The Effects of Literacy on the Recognition of Dichotic Words

José Morais
Université libre de Bruxelles, Belgium

Sao Luis Castro
Universidade do Porto, Portugal

Leonor Sciliar-Cabral
Universidade Federal de Santa Catarina, Brazil

Régine Kolinsky and Alain Content
Université libre de Bruxelles, Belgium

The hypothesis that awareness of phonemic segments influences the way in which speech is perceived was examined. Illiterate adults, who generally lack awareness of segments, were compared with literates, who are aware of the segmental structure of speech, on the recognition of words presented dichotically. A group of people who learned to read and write but who do it only occasionally was also tested. The results indicated much better performance in literates than in illiterates or semiliterates. In addition, literates made proportionally more single-segment errors, especially those limited to the first consonant, and fewer global errors, i.e. on all the segments of a syllable, than illiterates. On the other hand, phonetic feature blendings were as frequent in illiterates as in literates. It is suggested that awareness of segments allows attention to be focused on the phonemic constituents of speech and thus contributes to better recognition in difficult listening conditions. However, awareness of segments does not influence the preattentive extraction of phonetic information.

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

This work has been partially subsidized by the Belgian “Fonds de la Recherche fondamentale collective” under conventions nos. 2.4505.80 and 2.4562.86. The two last authors hold grants from the “Fonds National de la Recherche Scientifique”. We wish to thank Bruno Repp, Quentin Summerfield and an anonymous reviewer for very helpful comments on a previous version.
Introduction

Studies on speech perception in the adult have generally employed the university student as subject. The assumption that speech perception is independent of educational level should, however, not be accepted without further investigation. Illiterate adults have been shown to be unable to perform explicit manipulations of subsyllabic units, like deleting the initial consonant from an utterance provided by the experimenter (Morais, Cary, Alegria and Bertelson, 1979; Morais, Bertelson, Cary and Alegria, 1986). Though this inability does not preclude the possibility of unconsciously using these units in speech perception, it is also possible that the acquisition of metaphonological abilities permits the development of listening procedures focusing on the segmental structure of the utterances. This suggestion, which was already made in a somewhat veiled form by Mehler, Morton and Jusczyk (1984), may appear implausible to some theoreticians. But it should be noticed that conscious processes play an important role in the acquisition of many automatic abilities.

In the present study we compare illiterates, semiliterates—i.e. people who are able to read and write but who do it only sporadically—and literates on the recognition of words presented dichotically. There were several reasons for using the dichotic technique. First, dichotic competition makes recognition difficult. In addition, when isolated words are presented, no contextual cues are available, so that attention to the phonemic structure of the stimuli may be useful. Second, many errors may consist in combining parts or properties from each of the two dichotic stimuli and would thus provide an opportunity for examining the kind of units used in perception. We will examine in particular several phenomena of dichotic interaction, namely phonetic feature fusion or blending and phonological fusion (cf. Cutting, 1976), which seem to imply the extraction of phonetic features and phonemes, respectively. They will be discussed below. One additional reason for using the dichotic technique was our concern with literacy effects on ear differences. This issue is considered in a separate paper (Castro and Morais, in press).

When stop consonants are employed under dichotic stimulation, subjects (university students) tend to make blending errors that may be described as the wrong combination of one phonetic feature of one stimulus with another feature of the other stimulus; for instance, given the pair [ba]–[ta] they respond [da]–[pa] (combining the existing features inappropriately) more often than [ga]–[ka] (introducing a new feature). This effect has been thought to constitute evidence for the perceptual analysis of speech into phonetic features (cf. Studdert-Kennedy and
Shankweiler, 1970; Cutting, 1976; Tartter and Blumstein, 1981) and has therefore been called phonetic feature blending. However, Repp (1977) has shown that this interpretation is too simple. He found that blending errors are much more frequent when the pair includes [ba] than [pa] and are affected by within-category variations in both voice onset time and formant transitions. On the other hand, blending errors are equally frequent regardless of which stimulus begins first in a relative onset manipulation, do not decrease with an increase of difference in fundamental frequency (see Cutting, 1976) and are not affected by differences in both pitch and vocal tract properties (Tartter and Blumstein, 1981). It thus seems that an interpretation in terms of dichotic interaction of the sole auditory properties of the stimuli is also untenable. The blending effect should arise early in the processing of speech, but not before a stage where acoustic information is evaluated by reference to speech categories.

Phonological fusion is typically the combination of an utterance beginning with a consonant and of another utterance beginning with a different consonant so as to form a third utterance beginning with a consonantic cluster: for example, *banket/lanket* yields *blanket* (Day, 1968). The effect may also occur, though probably to a lesser extent, in other positions of the words (Cutting and Day, 1975). It is not much affected by small differences in relative onset time, intensity and fundamental frequency (Cutting, 1975; 1976). It depends, rather, on linguistic factors: number of syllables, whether items are words or non-words and whether or not phonotactic regularities are preserved (*lbanket* is never heard, but *tass/tack* can yield both *tacks* and *task* responses; cf. Day, 1970). The theoretical implications of phonological fusion have been somewhat overlooked in the literature that deals with the units of speech perception, in particular regarding the phoneme/syllable debate. Two CV (consonant–vowel) utterances could not give rise to perception of a CCV syllable if syllables were the smallest units extracted during perceptual processing. On the other hand, the occurrence of phonological fusions does not permit any inference about what unit, from syllable and phoneme, is processed first.

The dichotic test used here consisted of pairs of CVCV words. All words had stress on the first syllable and began with a stop consonant, and each pair differed in at least the three initial segments. It is thus possible to compare segmental intrusions of the first consonant or of the first vowel alone with intrusions of the first syllable; and, by contrasting manner of articulation of the second consonant, opportunities were provided for phonological fusions. On some of the trials—for instance on the trial “capa–bota”—the initial stop consonant differed both on voicing and place, thus permitting an estimation of feature blendings.
Our expectations were the following: (1) Assuming that literacy makes the subject more attentive to the segmental structure of the words he tries to recognize, we predicted proportionally more segmental errors (i.e. limited to one phoneme) and fewer global errors (i.e. on all the segments of a syllable) in literates than in illiterates. (2) Assuming that phonetic processing is a universal, built-in component of speech perception mechanisms, we predicted the same rate of phonetic feature blendings in all groups. Regarding phonological fusions, we had no precise prediction, as we did not know to what extent the illusion depends on attention to the segments. Finally, we expected semiliterates to display an error pattern between those of illiterates and literates. Their position between these groups would depend on the extent to which attention to the speech segments in difficult listening conditions is elicited by having learned to read and write or requires a great deal of literacy experience.

Method

Subjects

All the subjects were female, since it is difficult to find male illiterates in Portugal. There were 18 illiterates, 14 semiliterates and 14 literates. Illiterates were aged between 43 and 61 years (mean = 51.8, sd = 6.24). They were living in an agricultural area of northern Portugal where they were working in farming or as servants. For socioeconomic reasons they had not attended school as children and never received reading instruction. The semiliterates were aged between 42 and 60 years (mean = 50.9, sd = 6.07). Eight of them were living in a rural milieu and working in farming or as servants; the remaining six lived in a small town and worked as housekeepers. They had had a maximum of four years’ schooling in childhood, which gave them basic reading and writing skills; these, however, were neither strictly necessary nor regularly used in their jobs. Literates were aged between 38 and 63 years (mean = 50.9, sd = 7.81). Two subjects had 9 years’ schooling, twelve had 12 years’ and two had reached university level. All were able to read and write fluently and made systematic use of these skills in their jobs. Two subjects, one literate and one semiliterate, who complained of reduced loudness in one ear, were replaced.

Stimuli and Procedure

The test consisted of 48 pairs of disyllabic CVCV Portuguese words presented dichotically at a rate of one pair every 5 sec. Each pair was preceded by a beep, which acted as a warning signal. The initial consonant was always a plosive. The words were uttered by a male Portuguese speaker. Each word of a pair had first been recorded on a different tape and marked at the onset of the explosion. Then, the two words were played back simultaneously on two synchronized tape recorders (Revox A77) and fed to the two channels of a third one. The final tape was later monitored on a Siemens Oscillomink graphic recorder to avoid synchronization errors that would exceed 10 msec. The words presented on each
trial differed in at least the three initial segments (for instance, "pato"-"cura", "capa"-"bota"). Initial consonants differed either in place alone, voicing alone or both place and voicing. The second consonant was either a plosive, or a liquid, or a fricative, or a nasal, and different types of contrast were used. The stimuli were presented to the subjects at approximately 70 dB SPL via stereo headphones and the corresponding high-fidelity playback equipment (CP-105 Hatadi Pearce Simpson). The subjects were tested individually in a quiet room. They were told that they would hear words spoken by a man through the headphones; each time there would be one word at one ear and another word at the other ear, as if two different men were speaking, one on the left side and the other on the right side. Their task consisted in paying attention to one ear only, as indicated to them by the experimenter, and in repeating what they had heard on that ear. If they felt unsure about what they heard, they should, nevertheless, "speak it out", even if it was nonsense. In the beginning, half of the subjects in each group were instructed to pay attention to the left ear, the other half to the right. Every six trials the subjects were instructed to switch attention between the ears, and at the same time the headphones were reversed.

All subjects were given, after the test described above, a second set of 48 dichotic trials using the same words but paired in such a way as to differ in the first consonant only (e.g. "pato-gato"). These trials were of no interest for the main purpose of the present investigation. They were included because different ear advantages in illiterate adults have been reported for phonetically similar and dissimilar words (Damásio, Damásio, Castro-Caldas and Hamsher, 1979). Thus, the results relative to the second test are reported only in Castro and Morais (in press). Lastly, subjects' audition was assessed by a monaural test consisting of 18 (9 per ear) digits presented at three decreasing levels of intensity (about 75, 65 and 55 dB SPL). No unilateral or bilateral hearing impairment was noticed.

**Results**

Overall performance was better in the literates (mean = 64.4%, ranging from 35.4% to 87.5%) than in the semiliterates (mean = 48.7%, ranging from 22.9% to 77.0%) and the illiterates (mean = 44.4%, ranging from 25.0% to 68.8%). Analysis of variance on the raw data showed a significant effect of group, $F(2,43) = 7.49$, $p < 0.005$, and ear, $F(1,43) = 13.15$, $p < 0.005$. Scheffé's contrasts indicated that the literates were significantly superior to either the illiterates or the semiliterates, but that these two groups were not significantly different from each other. The ear effect consisted of a right-ear advantage, and there was no Ear x Group interaction ($F < 1$).

**Local vs. Global Errors**

Table I presents the proportions of several error categories. Relative rather than absolute error rates are examined because we are interested in the relative importance of different sources of error. The assumption
Table I

Proportions of Error Categories (Mean Percentages)

<table>
<thead>
<tr>
<th>Errors</th>
<th>Illiterates</th>
<th>Semiliterates</th>
<th>Literates</th>
</tr>
</thead>
<tbody>
<tr>
<td>WWI</td>
<td>7.1</td>
<td>9.3</td>
<td>3.8</td>
</tr>
<tr>
<td>C1</td>
<td>23.8</td>
<td>29.1</td>
<td>38.7</td>
</tr>
<tr>
<td>V1</td>
<td>5.7</td>
<td>6.3</td>
<td>5.4</td>
</tr>
<tr>
<td>C2</td>
<td>7.6</td>
<td>7.5</td>
<td>8.8</td>
</tr>
<tr>
<td>ClV1</td>
<td>10.9</td>
<td>10.4</td>
<td>7.5</td>
</tr>
<tr>
<td>C1C2</td>
<td>6.4</td>
<td>8.7</td>
<td>6.6</td>
</tr>
<tr>
<td>ClV1C2(V2)</td>
<td>12.0</td>
<td>8.1</td>
<td>6.7</td>
</tr>
<tr>
<td>Structural</td>
<td>6.9</td>
<td>10.1</td>
<td>10.0</td>
</tr>
<tr>
<td>Others</td>
<td>19.6</td>
<td>10.5</td>
<td>12.5</td>
</tr>
</tbody>
</table>

Note:
Errors consisting in
WWI: whole-word intrusions,
C1: changing the initial consonant only,
V1: changing the initial vowel only,
C2: changing the second consonant only,
ClV1: changing the initial syllable only,
C1C2: changing both consonants only,
ClV1C2(V2): changing the three initial segments at least but excluding whole-word intrusions.
Structural: Errors that did not respect the CVCV structure of the words;
Other: Errors on the second vowel only, and all the categories of errors on two or three segments that include the second vowel.
All the percentages are calculated relative to the total number of errors in each group.

underlying this analysis is that error categories may reflect different recognition mechanisms or strategies. More precisely, the analysis of relative error rates brings us information about the relative capacity of different mechanisms to determine the subject’s response. Examination of Table I suggests that literates make, proportionally, more C1 errors than illiterates, while illiterates make more ClV1 and ClV1C2(V2) errors. An analysis of variance on the arcsine transformed proportions of errors involving only the initial consonant showed a significant effect of group, \( F(2, 43) = 3.87, p < 0.05 \). Scheffé’s tests indicated that only the literates vs. illiterates comparison was significant. There was no ear effect (\( F < 1 \)), despite a tendency for the proportions averaged over the three groups to be higher in the right (0.321) than the left (0.287) ear. This tendency was roughly the same in the three groups. Accordingly, there was no Ear × Group interaction (\( F < 1 \)).

Among C1 errors, one type was particularly frequent: giving the initial consonant from the word presented to the irrelevant ear. These
initial consonant intrusions accounted for 9.2%, 13.3% and 15.5% of all the errors in illiterates, semiliterates and literates, respectively. However, their proportion among C1 errors did not differ between the groups: 0.39, 0.46 and 0.40, respectively.

The correlation between the arcsine transformed proportions of C1 errors and the number of correct responses was positive in each group: non-significant in the literates (0.28), it fell short of significance at $p<0.05$ in the illiterates (0.44) and reached $p<0.01$ in the semiliterates (0.75).

Since all correlations were positive, suggesting that the proportion of C1 errors increases with overall performance, it is important to ascertain whether an effect of literacy on the proportion of C1 errors still occurs when subgroups are equated for overall performance. The seven poorest literate subjects and the ten best illiterates produced the same mean percentage of correct responses (53.3%). A one-tailed $t$-test performed on the arcsine transformed proportions of C1 errors showed that these errors remain proportionally more frequent among the literates, $t(15)=1.75$, $p=0.05$. On the other hand, literacy influences overall performance even when subgroups of illiterates and literates are roughly equated for the proportion of C1 errors (0.286 and 0.267, respectively). The thirteen illiterates with the highest proportions of these errors performed significantly lower than the seven literates with the lowest proportions, $t(18)=2.35$, $p<0.025$.

An analysis of variance was performed on the arcsine transformed proportions of global errors, i.e. the sum of C1V1 and C1V1C2(V2) errors (excluding whole-word intrusions, which may reflect a failure of spatial selective attention). This analysis showed an effect of group, $F(2,43)=3.71$, $p<0.05$). Only the Scheffé test contrasting literates and illiterates was significant. There was, in addition, an effect of ear, $F(1,43)=4.41$, $p<0.05$, reflecting the fact that the proportion of global errors was higher on the left (0.197) than the right (0.140) ear. All the groups exhibited a tendency in this direction. Accordingly, there was no Ear $\times$ Group interaction ($F<1$).

The correlation between the arcsine transformed proportions of global errors and the percentage of correct responses was negative in each group: $-0.45$ (non-significant), $-0.68$ and $-0.67$ (both significant at $p<0.01$), in the literate, semiliterate and illiterate groups, respectively.

One rather uninteresting possibility is that the differences between groups in the proportion of local and global errors are related to illiterates’ reluctance to produce non-word responses. If C1 misperceptions lead frequently to non-words, illiterates might then have produced lower proportions of C1 errors and higher proportions of global ones.
However, contrary to this hypothesis, the percentage of non-word responses among the errors decreased, rather than increased, with literacy: it was, on average, 26.1%, 22.6% and 16.4% in illiterates, semiliterates and literates, respectively. An analysis of variance on the arcsine transformed proportions of non-word responses showed a significant effect of group, $F(2, 43) = 3.45, p < 0.05$. Scheffé's tests indicated that only the literates and the illiterates differed significantly. This group effect may result from a trivial response bias. Since all the items were words, the literates, having identified many more items than the illiterates, might have tended to produce words when uncertain.

**Feature Analysis**

Errors limited to the initial consonant when this differed on voicing only showed an effect that is linked presumably to the characteristics of Portuguese. There were four times more errors on voicing when the attended stimulus was unvoiced than when it was voiced. In Portuguese, voicing is signalled by the presence of periodicity, generally for more than 0.1 sec, before the explosion. Most errors may thus have resulted from a wrong interpretation of the spatial origin of the periodicity in the situation of dichotic competition. The effect may be analogous to one discovered by Repp (1975): in a situation where English CV targets are presented to one ear and isolated vowel masks to the other ear, there is a tendency to give voiced responses when the onset of the vowel precedes the onset of the syllable. This temporal relationship across ears competes with the temporal cues of voicing in the target.

Errors limited to the initial consonant when this differed in both voicing and place between the ears were examined for the occurrence of phonetic feature blendings. Given, for instance, the pair [b]–[t], the probability of a blending ([p] or [d]) equals, by chance, that of an anomalous stop response ([g] or [k]). The results show that the former responses are more frequent than the latter. Illiterates made 29 blendings but only 7 anomalous stop responses ($p < 0.001$, by the sign test); semiliterates, 18 and 7, respectively ($p < 0.05$), and literates 20 and 5 ($p < 0.005$). The ratio between the number of blendings and the number of blendings plus the number of anomalous stop responses is not higher in literate (0.80) or semiliterate subjects (0.72) than in illiterates (0.81).

We felt, in this analysis, the necessity of separating voiced and unvoiced responses. The reason is that voiced blendings may eventually result from a tendency to accept periodicity whichever ear was stimulated by it. In the example above, the response [d] may result from combining prevoicing to the acoustic cues of [t] regardless of whether prevoicing is presented in the attended or the unattended ear. Fortunately, the number of unvoiced blendings is high enough to support, by comparison
with the number of unvoiced anomalous stop responses, the notion of phonetic feature blending. Illiterates made 14 unvoiced blendings but only 1 unvoiced anomalous stop response ($p < 0.001$, by the sign test); semiliterates, 10 and 2, respectively ($p < 0.05$), and literates 15 and 2 ($p < 0.001$).

The correlation between the number of feature blendings minus the number of anomalous stop responses on the one hand, and the number of correct responses on the other hand, was low and not significant in any group: 0.18, 0.14 and $-0.01$, in the literate, semiliterate and illiterate groups, respectively.

*Phonological Fusions*

Finally, errors were also examined for the occurrence of phonological fusions. The response was considered as a phonological fusion if it included an additional segment to the four in the target word and if the five segments were present in the dichotic stimulus regardless of the ear stimulated. Transformations of segments into very similar ones were, however, also accepted. All the phonological fusions observed were consistent with Portuguese phonotactic rules. All but three were of the type CVCCV; in two cases the additional C was transformed into the semivowel /w/, and in one case the structure was CCVCV. About 20% were nonsense words. The total number of phonological fusions was 11, 18 and 10 in the illiterate, semiliterate and literate groups, respectively. They are detailed below.

Six trials contrasted the liquid /l/ and a stop consonant. Among them, /palA/-/boda/ yielded the response /baldA/ in eleven subjects, and /kalu/-/toda/ yielded either /KaldA/ or /taldA/ in seven; in the last trial two subjects responded /kawdA/.

Four trials contrasted /r/ and a stop consonant. The pair /gadu/-/koru/) yielded one response /gordu/, and the pair /tarA/-/koku/ one response /porku/.

Five trials contrasted /ɾ/ and a stop consonant. The pair /poku/-/baru/ yielded three responses /barku/, and the pair /buɾA/-/põtu/ yielded two responses /portu/, one /bortu/ and one /brutu/. These responses involve a phoneme change, from /ɾ/ to /r/, which is required by Portuguese phonotactics. The last pair yielded, in addition, four responses /boltu/ and two /voltu/, and the pair /gaɾA/-/kotu/ yielded one response /goltu/.

Eight trials contrasted a fricative and a stop consonant. Only one phonological fusion was observed: /bufu/-/pitu/ gave one response /bultu/.

Three trials contrasted a fricative and a liquid. The pair /tuɾA/-/kasA/ yielded one response /karta/, and the pair /gulA/-/posA/ one response /bolA/.
Discussion

In a task requiring the recognition of words presented dichotically, literates produced more correct responses than either semiliterates or illiterates. However, both semiliterates and illiterates understand speech under normal circumstances. Their behaviour in the present situation, particularly that of the illiterates, suggests that they may not process speech in exactly the same way as literates.

The present experiment brings some evidence both on preattentive processes of phonetic decoding and on attentional processes. Preattentive processes of speech decoding do not seem to contribute in a significant way to differences in performance. First, we found no correlation between the estimation of phonetic feature blendings, indicative of extraction of phonetic information, and overall performance. Second, the proportion of feature blendings was not lower in illiterates than in literates. Thus, illiterates would not employ the automatic processes of phonetic decoding to a lesser extent than literates. The highly analytical operations of phonetic decoding involved in speech recognition precede, in all likelihood, the onset of literacy.

With respect to attentional processes, one could speculate that illiterates and semiliterates are poorer performers because they are deficient in the spatial focusing of attention. However, this hypothesis may be discarded. Though they make more whole-word intrusion errors than the literates, their surplus of such intrusions is small in comparison with the difference in correct responses. A more likely interpretation will probably be found by reflecting on the possible sources of variation in the proportion of single-segment and global errors. In fact, the proportion of errors limited to the initial consonant was positively correlated with overall performance, while the proportion of global errors was negatively correlated to the latter. In addition, literates displayed both a higher proportion of errors limited to the initial consonant and a lower proportion of global errors than illiterates.

Our suggestion is that a high proportion of single-segment errors may reflect a mechanism of attention to the segmental structure of speech. This mechanism would be particularly useful when the acoustic signal is not clear and contextual cues are lacking. Its development might be stimulated, if not triggered, by segmental awareness. As previous studies have indicated, segmental awareness is absent in most illiterates (Morais et al., 1979; 1986); it usually emerges during the first stages of the acquisition of literacy, though this need not be the only way to promote segmental awareness. Being unable to elaborate conscious representations of segments (illiterates) or simply lacking practice in the explicit analysis of speech into segments (semiliterates), illiterates and semiliterates would not focus on the segmental structure of the stimulus.
The mere occurrence of single-segment errors is not indicative of attention to segments. Attention to the whole stimulus may lead to some proportion of errors limited to one phoneme, and to more errors on consonants than on vowels if utterances with common vowels are perceptually less discriminable than utterances with common consonants. Moreover, the proportion of single-segment errors may increase as a result of a more accurate attentional effort. However, overall accuracy being the same, a greater proportion of single-segment errors should probably be taken as indicative of the utilization of a special strategy. In our experiment, literates displayed a higher proportion of errors limited to the initial consonant than illiterates even when equated for overall performance. Thus, it seems that literacy has a specific effect on the mechanisms of speech comprehension, which, given what we know about the relationship between literacy and segmental awareness, would consist of inducing greater resort to segmental focusing. Of course, the demonstration that the effect of literacy obtained here is really mediated by segmental awareness needs experimental rather than correlational evidence.

On the other hand, literates performed better than illiterates even when equated in terms of the proportion of errors limited to the initial consonant. This result suggests that attention to the segmental structure of speech, though contributing to improved performance, is not the only factor in performance differences. Whatever the object of attention, this may be more or less intensive, and it may also draw upon more or less cognitive resources, and therefore be more or less accurate. The greater overall accuracy obtained by the literates probably results both from their ability to attend to segments and from other factors related to the exercise of attention.

The model of speech perception we are inclined to believe in is one that includes automatic, preattentive processes of phonetic decoding, which would be insensitive to the effects of phonological awareness and so of literacy, and attentional processes, which contribute to the final perceptual interpretation of the outputs of the former. Phonetic decoding would be accomplished in a biologically distinct system, a "module" (cf. Fodor, 1983; and Liberman and Mattingly, 1985), existing in every human being. Though exhibiting variability as any other biological system, its role as a source of individual differences in performance would be negligible. On the contrary, attentional processes—in particular attention directed towards one or another kind of unit within the utterance—would be responsible for large inter-individual differences in performance. Awareness of segments, which is acquired during learning to read and write in an alphabetic system (Morais et al., 1979, 1986; Read, Zhang, Nie and Ding, 1986) would allow attention to be directed towards the phonemic constituents of speech.
Other studies begin to stress the importance of attentional strategies in auditory word perception, regarding more specifically the possibility of attending to different levels of information. Working on the phonemic restoration illusion, Nusbaum, Walley, Carrell and Ressler (1982) found that literate subjects could learn to distinguish between the addition of a bit of noise to one phoneme and replacement of the latter by the former by attending to the phonemic structure of the word. More recently, Samuel and Ressler (1986) obtained a similar effect when the identity and location of the critical phoneme was cued, provided that the identity of the test word was previously known. The present study, using a situation that is less constrained by lexical factors and that renders identification of the word difficult, leads to a stronger suggestion: attention to the segments may contribute to the very identification of the word.

Do the present results have any implications beyond the highly artificial dichotic listening situation? We believe that artificial situations may help to disclose processes actually used in natural circumstances. The idea that literates had developed a mechanism of segmental attention to perform better on our dichotic tasks is, of course, unreasonable. But people who, in the context of literacy acquisition and practice, have become able to pay attention to the segments in speech may profitably use this ability in some situations, for instance when listening to words in noise or to an ill-known foreign language while lacking enough pragmatic information. We predict, and plan to test, that literates should be better at and should make proportionally more single-segment errors than illiterates in perceiving lists of unrelated words in noise.

Semiliterates as a group behaved in the present task more like the illiterates than the literates. This may reflect the action of either or both of two factors. First, the full development of an attentional mechanism directed towards the segmental units of speech may need more than the segmental abilities required by rudimentary reading and writing. We plan to address this question by running similar studies on children from various school grades. Second, the preservation of those mechanisms may need a relatively constant practice of literacy. A more detailed assessment of the actual literacy level of semiliterate populations, as well as of their segmental analysis abilities, seems to be worthwhile for further investigations in the present matter.

A right-ear advantage was observed here, which is discussed in a more comprehensive study dealing with the laterality issue (Castro and Morais, in press). It is important to note that neither the ear effect on overall accuracy nor the ear differences observed on the proportion of either kind of errors, local and global, interacted with group. Thus, the
group differences discussed above are not contaminated by laterality factors. A further, potentially interesting, observation is that the proportion of global errors was higher in the left than in the right ear. At first sight, this result is consistent with the view that right-hemisphere processing is less analytical than left-hemisphere processing (cf. discussions in Bradshaw and Nettleton, 1981; and Morais, 1982). Global errors might, at least in part, result from a right-hemisphere strategy consisting in trying to match some global acoustic properties of the input with corresponding properties in the mental representation of lexical items. However, another possible interpretation is that global errors are more frequent in the left ear, not because this ear reflects mainly right-hemisphere processing, but because it is the ear to which attention is poorer in the majority of the subjects and on which, consequently, performance is lower. In the absence of compelling evidence, the interpretation in terms of hemispheric differences should probably be considered as less parsimonious.

Finally, the occurrence of phonological fusions deserves some commentary. They were observed in all groups. Thus, they would not depend on literacy, nor on phonological awareness. However, there was some hint of a greater tendency in semiliterates to produce phonological fusions, and since the number of such responses was relatively small in the present study the question remains open to further examination. Tangentially, some aspects of these results are worth noting. As already noticed in previous investigations, stop-liquid pairs fuse, yielding a liquid-stop sequence, more readily than fricative-liquid pairs, stop + /l/ pairs produce more typical fusions (i.e. without phoneme transformation) than stop + /r/ pairs, and /l/ is more frequently substituted for /r/ than the reverse (Cutting, 1975). More importantly (incidentally, previous studies frequently used monosyllables; when they did use polysyllables they usually contrasted initial or final positions), it now appears quite clear that phonological fusions do not respect syllable boundaries: the consonant of the final syllable of one stimulus may be added to the initial syllable of the other stimulus so as to produce an initial CVC syllable. One would not expect such errors if syllables were the basic units of speech processing.

In sum, the present work suggests that whereas preattentional processes of speech perception are not influenced by literacy, literacy does contribute to the development of mechanisms of attention to the segmental structure of speech. These mechanisms, which probably play only a negligible role when the speech signal is clear and both syntactic and pragmatic factors intervene, may be instrumental in more difficult listening conditions.
REFERENCES


*Revised manuscript received 9 September 1986*