

Efficiency of arithmetic operations in bilinguals as a function of language

LINDA GUTIÉRREZ MARSH and RUTH HIPPLE MAKI
North Dakota State University, Fargo, North Dakota 58102

English-Spanish bilinguals solved simple arithmetic problems and were required to respond in their preferred (P) language, the language in which they originally learned arithmetic, or in their nonpreferred (NP) language. Each arithmetic problem required one, two, or three addition operations. Reaction time was a linear function of number of operations. The intercept for the P language was lower than that for the NP language, but there were no differences in slope. The intercept difference was interpreted in terms of translation time, either as translation of the sum from the P to the NP language or as translation from an abstract representation to the NP as opposed to the P language.

Recently much attention has been focused on determining the role that language plays in the storage of information in memory. Many researchers have attempted to investigate this problem by studying the mental processing of bilinguals, people who are capable of understanding and expressing themselves in two languages.

Kolers (1968) has proposed that bilinguals can store information in two different ways. One hypothesis states that information is stored in a common memory, regardless of the language in which the information was originally encoded. Thus, any information contained in the common store would be equally available to either of a bilingual's languages. The second hypothesis states that the storage of information depends upon which language was used to encode the information. This would imply that bilinguals have two separate memory stores, one for each of their languages. Information encoded in one language, therefore, would not be readily accessible to the second language.

Although studies (e.g., Kintsch, 1970) have shown that neither the common nor the separate stores hypothesis can completely account for a bilingual's storage of information, there appear to be some instances where information available in one language is not easily accessible to the second language. A phenomenon often mentioned by bilinguals (Kolers, 1968) is their tendency to perform arithmetic operations in the language in which they were first learned. Kolers cites the case of one of his colleagues who, having moved from France to the United States during his childhood, does his arithmetic in French and his calculus in English. Since such subjective reports have never been tested empirically, the main purpose of the present experiment was to demonstrate that there exists a difference in time to

perform simple arithmetic in the preferred (P) and the nonpreferred (NP) languages. The preferred language will refer henceforth to the language in which a bilingual learned arithmetic and the nonpreferred language will refer to a bilingual's other language. Preferred and nonpreferred may not necessarily coincide with dominant and nondominant languages in adulthood.

A reaction time (RT) task was chosen. Bilingual subjects were required to mentally add two digits (e.g., $2 + 3 =$), which requires one addition operation: three digits (e.g., $2 + 3 + 4 =$), requiring two operations: or four digits (e.g., $2 + 3 + 4 + 5 =$), which requires three addition operations. They were to give their answer as rapidly as possible in either their P or the NP language. A pilot study with monolinguals had indicated an almost perfect linear relationship ($r^2 = .99$) between RT and number of addition operations.

It was assumed that bilinguals would probably show similar linear relationships in both their P and NP languages; however, it was expected that responses would be faster in the P than in the NP language. Using Sternberg's (1969) methods of analyzing RT data to infer mental processes, one could assess whether bilinguals were translating responses from their P language to their NP language or whether they were performing their calculations in the language in which they were required to respond.

If subjects were translating their answers from the P to the NP language, one would expect to find parallel linear functions with a difference in intercept between the two languages. This difference would reflect the time necessary to translate the answer from the P to the NP language. The function with the lower intercept would correspond to the P language, because once subjects arrived at their answer, there would be no further delay involved in a translation process.

On the other hand, if subjects were performing each task mentally in the language in which the response was to be given, a difference in slope might occur between the two functions. If encoding each numeral into the NP

This paper is based on a thesis submitted by the senior author in partial fulfillment of requirements for the MS degree at North Dakota State University. The research was aided by a grant-in-aid of research from Sigma Xi, The Scientific Research Society of North America.

language or if carrying out each operation in the NP language took longer than in the P language, then, as the problems increased in complexity from one to three operations, it would take progressively longer to respond in the NP language relative to the P language. The increasing difference in RT between the two languages as the complexity of the problem increased would serve to steepen the slope of the function for the NP language.

Although contrary to anecdotal reports, the effect of the P and NP languages might be lost for bilinguals who are approximately equally proficient in their two languages. Therefore, the extent of each subject's degree of bilingualism was determined by using a self-rating scale. Classification of subjects based on self-ratings must be interpreted with caution. However, Macnamara (1967) mentions that self-rating scales on degree of bilingualism correlate highly with other measures of language background, and Fishman and Cooper (1969) have found that the best predictors for the measurement of bilingualism are obtained from retrospective reports. By classifying subjects according to their self-ratings, then, it was possible to determine whether the effect of the P and NP languages was related to degree of bilingualism.

A control group of English monolinguals also participated in the experiment to allow a comparison between them and bilinguals who learned their arithmetic in English. One might expect that, since both of these groups learned their arithmetic in the same language, the RT function yielded by the monolingual group would closely resemble that of the bilinguals. On the other hand, the bilinguals have the additional task of maintaining two linguistic codes relatively active in memory in this task, whereas the monolinguals need maintain only one. This task may add a component to the mental processing of bilinguals which may result in longer RTs for the bilinguals relative to those of the monolinguals. Inclusion of the monolingual control group allowed such a comparison to be made.

METHOD

Subjects

The subjects consisted of 20 bilingual adults, most of whom resided in the Fargo, North Dakota or Moorhead, Minnesota area. Ten subjects had English as their P and Spanish as their NP language (E-S bilinguals). Ten subjects had Spanish as their P and English as their NP language (S-E bilinguals). S-E subjects were drawn from a pool of foreign students attending North Dakota State University during the 1974-1975 school year. A minimum of 1 year's residence in the United States was required for participation. E-S subjects were drawn from foreign language teachers in the area or other qualified bilinguals. Nonlanguage teachers were required to have spent at least 1 year in a Spanish-speaking environment. All subjects were paid for their services.

An additional 20 English monolinguals were also run as a control group. The monolinguals were matched with the bilinguals on the basis of age and sex. Half of the monolinguals were introductory psychology students who received bonus points for their participation. The other half were mostly affiliated with the graduate departments of psychology and chemistry at North Dakota State University.

Questionnaire

Bilingual subjects answered a questionnaire at the beginning of the experimental session. Among a number of filler questions, subjects were asked what language they used during their early years of schooling and specifically which language was used when first learning arithmetic. The answer to the latter was used to designate each subject's P language.

The extent of degree of bilingualism of each subject was also determined. Subjects were asked to rate themselves on a scale from 1 to 9. A 1 on the scale indicated that the subject knew only English; a 9 indicated that the subject knew only Spanish. A score of 5 indicated equal ability with both languages. All subjects rated themselves between 3 and 7 on the scale.

All items on the questionnaire were written both in English and in Spanish. Half of the questionnaires had the English questions first, followed by the Spanish questions, and half were arranged in the reverse order.

Apparatus

Slides of addition problems were projected from a Kodak Carousel 760 slide projector onto a rear projection screen. The presentation of each problem started a Hunter Klockcounter; a Lafayette voice-activated relay stopped the timer and closed the shutter on the projector when subjects answered verbally. The recording of each response with its corresponding RT and the advancing of slides were done manually. A warning signal was displayed immediately before the presentation of each problem.

Materials

The problems were typed on slides. They consisted of one ($p + q =$), two ($p + q + r =$), and three ($p + q + r + s =$) operation problems. The answers ranged in values from 10 to 17. No two digits within a problem were repeated (e.g., $p + p =$ or $p + q + p =$), as the repetition of a digit has been shown to decrease RT (Parkman & Groen, 1971). Each answer was given an equal number of times under each digit condition. All problems were typed to occupy the same amount of space on each slide.

A session consisted of 72 problems, 24 problems under each digit condition, presented in a blocked random order. Each block consisted of two one-operation, two two-operation, and two three-operation problems. Within a session, bilingual subjects experienced four language shifts, E-S-E-S or S-E-S-E. Each language condition contained 18 problems, or three blocks of six slides each. Half of the subjects started the session with problems which required responses in English; the other half, in Spanish. Preceding the presentation of a set of problems to be answered in one language, a slide appeared with the word "English" or "Español," which the subject read aloud. This slide indicated to the subject in which language he was to answer the next set of problems.

Monolinguals performed the same task as the bilinguals but all of their responses were in English. However, when bilinguals saw slides indicating the language of the response, monolinguals were presented with slides containing the phrases Set 1, Set 2, Set 3, and Set 4.

Procedure

Upon completion of the questionnaire, the instructions were read to each subject both in English and in Spanish. Subjects who had English questions followed by Spanish ones on the questionnaire had Spanish followed by English when read the instructions. The reverse occurred for subjects who received Spanish followed by English in the questionnaire.

The subject's task was to solve each problem as fast and as accurately as possible and to respond, in the language specified, into the microphone which he was holding. Six practice trials were given to familiarize each subject with the apparatus and the procedure. Each slide remained visible until the subject responded or until 10 sec elapsed. If the subject did not respond

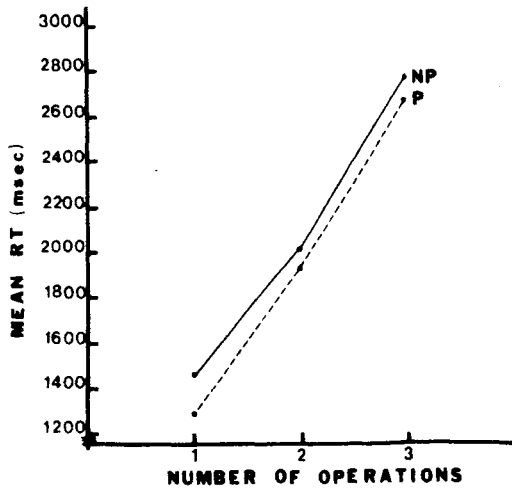


Figure 1. Mean reaction times of bilinguals as a function of language and number of operations.

within the allotted time, the trial was counted as an error. Very few subjects failed to respond within the allotted time. All false starts and incorrect answers were counted as errors and were omitted from the analysis. A subject's verbal response automatically closed the shutter on the slide projector to end the trial.

RESULTS

Type of Bilingual

Figure 1 shows RTs in the P and NP languages as a function of number of arithmetic operations. A mixed-design analysis of variance was conducted with Type of Bilingual (S-E and E-S), Language, and Number of Operations as the factors. Reaction time increased significantly with Number of Operations, $F(2,36) = 171.90$, $MS_e = .10$ sec, which is significant beyond the .05 level, the level of significance to be used in all subsequent tests. Subjects were faster in the P than in the NP language, $F(1,18) = 14.50$, $MS_e = .04$. There was no significant interaction of Number of Operations by Type of Language, $F < 1$.

A strong linear relationship between RT and number of operations was found for both the P and NP language ($r^2 = .99$). The linear equation for the P language was $\overline{RT} = 596 + 684s$; for the NP language, the equation was $\overline{RT} = 796 + 654s$. The difference between the intercepts for the P and NP languages was significant, $t(19) = 2.33$. No significant difference was found between the slopes, $t(19) = .91$. This is in accordance with the finding that no interaction occurred between Type of Language and Number of Operations.

Although the overall mean for the S-E group (1,882 msec) was lower than the overall mean for the E-S group (2,190 msec), this difference was not significant, $F(1,18) = 3.47$, $MS_e = .82$. However, an interaction of Type of Bilingual by Number of Operations was found, $F(2,36) = 4.08$, $MS_e = .10$. As the problem became more complex, the E-S bilinguals took longer than the S-E bilinguals to

respond. No other interactions reached significance, all $F_s \leq 2.88$.

Degree of Bilingualism

Subjects were also divided into high and low bilinguals. Those who rated themselves as equally proficient in both languages ($N = 8$) were regarded as having a high degree of bilingualism, and those who rated themselves as being more proficient in one language than in the other ($N = 12$) were regarded as having a low degree of bilingualism. Table 1 shows RTs for the P and NP language as a function of Number of Operations and Degree of Bilingualism. An unweighted means analysis of variance indicated that there were no differences in performance between the groups differing in degree of bilingualism, $F < 1$. As expected from the previous analysis, RT in the P language was faster than in the NP language, $F(1,18) = 12.22$, $MS_e = .05$, and RT increased with number of operations, $F(2,36) = 134.76$, $MS_e = .12$. There was no interaction of Degree of Bilingualism by Language, $F < 1$, indicating that whether a subject was a high or low bilingual did not affect the P-NP language difference. None of the other interactions was significant, $F_s \leq 1.23$.

Comparison of Bilinguals and Monolinguals

A third analysis compared 10 monolingual English speakers matched by age and sex to the bilingual E-S subjects on English, their P language. Although monolinguals performed all 72 problems in English, for purposes of this analysis, only the data obtained from the problems corresponding to those which the bilinguals performed in English were used. Figure 2 shows RT as a function of number of operations for the bilingual and the monolingual subjects. Mean RT for the monolingual subjects (1,661 msec) was significantly lower than that of the E-S subjects (2,089 msec), $F(1,18) = 5.91$, $MS_e = .47$. Again, as the number of operations increased, RTs also increased, $F(2,36) = 121.50$, $MS_e = .08$, but no significant interaction occurred between the Number of Languages and Number of Operations, $F(2,36) = 2.73$, $MS_e = .08$. The linear equations for the monolingual and bilingual subjects were

Table 1
Mean Reaction Times (Milliseconds) and Errors as a Function of Degree of Bilingualism, Language, and Number of Operations

| | P Language | | | NP Language | | |
|-----------------|----------------------|---------------|----------------|---------------|---------------|----------------|
| | Number of Operations | | | | | |
| | 1 | 2 | 3 | 1 | 2 | 3 |
| High Bilinguals | 1332 (.8) | 1980 (2.5) | 2562 (10.0) | 1476 (8.3) | 2081 (9.2) | 2761 (9.2) |
| Low Bilinguals | 1267 (1.7) | 1916 (5.8) | 2730 (11.7) | 1448 (7.5) | 2097 (5.0) | 2774 (10.8) |

Note—Percentage errors are indicated in parentheses.

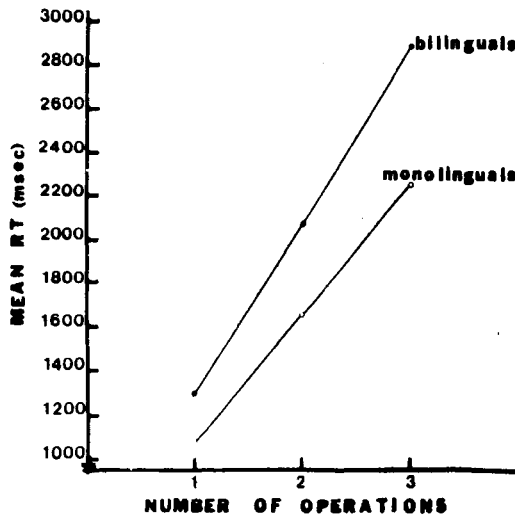


Figure 2. Mean reaction times of English monolinguals and E-S bilinguals as a function of number of operations.

$\overline{RT} = 479 + 590s$ and $\overline{RT} = 491 + 799s$, respectively. Although there appears to be a difference in slope, this was not significant, $t(19) = 1.91$. The difference between intercepts was also nonsignificant, $t(19) = .07$.

Practice Effects

The final analysis attempted to determine whether language shifts disrupted performance, as well as whether working within one language improved performance. Eighteen problems occurred in each language before a language shift. Each of these was split into three blocks of six problems. The \overline{RT} for each block and the order of the block (first, second, or third) was determined for the P and the NP languages. Because each language was used twice during the session, it was further determined whether a block of problems belonged to the first occurrence or the second occurrence of each language. Thus, three variables were analyzed—type of language, order of block within the presentation of a language, and occurrence of the language. Three bilingual subjects were dropped from this analysis because they made more than three errors in at least one block of problems.

A triple repeated-measures analysis of variance indicated that the P language was significantly faster than the NP language, $F(1,16) = 17.96$, $MS_e = .10$. Although the mean for the first occurrence of a language was slightly larger (2,051 msec) than the mean for the second occurrence (1,983 msec), the analysis indicated that performance did not change significantly as a function of whether the subject encountered a language for the first or second time, $F(1,16) = 1.93$, $MS_e = .12$.

Figure 3 shows that RT varied as a function of Block for bilingual subjects. The main effect of Block was significant, $F(2,32) = 14.63$, $MS_e = .06$. Subsequent analyses demonstrated that RT in the first block was significantly longer than RT in the second block, $F(1,16) = 16.11$, $MS_e = .02$, but that the difference

between the second and third blocks was not significant, $F < 1$. It appears, then, that performance improves from the first to the second block of problems and then levels off.

It may be argued that the results which show an increase in RT after a language shift or during the first block of problems may be due merely to the fact that subjects experience a small break when changing languages. The slide which indicates in which language the response is to be given may be causing the increase in RT by virtue of the fact that it interrupts the ongoing task. A similar analysis was performed on 19 monolinguals matched both to E-S and S-E subjects on the basis of age and sex to determine if breaks occasioned by the slides Set 1, Set 2, Set 3, and Set 4 yielded similar functions to the breaks occasioned by the slides indicating the upcoming language. One monolingual subject was dropped from this analysis because he had more than three errors within one block of problems.

A double repeated-measures analysis of variance on the monolingual data indicated that performance did not change as a function of Set, $F(3,54) = 1.31$, $MS_e = .08$, nor as a function of Blocks within a set, $F < 1$. Figure 3 illustrates that, in contrast to the bilinguals' performance, the monolinguals' performance was approximately the same in all blocks. It may be reasonably concluded that the breaks experienced by the monolinguals were not similar to breaks occasioned by language changes in the bilingual condition. The latter produced consistent changes in performance between Blocks 1 and 2; the former did not.

Errors

The overall error rate produced by bilinguals was 6.9%. Table 1 shows the percentage errors as a function

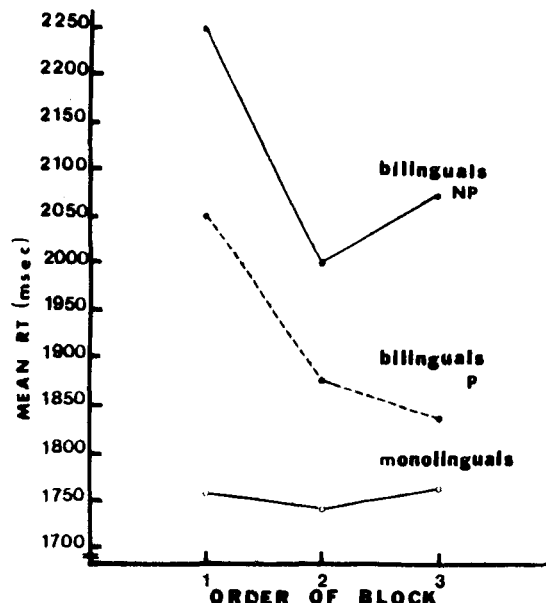


Figure 3. Mean reaction times of bilinguals as a function of language and order of block and of monolinguals as a function of order of block.

of Degree of Bilingualism, Language, and Number of Operations. A mixed-design analysis of variance indicated a significant main effect on the number of errors for Number of Operations, $F(2,36) = 6.02$, $MS_e = .90$. Subsequent analyses demonstrated that problems involving one or two addition operations did not differ significantly in the numbers of errors they produced, $F < 1$, but that problems involving two or three operations did produce different numbers of errors, $F(1,19) = 9.58$, $MS_e = 1.38$. Longer problems, i.e., those with more addition operations, appeared to be more difficult, and this was indicated both in terms of more errors and in longer RTs.

The NP language produced longer RTs than the P language. However, the number of errors produced by both languages did not differ significantly from each other, $F(1,18) = 3.91$, $MS_e = 1.19$. Therefore, the RT effects found between the two languages cannot be attributed to speed-accuracy trading. None of the other main effects or interactions proved to be significant. Although an inspection of Table 1 might indicate that Language and Number of Operations interacted, this interaction did not reach significance, $F(2,36) = 2.99$, $MS_e = .74$.

The monolinguals' overall error rate was 7.6%. The percentage of errors was 5.0, 6.0, 11.7 for one, two, and three operations, respectively. Like the bilinguals, a significant main effect of number of errors was found as a function of Number of Operations, $F(2,19) = 7.37$, $MS_e = 2.01$. Again, problems involving one or two operations did not produce significantly different numbers of errors, $F < 1$, whereas problems involving three operations produced more errors than those involving two, $F(1,19) = 9.54$, $MS_e = 1.91$.

DISCUSSION

Support was found for the hypothesis that arithmetic operations can be performed more rapidly in a bilingual's P than in his NP language, confirming many anecdotal reports by bilinguals. One explanation for the parallel functions obtained between the P and NP languages is that subjects perform their arithmetic in their P language and merely translate the answer into their NP language. The intercept difference of 200 msec would be the time necessary for the translation process to occur. This explanation coincides with many bilinguals' subjective reports.

Another alternative explanation is that subjects perform their arithmetic in the abstract without the aid of language. Translating the sum from the abstract to the NP language could take longer than translating to the P language. If one assumes that a numeral such as 7 is an abstract representation of the number and that the representation of such a number in memory is not coded by a specific language, then a task where subjects merely

name the abstract concept might indicate whether such a translation difference could explain the intercept difference. If there was no difference between naming numbers in the two languages, then one could rule out the explanation that arithmetic was performed in the abstract and that the intercept difference was the result of translation from the abstract representation.

A brief task requiring bilinguals to name the numbers used in the answers of the experimental problems was conducted prior to the main experiment. This allowed subjects to familiarize themselves with the equipment. The numbers 10 through 17 were presented four times in a random sequence and were named twice in each language. The procedure followed was similar to the main experiment except that no problems were solved. The results indicated that the RT for the P language, 732 msec ($\sigma = 142$ msec) was significantly lower than the RT for the NP language, 822 msec ($\sigma = 192$ msec), $t(19) = 2.36$.

The difference between the two languages in this task was 90 msec, whereas, the intercept difference obtained in the main experiment was 200 msec. One task merely required subjects to name a number; the other required operations to be performed on two or more numbers. Because the tasks were different, the difference between the two cannot be interpreted unequivocally. However, the number naming task did suggest that, even at a very elementary level, differences between the two languages can be observed. Therefore, both hypotheses—translation from the P to the NP language and translation from the abstract representation to the P and NP language—provide plausible interpretations of the obtained results.

The control comparison between the E-S bilinguals in their P language and the English monolinguals in the same task indicated that, at least for the E-S subjects, maintaining two languages actively in memory increased RTs. Kollers (1966) showed that reading aloud unilingually provided better performance than reading aloud bilingually. The results of the present study extend his conclusion to include the use of two languages to solve arithmetic problems in the same session.

Further supporting the hypothesis that maintaining two linguistic codes active in memory hampers performance, it was found that immediately after a language shift, bilinguals' RTs increased. Subsequent performance within one language improved after a few trials and then leveled off, but the bilinguals' RTs remained longer than those of the matched monolinguals. These results also support those found by Kollers (1966) and by Macnamara, Krauthammer, and Bolgar (1968), who demonstrated that changing languages takes time and that working unilingually is generally more efficient than working with two languages simultaneously. The increase in RTs during the first block of trials following a language shift is probably

due to changes which were necessary to respond in the new language. Since the monolinguals' performance was not disrupted systematically when similar breaks occurred during the same task, the bilinguals' increase in RT may be attributed to the language change.

REFERENCES

- FISHMAN, J. A., & COOPER, R. L. Alternative measures of bilingualism. *Journal of Verbal Learning and Verbal Behavior*, 1969, 8, 276-282.
- KINTSCH, W. Recognition memory in bilingual subjects. *Journal of Verbal Learning and Verbal Behavior*, 1970, 9, 405-409.
- KOLERS, P. A. Reading and talking bilingually. *American Journal of Psychology*, 1966, 79, 357-376.
- KOLERS, P. A. Bilingualism and information processing. *Scientific American*, 1968, 218, 78-86.
- MACNAMARA, J. The bilinguals' linguistic performance—A psychological overview. *Journal of Social Issues*, 1967, 23, 58-77.
- MACNAMARA, J., KRAUTHAMMER, M., & BOLGAR, M. Language switching in bilinguals as a function of stimulus and response uncertainty. *Journal of Experimental Psychology*, 1968, 78, 208-215.
- PARKMAN, M., & GROEN, J. Temporal aspects of simple addition and comparison. *Journal of Experimental Psychology*, 1971, 89, 335-342.
- STERNBERG, S. Memory-scanning: Mental processes revealed by reaction-time experiments. *American Scientist*, 1969, 57, 421-457.

(Received for publication July 22, 1975;
revision accepted November 24, 1975.)