones repeating words ($p < .0001$).

It may be concluded that illiterate participants repeated words similarly to literate ones, and that repeating words was not different from repeating pseudo-words for literate participants. In relation to the pseudo-words, illiterate participants made more errors, and tended to transform pseudo-words into meaningful words—26% of the errors were transformations from pseudo-words into meaningful words. This was difficult to evaluate in the group of literate participants due to the rarity of errors produced.

EXPERIMENT 2–WORD PAIR ASSOCIATION

Materials and Method

This second task is a variant of the word pair association test of the Wechsler Memory Scale (Wechsler, 1945): Two sequences (Version A and Version B) of 10 pairs of words were prepared. In each sequence, five words were semantically related (for example, “garfo–colher” which means “fork–spoon,” or “rosa–cravo,” which means “rose–carnation”), and the other five were phonologically related and semantically unrelated (e.g., “mala–pala” [mala]–[pala], which means “bag–eyeshade,” or “lua–rua” [lua]–[rua], which mean “moon–street”). Participants were tested both with Version A and Version B. The method of testing was performed according to the instructions of the Wechsler Memory Scale. The final score was the sum of the results on both versions. Version A and Version B were not presented sequentially, but separated, by other tests within the same session.

<table>
<thead>
<tr>
<th>Source of variance</th>
<th>$F$</th>
<th>$p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group</td>
<td>19.74</td>
<td>.0001</td>
</tr>
<tr>
<td>Repetition</td>
<td>20.49</td>
<td>.0001</td>
</tr>
<tr>
<td>Group × Repetition</td>
<td>16.22</td>
<td>.0002</td>
</tr>
</tbody>
</table>

Research Participants

Participants included in this experiment were the same as for Experiment 1, and both experiments were performed in the same session.

Results

A two-way ANOVA was conducted (Table 2 and Figure 2) for Group (literate vs. illiterate) × Type of Word Association (semantic vs. phonological). Results showed a literacy group effect $[F(1,56) = 84.58, p < .0001]$, a word association effect $[F(1,56) = 9.83, p = .003]$, and an interaction of Group × Type of Word Association $[F(1,56) = 6.10, p = .02]$. A post hoc test (Scheffé test) revealed that the only non-significant interaction was that of Semantic Association × Phonological Association in the literate group ($p = .98$).

Results showed that illiterate participants performed worse than literate ones in both the phonologically related pairs and the semantically related ones, while literate participants showed no difference in performance between the two types of word association. Results also indicated that illiterate participants were much weaker on the evocation of phonologically related words than they were on the
evocation of semantically related words. The difference in performance between the two ways of storing related information was greater in illiterate participants. The error analysis showed a tendency to produce semantically related words instead of phonologically related ones, both in illiterate and literate participants. Among the illiterate participants, 15.4% of errors were semantically related, versus 15.9% in the literate group.

EXPERIMENT 3 – VERBAL FLUENCY

Materials and Method

Two tasks of verbal fluency were designed. In the first, the participants were asked to produce names of animals (Subtask 1) and furniture (Subtask 2); in the second, the participants were asked to produce words beginning with the phoneme /p/ (Subtask 1) and /b/ (Subtask 2). The total score was composed by the total numbers of words produced during a 2-min period, for each criterion. Performance on Subtasks 1 and 2 in each task were summed for the treatment of the results.

Research Participants

The sample of participants taking part in this study was somewhat different from that included in the other two studies, although the criteria of inclusion were similar; that is, participants were from the same sociocultural environment, and the same criteria for literacy and illiteracy were also employed. Twenty-three illiterate female participants were included (mean age 62.2 years), and their performances were compared to those of 16 literate women (mean age = 58.4 years, mean years of schooling = 9).

Results

A two-way analysis of variance (Table 3 and Figure 3) for Group (literate vs. illiterate) X Criterion of Verbal Fluency (semantic vs. formal) was calculated. Again, a group effect \( F(1,74) = 129.25, p < .0001 \), verbal fluency effect \( F(1,74) = 77.96, p < .0001 \), and an interaction between factors \( F(1,74) = 20.15, p < .0001 \) were found. Post hoc analysis (Scheffé test) showed that the nonsignificant interactions were between the formal criterion in literate participants and the semantic criterion in illiterate participants \( (p = .37) \), which was irrelevant to our discussion, and in the comparison between the formal and the semantic criterion in the literate group \( (p = .054) \).

Table 3 and Figure 3 shows the results obtained by the two groups of participants in each test. Illiterate participants had poorer scores in all tasks when compared to literate ones. The comparison of the performances in the tasks in which a semantic cue was used with those in which a formal cue was used showed that literate participants performed similarly in both tasks, whereas illiterate participants clearly performed more poorly in tasks in which a formal cue was given.

DISCUSSION

The results of all three experiments demonstrated differences in performance between literate and illiterate participants. The selection of a sample among those living in the same sociocultural environment reduces the risk of influence of many variables related to learning in general, and sophistication of cognitive life. It is natural that literate individuals have access to more information through reading. However, the knowledge we have about this particular community allows us to consider that, although some participants have a number of years of schooling, they do not have frequent reading habits, and their daily life is generally similar to that of the illiterates. Some of the participants in both groups were selected from within the same family. Therefore, the differences found may be considered as related to the knowledge of the grapheme–phoneme correspondence.

If the learned capacity of matching graphemes and phonemes is the major difference between the two groups, we have first to understand how this capacity influences the

<table>
<thead>
<tr>
<th>Source of variance</th>
<th>( F )</th>
<th>( p )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group</td>
<td>129.25</td>
<td>.0001</td>
</tr>
<tr>
<td>Verbal Fluency</td>
<td>77.96</td>
<td>.0001</td>
</tr>
<tr>
<td>Group × Verbal Fluency</td>
<td>20.15</td>
<td>.0001</td>
</tr>
</tbody>
</table>
strategy used in repeating pseudo-words, since participants’ performances varied on this task. In order to repeat pseudo-words, it is necessary to use a phonological route that has been described in several studies with aphasic subjects (Coslett et al., 1987; 1991; see Butterworth & Warrington, 1995 for general reference). The phonological route starts, naturally, by a phonological analysis of the auditory input. A phonological input buffer permits the storage of phonological information (segmented and correctly sequenced) for a short period. This information will be used in future neural operations, such as auditory comprehension, repetition, or such other experimental tasks, as auditory stimuli comparisons or the detection of the specific phonemic characteristics of words. It seems that it is precisely this analysis that is problematic in illiterate individuals, as was shown in several studies performed on illiterate persons, who demonstrated difficulties in certain tasks that required phonological awareness (Morais et al., 1979; Adrián, 1993; Morais, 1993).

If we want to fully understand the relationship between phonological processing and reading capacity, it is necessary to go back to the preschool years. Studies performed on this age group showed that children are aware of some phonology before learning to read (Wimmer et al., 1991). We can thus consider this ability as a prerequisite for learning the grapheme–phoneme association. In order to make an association between an auditory component of the language and its written counterpart, it is necessary to be able to identify it as an isolated unit. This compartmentalized auditory unit will become meaningful to the system, as having the quality of visual representability. This will happen at several levels, from the single phoneme, to the larger cluster of phonemes. Recent results on developmental dyslexia may also contribute to better elucidating our point. Master-son et al. (1995) recently reported deficits in phonemic processing in some cases of developmental phonological dyslexia. Their participants, S.B. and R.E., who were adult phonological dyslexics, made errors in pseudo-word repetition, suggesting a close connection between reading and auditory phonemic discrimination. On the other hand, Gathercole (1995) suggests that learning to read does not influence the phonological working memory in children, although it depends on it. Our results with pseudo-word repetition suggest that the absence of a visual reinforcement of a phonological analysis for reading purposes decreases the efficiency of the explicit phonological processing of oral language in adult life. Nevertheless, there is some evidence that phonological awareness can be learned late in life by illiterate persons (Morais et al., 1988), which should enable illiterate individuals to become literate during adult life. It would be valuable to know whether their ability to repeat pseudo-words improves after such late learning to read and write.

There is therefore one level of phonological analysis shaped by orthography in the input of the system for which illiterates are not skilled. The same occurs when the process of reading and writing is based on a grapheme–semantic, and not on a grapheme–phoneme, association (see Prakash et al., 1993).

Thus, processing of the written representation of language operates as a parallel system for phonological auditory processing. One major difference between these two processing mechanisms may rely on the fact that one is paced by semantics and is implicit, whereas the other one is bound to visual experience, and is explicit. During the early stages of language development, the stabilization of rules stems from the interaction of three distinct operations: (1) phonological analysis, (2) semantic reinforcement, and (3) phonological production. The interaction of these three operations is mainly based on semantic reinforcement, as can be deduced from the work of Patterson et al. (1994), in cases of semantic memory loss. In those rare cases of semantic dementia, the authors showed that the errors produced in recalling unknown words reveals a deterioration of the phonological composition of the words. They conclude that meaning plays a crucial role in binding the elements of phonological word forms. This is what we consider to be the implicit mechanism of phonological processing. Learning to read and write brings an explicit dimension to the process.

To repeat pseudo-words it is necessary to acquire a certain sophistication of explicit phonological analysis. To perform this task, illiterate individuals use solely the implicit mechanism that recognizes and submits the information to the semantic system, whereas literate individuals use both this mechanism and an explicit phonological analysis. The semantic errors of illiterate participants while repeating pseudo-words reflect the influence of the semantic system.

The working hypothesis that can be developed from our results is that semantics is the major attractor, or reference system, for language processing. Subsidiary or secondary
systems are developed through formal learning. The written form of the words constitutes an important secondary reference system. Literate adults process words using all systems simultaneously and in parallel, whereas illiterates base most of their operations on the semantic one.

The study of aphasic illiterate adults may greatly contribute to a further understanding of the analysis of the mechanisms of language production; namely, to the language output system, and the question of the brain mechanisms underlying language processing. The study of Lecours et al. (1988) reflected the findings we obtained in this study. One of the main differences the authors found was in tasks of oral repetition. Our preliminary results with illiterate aphasics showed that the repetition of real words correlated with oral comprehension, which was not true for literate aphasics. This implies that access to semantics is particularly important for accurate oral repetition in the case of illiterate persons (Castro-Caldas et al., 1994).

Results in Studies 2 and 3 showed that illiterates perform better when semantic strategies are compared to strategies related to the phonological aspects of words. The performance of illiterates, even at the semantic level, is worse than that of literate individuals. It may be that the parallel running strategies are interactive, and not simply alternatives. Such an interaction might facilitate the performance of individuals who have more available strategies.

Thus, all our results support the hypothesis that the missing of a single skill (grapheme–phoneme association) interferes significantly in the higher development of the language system.

ACKNOWLEDGMENTS

Supported in part by Junta Nacional de Investigación Científica e Tecnológica (JNICT) (A.R.–BD/M2361) and a grant from project STRIDE (352/92).

REFERENCES


