Recognition memory for items from unilingual and bilingual lists

P. D. McCORMACK and S. P. COLLETTA
Carleton University, Ottawa K1S 5B6, Canada

It was predicted that recognition memory for items from bilingual lists would be superior to that from unilingual lists. Each of 72 bilingual subjects, fluent in English and French, was tested on two unilingual and two bilingual lists. Each list involved a study phase of 20 targets and a test phase of these and 20 distractors. A signal-detection analysis revealed a significantly larger d' value for the bilingual than for the unilingual treatment.

Free recall from a bilingual list appears to be equivalent to that from a unilingual list of the same length, in spite of the fact that the former setting imposes the additional demand on the subject of having to recall in the appropriate language (Kolers, 1965; Lambert, Ignatow & Krauthamer, 1968; Liepmann & Saegert, 1974; Nott & Lambert, 1968; Saegert, Obermeyer, and Kazarian, 1973). The only exception to this finding has been reported by Tulving and Colotla (1970) where recall was observed to be superior from a unilingual list.

It might, however, be predicted that recognition memory should be better for the bilingual than for the unilingual setting. First of all, and of foremost importance, is that the requirement to recall in the correct language is no longer in effect. Secondly, since language has been demonstrated to be an effective encoding attribute (Dillon, McCormack, Petrusic, Cook, & Lafleur, 1973; Goggin & Wickens, 1971), items from a bilingual list should have the additional advantage of a linguistic tag. This would presumably provide an additional basis for discrimination among alternatives on the test. Thirdly, should recognition memory involve search operations (see McCormack, 1972), and should the search be selective with respect to language, then the search set size for the bilingual list is half that of the unilingual list. This is also, of course, the case for the recall setting, but this advantage is likely counteracted by the requirement to recall in the correct language. Finally, the bilingual setting should involve less storage and test-phase interference of the McCormack-Swenson (1972) variety (see also: McCormack, 1974) than that encountered in the unilingual situation, again because the set size is halved.

Support for this research was received under National Research Council of Canada Grant APA-78. The authors are indebted to Joan McCormack for computer programming as well as to H. P. Edwards and R. Vaillancourt, both of the University of Ottawa, for kindly providing access to subjects. They also wish to thank the 10 bilingual adult translators, all of whom were former neighbours of the senior author. Requests for reprints should be sent to P. D. McCormack, Department of Psychology, Carleton University, Ottawa, Ontario K1S 5B6, Canada.

METHOD

Subjects

The 72 subjects were 43 females and 29 males, all of whom were presumably equally fluent in English and French. They were paid \$2.00 each for participating. Forty-eight of these were enrolled in summer school at the University of Ottawa, with the remaining 24 being students in fall-term classes at Carleton University. The data of four subjects were subsequently discarded, since recognition performance was no better than chance.

Design and Materials

All subjects were tested once on each of four lists. The 24 possible orders of these four lists were each used three times. One list was unilingual English (U_E), 20 English target nouns being presented for study followed by a test phase involving these targets and 20 English distractor nouns, in random order. Another list was unilingual French (U_F), the 20 target and 20 distractor nouns being presented in the French language. The remaining two lists were bilingual (B_1 and B_2), the study phase involving 10 English and 10 French targets in random order. These targets were presented with 10 English and 10 French distractor nouns, again in random order, during the test phase.

The original word pool consisted of 264 English nouns taken from the norms of Paivio, Yuille, and Madigan (1968). All had A or AA frequency ratings, and their mean meaningfulness and concreteness values were in the middle range of these dimensions. These items were presented to 10 bilingual adults for translation into French. Of the 264 words, 104 could not be translated reliably. This left 160 nouns which were translated identically by the 10 judges. The 160 English and 160 French items were then used to construct the four lists, with a different randomization for each subject. Once a word was selected for any given list, neither it nor its translation could be used again for that subject.

Apparatus and Procedure

All materials were presented by means of 35-mm slides mounted in trays in a projector located behind the display panel. The subject was located 46 cm in front of the panel. Items appeared at eye level on a 28 x 23 cm surface. A movable 10.5 x 5.0 cm response unit was positioned horizontally on a desk between the subject and the panel. On this unit were mounted two pushbuttons, 5 cm apart, one labeled "old/vieux" and the other labeled "new/nouveau." The experimenter was located between the subject and the projector at a distance of 1.1 m from him and typically out of his visual field. The experimenter's panel consisted of two timers, one latency recording device, and a control unit which allowed for a timer or a manually controlled operation of the projector.

For half of the subjects, the instructions were taped in English with a French summary, while the reverse was the case for the remaining subjects. Following the instruction period, the subject was given a single practice list with nine two-digit numbers serving as targets and appearing consecutively at 3-sec intervals. The subject was told that this task would be more difficult than the one involving words. Following the ninth target, a blank slide appeared, at which time the subject began counting backwards as fast as he could by threes from a three-digit number. Following 75 sec of backward counting, the nine targets and nine distractors were presented in a manual mode and in random order, with the subject pressing one of the two buttons each time an item appeared. This enabled the experimenter to record the subject's response as well as its latency. When the experimenter was assured of the subject's understanding of the task, the instructions were repeated in modified form, and the four experimental lists were presented, with a pause of approximately 45 sec between lists. For this experimental phase, the 20 targets were presented every 1.3 sec, otherwise the procedure was essentially identical to that of the practice task. The subject was told that a word, or its translation, would never be repeated in any other list and that, during the test phase of any given list, he should respond "old" if, and only if, the target was presented in its original language.

RESULTS AND DISCUSSION

An analysis of variance of the hit data revealed no reliable differences among the four lists ($F < 1.00$). A similar analysis of false alarms, on the other hand, resulted in a significant list effect, $F(3,201) = 2.71$, $p < .05$. This was in the predicted direction since the mean number of false alarms for the combined unilingual lists was 4.71, whereas that for the combined bilingual lists was 4.01, $t(67) = 1.97$, $p < .10$.

In the next analysis, the two bilingual lists were treated as one, as were the two unilingual lists, with the signal-detection parameters d' and β being calculated for the combined hit and false-alarm data. The normality and homoscedasticity assumptions appear to be reasonably well satisfied with this paradigm (see McCormack & Swenson, 1972). The critical comparisons, therefore, become d' for the unilingual treatment vs. that for the bilingual condition, as well as the unilingual vs. bilingual β values. These were calculated for each subject and difference scores were generated. Whenever a subject exhibited no false alarms or no misses, which occurred in only 15 out of 272 instances, the p value was set at .01 instead of at zero, otherwise $z = \infty$, $y = 0$. The distributions of both the d' and β difference scores appeared to be reasonably Gaussian. The mean β values for the bilingual and unilingual treatments were 2.22 and 1.56, respectively; $t(67) = 2.55$, $p < .02$. The mean d' values were 2.58 for the bilingual and 2.40 for the unilingual treatment, $t(67) = 2.60$, $p < .02$. The subjects, therefore, appeared to be more conservative, as well as more sensitive, in the bilingual than in the unilingual setting. The latter finding is precisely what was predicted. Similar conclusions were reached as a result of a nonparametric analysis of these data (Luce, 1963). The η means for the unilingual and

bilingual treatments were .14 and .11, respectively; $t(67) = 2.60$, $p < .01$. The b mean for the bilingual condition was 1.90, whereas that for the unilingual treatment was 1.49; $t(67) = 2.06$, $p < .05$.

The significant list effect, referred to earlier, was primarily due to Condition U_E , where the mean number of false alarms was 2.63. The false-alarm means for each of the other three conditions were as follows: $U_F = 2.07$, $B_1 = 1.96$, $B_2 = 2.06$. The critical difference between means for pairwise comparisons ($\alpha = .05$) was .53, and so Condition U_E differed reliably from each of the others, but they did not differ among themselves. A high false-alarm rate may be interpreted as being indicative of test-phase interference (McCormack, 1974; McCormack & Swenson, 1972). The primary underlying mechanism is likely related to degree of feature overlap among distractors, as contrasted with the more commonly accepted notion of features in common with targets (for example, see Tulving & Bower, 1974). Given that the subjects of the present investigation are, in their current milieu, more exposed to the English than to the French language, then their high false-alarm rate under Condition U_E is understandable in the light of the likelihood that English-language distractors share a greater degree of feature overlap than those from the French language. We, thus, have yet another instance of a higher false-alarm rate for common than for rare items (for example, McCormack & Swenson, 1972).

In another analysis, the two bilingual lists were examined with respect to both hits and false alarms for the English and French items separately. No reliable false-alarm differences were revealed between the two classes of items; $t(67) < 1.00$. There were, however, reliable hit differences, the mean number for the combined lists being 17.79 for French and 16.90 for English items; $t = 3.23(67)$, $p < .01$. A high hit rate has been interpreted as an indicator of interference which develops during the storage of targets (McCormack, 1974; McCormack & Swenson, 1972), the underlying mechanism likely being related to the degree of overlap among target features. The higher hit rate for the French than for the English targets may, therefore, be accounted for in terms of relative differences in exposure to the two languages and, consequently, relative differences in degree of target feature overlap.

The data of the present investigation are, therefore, interpretable in terms of two interference loci, one at storage and the other at the test phase. The English language items are those that appear to be most vulnerable, presumably because they are high-frequency words and therefore share certain critical features. For this interpretation to be logically consistent, a number of additional assumptions must be made. There appears to be, from an examination of the false-alarm data of Condition U_E , a relatively large amount of test-phase interference involving distractors. This is not the case for the English language items of the bilingual treatments.

Since the distractor set size is 20 in the former situation and only 10 in the latter, it is therefore assumed that degree of feature overlap is some direct function of distractor set size. The hit data, however, present a somewhat more complex picture. Storage interference among English language targets appears to be more pronounced in the bilingual than in the unilingual conditions, even though the set size in the former setting is half that of the latter. These data may be accounted for by assuming that the subject, when storing targets in two languages, devotes more processing time to his least encountered language, which is French in the case of the present study. The English language targets will, therefore, presumably share more features than the French items. This, in turn, should produce differential amounts of storage interference. It is, therefore, concluded that, while distractor set size appears to be the chief determiner of test-phase interference, depth of processing is likely the major contributor to interference which takes place during storage (for example, see Craik & Lockhart, 1972).

Further analyses of the data revealed no reliable differences among the four lists with respect to either hit reaction times (RT) ($F < 1.00$) or correct-rejection RTs, $F(3,201) = 1.65$, $p < .20$. This failure to observe faster RTs for the bilingual than for the unilingual lists militates against any notion that the search set is selective with respect to language. However, the usual finding of faster RTs for hits than for correct rejections was obtained, the two means being 1.18" and 1.44", respectively; $t(67) = 5.44$, $p < .001$. These results, of course, are not inconsistent with the notion of serial self-terminating search (see McCormack & Swenson, 1972).

REFERENCES

- Craik, F. I. M., & Lockhart, R. S. Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior*, 1972, 11, 671-684.
- Dillon, R. F., McCormack, P. D., Petrusic, W. M., Cook, G. M., & Lafleur, L. Release from proactive interference in compound and coordinate bilinguals. *Bulletin of the Psychonomic Society*, 1973, 2, 293-294.
- Goggin, J., & Wickens, D. D. Proactive interference and language change in short-term memory. *Journal of Verbal Learning and Verbal Behavior*, 1971, 10, 453-458.
- Kolers, P. A. Bilingualism and bicodality. *Language and Speech*, 1965, 8, 122-126.
- Lambert, W. E., Ignatow, M., & Krauthamer, M. Bilingual organization in free recall. *Journal of Verbal Learning and Verbal Behavior*, 1968, 7, 207-214.
- Liepmann, D., & Saegert, J. Language tagging in bilingual free recall. *Journal of Experimental Psychology*, 1974, 103, 1137-1141.
- Luce, R. D. Detection and recognition. In R. D. Luce, R. R. Bush, and E. Galanter (Eds.), *Handbook of mathematical psychology* Vol. 1. New York: Wiley, 1963. Pp. 103-189.
- McCormack, P. D. Recognition memory: How complex a retrieval system? *Canadian Journal of Psychology*, 1972, 26, 19-41.
- McCormack, P. D. Recognition memory for common words as a function of target/distractor ratio. *Bulletin of the Psychonomic Society*, 1974, 3, 445-446.
- McCormack, P. D., & Swenson, A. L. Recognition memory for common and rare words. *Journal of Experimental Psychology*, 1972, 95, 72-77.
- Nott, C. R., & Lambert, W. E. Free recall of bilinguals. *Journal of Verbal Learning and Verbal Behavior*, 1968, 7, 1065-1071.
- Paivio, A., Yuille, J. C., & Madigan, S. A. Concreteness, imagery, and meaningfulness values for 925 nouns. *Journal of Experimental Psychology*, 1968, 76 (1, part 2).
- Saegert, J., Obermeyer, J., & Kazarian, S. Organizational factors in free recall of bilingually mixed lists. *Journal of Experimental Psychology*, 1973, 97, 397-399.
- Tulving, E., & Bower, G. H. The logic of memory representations. In G. H. Bower (Ed.), *The psychology of learning and motivation* Vol. 8. New York: Academic Press, 1974. Pp. 279-281.
- Tulving, E., & Colotla, V. A. Free recall of bilingual lists. *Cognitive Psychology*, 1970, 1, 86-98.

(Received for publication May 5, 1975.)