Psychologica Belgica
1998, 38-3/4, 231-250.

BILINGUALISM AND NUMERIC COGNITION

Marie-Pascale NOËL
Catholic University of Louvain

Wim FIAS
University of Leuven

This article is a critical synthesis of the work which has been done in the field of bilingualism and calculation. The theoretical frameworks dealing with the question of the possible influence of the format or of the language on the retrieval of arithmetical facts are described. The results of the research are discussed in detail by examining the possible alternative interpretations. Encoding and production processes, as well as the possible influence of short-term memory (which may vary according to the subject's language), are the main considerations.

Introduction

The study of mathematics has taken a particular importance in the debate trying to determine the characteristics of information storage and manipulation in bilinguals for two main reasons. Firstly, anecdotal reports on bilinguals have repeatedly suggested that, despite their high proficiency in their second language, they keep performing arithmetic in their mother tongue, or more precisely, in the language in which they were taught the arithmetical operations (Kolers, 1968; Shannon, 1984). Secondly, research on simple additions or multiplications (up to 9 + 9 or 9 x 9) has converged into a global consensus according to which most humans from about 8-10 years of age retrieve arithmetical facts from memory (see Cornet, Seron, Deloche, & Lories, 1988; McCloskey, Harley, & Sokol, 1991; Ashcraft, 1992 for recent reviews, but see LeFevre, Bisanz, Daley, Buffone, Greenham, & Sadesky, 1996a and LeFevre, Sadesky, & Bisanz, 1996b, for a larger emphasis on the use of procedural strategies rather than simple retrieval from memory). Resolving simple mathematical problems is thus an excellent way of approaching general questions about the type of representations underlying the retrieval of information from long-term memory in bilinguals, particularly the issue of language specificity in the storage of semantic information. Furthermore, research on bilinguals offers good opportunities to test and disentangle different theories on the organisation and functioning of the cognitive numerical system This article aims at presenting a review of the research dealing with these issues.

This work was made possible by grants from the Belgian Federal Government (I.U.A.P. on Temporal Control of Dynamic Task Situations and the Nature of Human Knowledge Representation).

Correspondence concerning this article should be addressed to Marie-Pascale Noël, NECO/EXPE Faculté de Psychologie, 10 place du Cardinal Mercier, 1348 Louvain-la-Neuve. Electronic mail may be sent to either noel@neco.ucl.ac.be or to wim.fias@psy.kuleuven.ac.be.

In a certain way, we are all perfect bilinguals regarding the numerical domain. Indeed, everyone of us is able to express numbers in both the verbal code (e.g., twenty-three) and the Arabic code (e.g., 23). However, the term "bilingualism" will be considered here in a more restricted way, i.e., in the mastery of two *verbal* codes for expressing numbers.

Global Effect of Bilingualism

The first studies which looked at mathematics in bilinguals considered the theoretical alternatives opposing the independence and the interdependence hypotheses. According to the former, bilinguals have two distinct semantic memory stores for their two languages and these stores are relatively isolated from one another (Macnamara & Kushnir, 1971; Tulving & Colotla, 1970) whereas in the interdependence hypothesis, bilinguals are assumed to have only one single storage system with two representational lexical systems of access to that storage (e.g., Glanzer & Duarte, 1971). According to the first hypothesis, the performance of mono- and bilinguals should not differ significantly provided that no language switch is necessary. For the second hypothesis, however, accessing the storage necessitates the choice of representation which takes time and which gives rise to interferences.

Two studies have compared strong bilinguals to weak bilinguals or monolinguals in the realisation of mathematical operations. In Mägiste's (1980) study, weak and strong German-Swedish bilinguals (who were living in Sweden but who had learned their mathematics in German) were compared when performing four arithmetical operations (addition, multiplication, subtraction, division) of varying complexity (i.e., from 1 to 3 operations; e.g., "11 + 4 = ?", "22 + 2 + 3 = ?" and "45 + 3 + 1 + 5 = ?"). All the problems were presented in the Arabic format and responses were to be written in the same code. Under these conditions, the weak bilinguals were shown to take less time than the strong bilinguals and this advantage increased with the number of operations performed in both the additions and the subtractions but not in the divisions and in the multiplications. The disadvantage of the balanced bilingual group was interpreted as being due to their special difficulty in disregarding one of the languages when solving arithmetical problems. The authors considered this interference between the two languages as an argument in favour of the interdependence hypothesis which assumes the existence of a single storage system.

Somewhat different results were obtained by Geary, Cormier, Goggin, Estrada and Lunnhave (1993). In their two studies, problems were presented in the Arabic code and subjects were asked to judge the correctness of the given solution (also written in digits). In the first experiment, simple and complex additions (e.g., 8 + 6 = 14 and 7 + 5 + 3 + 6 = 21) as well as simple and complex multiplications (e.g., 8 x 4 =32 and 6 x 3 x 7 x 2 = 252) were compared in three groups of subjects: English monolingual, weak French-English bilinguals and strong French-English bilinguals. The effects obtained by Mägiste (1980) were not replicated: the difference between the groups did not reach statistical significance. A second experiment was then carried out which compared English monolingual, weak Spanish-English bilinguals and strong Spanish-English bilinguals in the verification of simple and complex additions and multiplications but this time complex operations consisted of a single operation performed on large numbers (e.g., 24 + 59 = 83 and 38 x 6 = 228). Once more, the groups did not differ significantly in terms of response times. However, when modelling the response times of each subject by computing regressions with the different sub-processes of the task as predictors (i.e., total number of digits for the encoding step, product of columnar digits for the memory retrieval process and absence vs. presence of a carry operation for the carry step), Geary et al. found relatively slow processing speeds for the carry operation in strong bilingual subjects when compared with the two other groups in the two complex operations (however, the level of significance was only 0.061). The authors thus concluded that bilingualism does not affect the long-term memory of arithmetical facts but does affect operations such as carry which require working-memory resources.

The Nature of Numerical Representations

The Theoretical Models

A second line of research on bilinguals' mathematics did not aim at comparing different groups of subjects (e.g., mono- and bilinguals) but rather aimed at measuring within-group differences by comparing subjects' performance when directed to calculate in one or other of their two languages. From this perspective, bilingualism is not taken as the primary question under study but as an opportunity to test general models of human mathematical cognition. The main issue tackled here is the nature of the information stored in long-term memory: Are the arithmetical facts stored as a verbal code or are they stored as an abstract language-independent code? And, if they are stored as a verbal code, do bilinguals possess two different stores according to their languages or only one (probably the native language)?

Three main theoretical models are proposed: The abstract model of McCloskey, Caramazza, and Basili (1985), the triple-code model of Dehaene (1992; Dehaene & Cohen, 1995) and the encoding-complex view of Campbell and Clark (1988). According to McCloskey, Caramazza, and Basili, (1985;

McCloskey, 1992), arithmetical processing operates upon a single and abstract format that constitutes the sole support for internal representations of numerical quantities. Comprehension and production processes, specific to each input and output format are dedicated to the translation of numerical input into the abstract representation of quantity or, to the reverse operation for the production modules. In other words, arithmetical-fact retrieval operates on abstract representations and is totally independent of the format used for the presentation of the problem or the production of the response. Although this theoretical framework was not originally constructed to account for the bilingual's cognitive system, extrapolation in this direction is possible. More precisely, one may assume that the bilingual would need a supplementary comprehension system for his/her second language and, similarly, another production system. These systems would be devoted to respectively, the translation of verbal numeral presented in the subject's second language into the format-independent internal code and the translation of a quantity expressed in the abstract code into the appropriate verbal sequence in the subject's second language. However, since calculation is supposed to operate on abstract codes, effects of language are considered to be possible only at the encoding (or comprehension) and production stages but not at the calculation stage.

The triple-code model proposed by Dehaene (1992; Dehaene & Cohen, 1995) posits the existence of three different internal codes for mentally representing numbers: An auditory verbal code, a visual Arabic code, and an analog magnitude code, each one being dedicated to specific processing. For instance, the Arabic code is supposed to mediate multi-digit calculation and parity judgment, the analog magnitude code is used for numerical comparison or quantity approximation and the verbal code is used for verbal counting and arithmetical-fact retrieval. According to task requirements, input is translated into the required code. Although Dehaene postulates the existence of specific and multiple internal codes, it should be made clear that he does not predict different performance patterns as a function of the presentation format. In the case of arithmetical-fact retrieval, for instance, he assumes that every type of input is first translated into the verbal code before arithmetical-fact retrieval takes place. "We suppose that arithmetic facts such as 2 x 3 = 6 cannot be retrieved unless the problem is coded into a verbal code "two times three ...", which then triggers the retrieval of the result "six" in the same verbal format. " (p. 87, Dehaene & Cohen, 1995). In the case of bilinguals, a fourth internal code, i.e., a verbal code for the second language, should obviously be introduced. However, besides comprehension and production of numerals in that language, the role that this fourth internal code would play in numerical operations is not obvious. Concerning arithmetical facts in particular, Dehaene (1992) argues that "bilingual subjects access them [i.e., the arithmetical facts] in the language that they acquired first" (Dehaene, 1992, p. 33). This quote suggests that he considers the existence of a single verbal store for the arithmetical facts in bilinguals so that when calculations are presented in the other language, they would first be converted into the teaching language in order to access the solution. In other words, the verbal internal code in the subject's second language would not be used for the retrieval of arithmetical facts. It would only be used to comprehend and produce verbal numbers in that language and allow internal translation from this internal code to the other ones. In summary, the predictions from Dehaene and McCloskey's models are the same: Both would expect the absence of format-specific influences on retrieval itself. If format-specific effects are observed, they should be attributable to encoding and/or production.

In contrast, the encoding complex model of Campbell and Clark (1988; Clark & Campbell, 1991) emphasizes the importance of the presentation and production formats for arithmetical-fact retrieval. More specifically, they assume that arithmetical-fact retrieval is mediated by format-specific representations, which are thought to be numerous: Verbal (i.e., articulatory, auditory, orthographic, and motor codes for various spoken and written number words) and non-verbal (e.g., visual and motor codes for digits). The model thus upholds that numerical processing in general, and arithmetical-fact retrieval in particular, "may differ as a function of the presentation format because Arabic and verbal inputs, e.g., potentially provide independent pathways to number-fact representations" (Campbell, 1994, p. 6-7). Recently, Campbell and Kanz (1999) have proposed some clarifications of their theoretical considerations regarding the cognitive system of bilinguals. They assume the existence of two verbal number systems which include memory for number facts. Numerical stimuli may activate retrieval processes simultaneously in both language systems (selection procedures as well as inhibitory processes are thus hypothesized) but different surface forms can vary in their capacity to call operations and representations in the various systems. However, for those who have learned their second language relatively late, it is possible that only the small arithmetical facts are stored in their second language verbal store and that the solution of larger problems requires a translation into the first language.

In summary, both McCloskey et al. (1985) and Dehaene's (1992) model predict differences between the first and second languages of bilinguals in terms of encoding or production mechanisms but assume that the retrieval stage is unaffected by the language since calculation is carried out on a single internal code (i.e., the abstract code in McCloskey et al.'s perspective or the verbal code in the language of arithmetical schooling for Dehaene) no matter what the presentation or production languages are. On the contrary, Campbell and Kanz (1999) expect differences between the two languages that go beyond the encoding and output stages: The retrieval process itself is affected by the input or output language.

Empirical Evidence

Several studies have shed light on the validity of these perspectives. The first research was carried out by Marsh and Maki (1976). They presented addition problems in Arabic format to English-Spanish[1] and Spanish-English bilinguals who were instructed to respond in their first or second language. The number of additions per problem varied from one (e.g., 3 + 4) to three (e.g., 3 + 5 + 4 + 2). They reasoned as follows: If arithmetical operations are performed in the first language irrespective of response requirements, then the two language conditions should give rise to parallel linear functions according to the number of calculations required. If, on the other hand, subjects perform each addition mentally in the language in which the response has to be given, then a difference in slope might occur between the two functions. This difference could be due to the fact that encoding each numeral into the non-preferred language takes more time than into the preferred language, or to the fact that carrying out each addition takes more time in the preferred language than in the non-preferred language. As expected the authors found a strong linear relationship between RT and number of operations and this, for both the preferred and the non-preferred language. The intercept was also bigger in the non-preferred condition than in the preferred condition (i.e., subjects were faster to respond in their preferred language than in their non-preferred language) but no difference in slope was obtained.

The same paradigm was applied by McClain and Huang (1982) but, to put more constraint on the subject's language of processing, additions were presented orally in either the first or the second language. Participants were instructed to respond in the same language as that used to present the problem. The experiment was first run with Chinese-English and English-Chinese subjects and then with Spanish-English and English-Spanish bilinguals. Both studies replicated exactly Marsh and Maki's findings, i.e., faster response times in the preferred language than in the non-preferred language but no interaction with the number of operations required. The absence of interaction between the language and the number of operations in both Marsh and Maki and McClain and Huang makes an organisation of the bilingual's numerical cognitive system with a dual memory store (one for each language) less likely and favours a single memory store hypothesis. At least two possible interpretations of this hypothesis can be considered. According to the first one, subjects calculate on the basis of their preferred language and then, if necessary, translate into the second language for responding (as assumed in Dehaene's triple-code model). Alternatively, it is possible that the calculation is not carried out on verbal

[1] By convention, we will systematically write the subject's preferred language first, followed by his/her second language.

representations but on a format-independent code (as in McCloskey's perspective) but translating numbers from the preferred language into this abstract code or the reverse is faster than translating from and to the non-preferred language.

Frenck-Mestre and Vaid (1993) were interested in more subtle influences of the language: They measured interference effects which are generally taken as evidence of the automatic activation of arithmetical facts and their neighbours in the network. In the first experiment, participants were asked to add mentally two visually presented numbers and to decide whether the sum equalled a subsequent presented probe. This probe was either the correct sum, or a number that differed from the sum by a small or a large split. The presentation format of the probe and the pair of numbers was manipulated (digits, first, and second language), as well as the SOA between the presentation of the numbers and the probe. It was found that English-French bilinguals were slower to respond to problems presented in the second language than in the first language, and were slower for problems presented in word format than in digit format. The effect of the split (i.e., longer rejection times for small splits than for large splits) was observed in each of the three presentation formats and the strength of this effect was indistinguishable across the two language conditions. The only effect of the language noticed by the authors was a difference between true-sum probes and non-sum probes in the digit and the first language conditions but not in the second language condition. However, it should be noted that there was only a trend for an interaction between type of probe and presentation format ($p < .08$).

Virtually the same procedure was used in their second experiment except that this time there were both addition and multiplication problems and that the probe types were different. The proposed solution could either be the true answer to the problem -true probes-, the correct solution for the other operation (i.e., the sum of the addends in case of a multiplication problem and the product of the addends in case of an addition problem) -associatively-related probes-, or mathematically unrelated to the problem -neutral probes-. In addition to the main effect of presentation format, the authors found that the associative confusion effect (i.e., the difference between the rejection of associatively-related vs. neutral probes) was present in the digit and the first language condition but not in the second language condition. However, the authors recognized that the overall longer processing times in the second language condition may have masked the appearance of the associative confusion effect, which is short-lived.

Finally, Campbell and Kanz (1999) have recently presented a very detailed investigation of number processing in Chinese-English bilinguals. Among other tasks, they asked participants to carry out simple addition or multiplication problems. The presentation format varied randomly between Arabic digits and Chinese symbols and the response had to be given orally in either Chinese or

English according to the instruction given prior to each trial. Detailed analyses revealed several interactions with input or output formats. The main findings concerning response times can be summarized as follows: (1) Subjects were slower to respond in English than in Chinese especially when the problems were presented in Chinese symbols; (2) the difference between addition and multiplication was bigger in English than in Chinese output; (3) the size effect (i.e., the difference between the realization of small vs. large problems) was smaller for equations with Arabic digits than for Chinese symbols, and was also smaller for Chinese responses than for English responses. Effects of the input and output formats were also obtained in the analyses of arithmetical errors: (1) Chinese equations provoked more operation errors, more tabled with intrusion errors and more distant errors than equations written in Arabic digits; (2) English spoken responses promoted more miscellaneous errors than Chinese responses. Finally, some interesting effects emerged from the analysis of "language errors", i.e., the fact of answering in English when required to answer in Chinese or the reverse. Indeed, when cued to answer in English, language errors (i.e., answering in Chinese instead) occurred as frequently for small and large problems but when cued to answer in Chinese, they occurred more often on small than on large problems.

Intermediate Conclusion

All the research concerning the calculation abilities of bilinguals has shown an advantage for the preferred language over the non-preferred language. Thus, Marsh and Maki (1976) who manipulated the language of production of the answers and McClain and Huang (1982) who manipulated both the presentation and production formats observed better performance in the subject's first language. A main advantage of the preferred language was also obtained by Frenck-Mestre and Vaid (1993) in the verification of additions as well as by Campbell and Kanz (submitted) who observed faster answers in Chinese than in English.

Most of these effects can be explained in terms of encoding or production processes. Subjects should be slower to encode numerals presented in their second language (Frenck-Mestre & Vaid, 1993, McClain & Huang, 1982) and also slower to produce a verbal numeral in that language (Marsh & Maki, 1976; McClain & Huang, 1982; Campbell & Kanz, 1999). These effects on relatively peripheral processing steps might also account for some differences between the two languages which at first appear to bear on more central processing stages. For instance, the absence of a confusion effect in the second language reported by Frenck-Mestre and Vaid (1993) might have been due to the overall longer processing times in the second language condition which may have masked the appearance of the short-lived associative confusion effect. Similarly, the operation-by-output-language obtained by Campbell and Kanz can also be accounted for by the difference in the syntactical structure of Chinese and English.

In other words, these results do not challenge theoretical models assuming the existence of a single store for the arithmetical facts. Indeed, main effects of the language are easily explained in McCloskey et al.'s (1985) or Dehaene's (1992) perspective. In McCloskey et al., one could consider that translating a verbal numeral presented in the second language into the abstract code or the reverse operation takes more time than translating a verbal numeral of the first language into the abstract code or the reverse operation. In Dehaene's triple code model, arithmetical facts are supposed to be accessed through the verbal representations of the first language. Thus, problems presented in the subject's first language can directly access the arithmetical-fact store and these answers can be directly spoken out loud since they are already in the required verbal format. By contrast, problems presented in the subject's second language require a supplementary step, i.e., the translation from the second language internal code to the first language internal code. Similarly, when responses have to be produced in the subject's second language, a translation from the first language to the second language is necessary. All these assumptions therefore predict a global disadvantage when either presentation format or response production is in the subject's second language. Thus, all empirical data discussed so far can be reconciled within a single-store framework attributing all differences between first and second language to peripheral encoding or production stages of numerical processing. From these data no definite conclusions can be drawn regarding the question whether this single store is presented in a (or several) verbal or a non-verbal code.

These models are, however, more strongly challenged by the data obtained by Mägiste (1980) and, by Campbell and Kanz (1999). Indeed, Mägiste showed that strong bilinguals have a lower performance than weak bilinguals when calculations were presented in the Arabic code and responses had to be written down in the same format. Furthermore, this difference increased with the number of operations. These effects cannot be attributed to encoding or production processes since all the subjects had to encode and produce Arabic numbers. One could possibly argue that these differences are due to a bad selection of subjects, the two groups being not equivalent at first. However, this explanation is not very convincing. Indeed, the subjects were selected so that the two groups did not differ regarding their grades in the mathematic class and they proved to be equally effective in the realization of single operations (for addition and subtraction). Another interpretation of the results should therefore be sought.

With regard to Campbell and Kanz's research on Chinese-English bilinguals, they reported not only main effects of input and output formats but also several interactions of these variables with factors indicating the arithmetical-fact retrieval stage (such as the size effect), and even input-output interactions. Some of these interactions could possibly be explained by differences in encoding or production stages. For instance, they reported a larger difference between addition and multiplication in English than in Chinese output. However, answers to multiplications correspond to bigger numbers than answers to additions. If, for some reason (such as the simpler syntax of Chinese which corresponds more closely to the Arabic number system), numbers are faster to produce in Chinese than in English, and this, even more as we go up to the larger numbers, then one should expect an interaction between language and operation and it would be attributed to a difference in the production stage rather than to the retrieval stage. Other interactions reported by Campbell and Kanz (input format-by-size; input format-by-output format) are not easily attributed to encoding or production stages and provide a real challenge for the hypothesis that arithmetic facts are stored in and retrieved from a single store. This, however, does not mean that the single code hypothesis cannot be maintained. Indeed, we will argue that there is another way through which language can play a role during mental arithmetic, i.e., the temporary storage of information in the short-term memory system (STM). By considering the possibility that STM processes could be influenced by characteristics of the subject's language, we will examine whether this could cause language differences in arithmetical tasks, without having to assume that two language stores are involved in retrieval.

The next section will be structured as follows. Firstly, we will consider the research that has underlined the possibility of differences in STM storage capacities according to the subject's language. Thereafter we will discuss possible implications of these findings for mental arithmetic.

Language and Short-Term Memory

In comparing people speaking different languages, it has been observed that digit span may vary with the subject's language. For instance, Chincotta and Underwood (1997) tested speakers of Chinese, English, Finnish, Greek, Spanish and Swedish and showed that Chinese speakers had a larger digit span (mean of approximatively 8.9) than speakers of the other languages, which did not differ from one another (for instance, mean of digit span in English is approximatively 7.5).

More impressively, similar differences were observed within single individuals who were tested in different languages. So for instance, Chinese-English bilinguals were shown to obtain a higher forward digit span when measured in

Chinese than in English (Hoosain, 1979; 1984). Similarly, Elliott (1992) studied subjects who spoke both English as well as one Asian language (Malay, Mandarin or Hokkien). He found no difference between the groups when the digit span was measured in English but differences were seen when the span was measured in the Asian language with Malay spans being shorter than those for Mandarin and Hokkien. Globally, spans in Malay were shorter than spans in English (for the same subjects) and spans in Mandarin were longer than those taken in English.

The results reported by Ellis and Hennelly (1980) indicated that it was not so much a global advantage for the preferred language. Indeed, they showed that Welsh-English bilinguals presented larger digit spans in English than in Welsh. For Ellis and Hennelly, the superiority of the English over Welsh was due to a word-length effect: Welsh number names are longer than English number names and are, accordingly, more slowly read.

This suggestion was clarified by da Costa Pinto (1991) who showed that it was not so much the word length that was important but rather the articulation speed. For instance, number-names in Portuguese have more syllables than in English but Portuguese-English bilinguals were faster to read digits in Portuguese than in English and presented, consequently, larger digit spans in Portuguese than in English. Similarly, Chincotta and Hoosain (1995) observed that English-Spanish bilinguals presented both faster number reading and larger digit spans in English than in Spanish and this, regardless of self-rated proficiency in these two languages. Similarly, Chinese-English subjects presented a superior digit span in Chinese than in English but were also faster in reading numbers in Chinese than in English. However, for these two groups of subjects, articulatory suppression completely erased the difference between languages. This relationship between digit span and reading rate has also been obtained by Hoosain (1982) who found a high correlation between digit span and digit pronunciation in both Chinese and English for Chinese-English bilinguals (correlation of .66 in Chinese and of .70 in English). Hoosain and Salili (1987) replicated the advantage in Arabic digit pronunciation time for Chinese (314 ms) over English (375 ms). Moreover, these pronunciation times were differently affected by the magnitude of the numbers. In English, small numbers (from 2 to 5) were pronounced faster than large numbers (from 5 to 9) (respectively, 358 ms and 404 ms) whereas in Chinese, small numbers took more time to be pronounced than large numbers (respectively, 395 ms and 251 ms).

These particular effects are easily interpreted within the STM model proposed by Baddeley (1990). According to this framework, items are encoded in an articulatory code and need to be constantly refreshed to counter the natural decay of information. This can be done by subvocal rehearsal of the material in a phonological loop. It is assumed that the delay before decay in this system is approximatively 2 sec. The subjects' STM capacity is thus limited to roughly the amount of material that can be rehearsed subvocally in about 2 sec which can be predicted on the basis of the number of words the subject could read in that temporal delay. This model can also account for the fact that the differences obtained between digit spans measured in different languages disappear when articulatory suppression procedures are used (as in Chincotta & Hoosain, 1995, for instance) since in these conditions, the information is not stored in the articulatory loop system but rather in the visuo-spatial scratch-pad which does not use verbal mechanisms to refresh the information.

Bilingualism, Short-Term Memory and Calculation

Given that STM may play an important role in calculation, Ellis and Hennelly (1980), among others, have suggested that factors such as number-name articulation speed which limit STM storage could make mental arithmetic more difficult. In the context of their study with Welsh-English bilinguals they predicted that "Any operation which involves remembering numbers (anything from mental arithmetic to the remembering of telephone numbers) will be more difficult to perform in the Welsh language than in the English language" (p. 51).

Several attempts have been made to test this hypothesis. Firstly, Hoosain and Salili (1987) measured the correlation between the size of the digit span and the grade obtained by Chinese-English bilinguals on a mathematics achievement examination and found a significant correlation of .38. Unfortunately, these authors did not take advantage of the particularity of their population: both the span and the math test were carried out only in English. Secondly, the authors used a global examination of mathematical abilities rather than devising a special task in which stimuli and processes could be controlled in detail.

The investigation reported by Elliott (1991) suffered from similar weaknesses. He studied four groups of bilinguals: Malay-English, English-Mandarin, Mandarin-English and Hokkien-English. The digit span measured in English was comparable in the four groups but Malay spans were lower than English spans and spans in Mandarin and Hokkien were larger than those measured in English. Some of these differences were also reflected in a mathematical task (not specified): Malay-English subjects were slower to do maths problems in Malay than in English but the Mandarin-English and Hokkien-English groups did not show any advantage for the Mandarin or Hokkien when compared with the English. Finally, the English-Mandarin subjects even showed a disadvantage for the Mandarin problems when compared with the English problems. According to Elliott, familiarity should be taken into account for interpreting these results. Indeed, a measure of the difference between reading rate and articulation rate indicated that all, but the first group, were more familiar with digits in English than in the other language.

A third attempt to evaluate the impact of the digit span size on numerical tasks was reported by Ellis (1992). In the first research, Welsh-English and English children were compared in a calculation task in which working-memory load was manipulated. Although the items were presented in Arabic digits and responses had to be typed in the same code, Ellis predicted that Welsh-English children would be slower and less accurate than their English peers for problems which involved an appreciable STM load, because Welsh number-names are much longer and take more time to pronounce than the English ones. Accordingly, Welsh-English children proved to be as fast as the English in the simple multiplication task in which a single answer had to be retrieved from long-term memory (e.g., 5 x 3 =), and in multiple additions which did not require a strong memory load because no carry was required (e.g., 305 + 42 =). But the Welsh were slower than the English in the two types of problems which involved carrying and many interim stages (i.e., multiple addition with carrying: e.g., 134 + 88 = and multiple successive additions: e.g., 5 + 3 +7 + 4 + 9 + 8 + 6 + 5 + 3 =).

A second experiment was then devised in order to show that these differences were due to the temporary storage demands rather than to the complexity of the calculation itself. In this study three groups of subjects were compared (Welsh-English and English-Welsh bilinguals as well as English monolinguals) in two tasks. Firstly, they had to count as quickly as possible from 1 to 100. Secondly, they had to read series of numbers which corresponded to the addends of the problems presented in Experiment 1, the interim numbers that needed to be generated for realizing the calculation and the solution obtained (e.g., for the problem 204 + 41, the corresponding sequence of digits to read was 1 4 5 4 0 4 2 245 since 1 and 4 gives 5, 4 and 0 is 4, and 2, thus making a total of 245). It appeared that both counting time and reading time were much longer in Welsh than in English. More importantly, in the reading task, this difference was modulated by the type of problem with a larger difference between Welsh and English in the multiple addition with carry and in the multiple successive additions. A clear correspondence was thus found between the time it took to name the numbers involved in the interim calculations and the actual response times and error rates measured when people actually perform the sums in the two languages.

Finally, a third experiment was carried out in a more natural setting. Welsh and English children were given lists of calculations to perform (by writing) in school, without any speed pressure. There were both simple and complex additions, subtractions, multiplications and divisions. Under these conditions, the disadvantage of Welsh over English was once more replicated: Welsh children made an average of 9.7 errors whereas their English peers produced only 7 errors. It would have been ideal to constrain the language used by Welsh-

English bilinguals when they calculate and find within-subject differences showing that they are better when pushed to calculate in English than in Welsh but such a comparison was not carried out.

In summary, these studies indicate that (1) the size of the digit span varies with the articulation speed, and thus, may vary within a single individual according to the language of the test; (2) mathematical performance is affected by the size of the span, especially when memory load is important; (3) the short-term storage of the addends of the calculation or of the interim results seems to use an articulatory code, even when numbers are presented in the Arabic code (this proposition has also been suggested by Brysbaert, 1995, Experiment 2[2]):

These results may shed some light on the data obtained by Mägiste (1980). She found no difference between weak and strong bilinguals in performing simple addition and subtraction but a clear disadvantage for strong bilinguals when the calculations were more complex and thus, required more temporary memory resources. Now, we may wonder whether and why these two groups could differ in terms of working-memory capacity. Memory span was not measured but it is interesting to note that, after the experiment, the two groups differed concerning the language in which they believed they had performed the calculations. Most of the weak bilinguals thought that they had calculated in German but the strong bilinguals' reports were more heterogeneous: 41 % of them reported having used German, 31 %, Swedish and the remaining 28 %, a mixture of the two languages. It is plausible that these different profiles induced differential memory load for the two groups of subjects. Firstly, it is possible that the articulation speed differed according to the group and/or according to the language used to store information in STM. Unfortunately, we do not have any data on this point. Secondly, for 28 % of the strong bilinguals, two different languages were used to store the information while calculating. This double verbal coding may have increased the interference and complicated some situations. This might have been particularly the case for German-Swedish bilinguals since these two languages present a very different syntax for expressing verbal numbers. Indeed, whereas Swedish uses the general Decade-Unit structure for naming two-digit numbers, the opposite order is used in German (e.g., 43 is not said "forty-three" but "three and forty"). It is well

[2] In this experiment, subjects were presented with three successive Arabic numbers that they had to store and then indicate whether a comparison stimulus belonged to the memory set or not. In these conditions, encoding of the Arabic numbers in the set (measured by the gaze duration) was significantly affected by the syllable length of the corresponding verbal numerals. According to Brysbaert, these results indicate that at some point a visual number code was translated into an auditory code, probably because the auditory code facilitates short-term memory rehearsal in a rote-learning situation. It should be noted, however, that the implication of verbal translation of the visual stimuli for temporary storage does not imply an absence of parallel and redundant coding of the same information in the visual scratch pad.

possible that switching from one code to another and from one syntactical system to another has increased the response times of the strong bilinguals not because they had two different verbal arithmetical-fact stores, one in each language, but because these subjects suffered more interference when manipulating and storing information in their verbal working-memory system.

Similarly, working-memory factors could account for some results obtained by Campbell and Kanz (submitted), even though the task used consisted in performing simple additions or multiplications. Indeed, the involvement of working memory resources in simple arithmetical fact tasks should be considered. Two studies (Ashcraft, Donley, Halas, & Vakali, 1992; Lemaire, Abdi, & Fayol, 1996) have investigated this question by using a verification paradigm with concurrent tasks loading the articulatory loop (e.g., (1) repetition of a single letter) and the central executive (e.g., generating words beginning with a given letter). Their results suggest an involvement of the WM system, and in particular, of the central executive. However, further investigation is required to examine whether this implication of the central executive is at the level of the comparison/decision stage or whether it is also revealed in the retrieval stage. Secondly, Lefevre, Sadesky and Bisanz (1996b) have argued that a non-negligible part of single digit problems are solved by strategies other than mere retrieval, for instance counting and transformation. This is especially true for larger numbers. These strategies can be expected to require a substantial involvement of STM processes. Indeed, Healy and Nairne (1985) demonstrated the role of STM processes in keeping track of the location in the counting sequence.

In the case of Campbell and Kanz's investigation, "The problem-size-by-language interaction and the operation-by-language interaction indicate that typically-difficult arithmetic facts were especially difficult to access in English whereas the English language cost was relatively small for typically-easier facts " (p. 23). If, as suggested by Lefevre et al. (1996b), more difficult problems require a decomposition into smaller problems, then a retrieval of the smaller sub-problems and finally, a combination of these solutions in order to obtain the answer of the problem presented, they could possibly involve the WM system. This is even more clearly the case in Campbell and Kanz's experiment since subjects had to carry out these processing steps while keeping activated the instruction given prior to the trial "answer in Chinese" or "answer in English". If the subjects actually followed these instructions and manipulated the information accordingly in English or Chinese in their articulatory loop, then difference in the size of the span according to the language might have influenced the response times. We have presented numerous results showing that the size of the digit span could vary according to the language in which it is tested, even within a single individual. Among these studies, all those that have compared English and Chinese have shown larger spans in Chinese than in English (Chincotta & Hoosain, 1995; Chincotta & Underwood, 1997; Hoosain, 1979, 1982, 1984). Accordingly, one could expect that, as the task requires more STM capacity, the influence of the size of the span would be greater. Since span is probably greater in Chinese than in English, processing in English should show a disadvantage compared to Chinese which would grow as the complexity of the calculation increases. This is exactly what Campbell and Kanz's results show.

Conclusion

This article presents an overview of the research which has examined the performance of bilinguals in calculation tasks. The core theoretical question addressed by these studies is whether the arithmetical-fact store uses verbal representations and, in the affirmative, whether bilinguals possess one sole memory store, or whether a second language-specific store develops apart from the first one, during the acquisition of the second language. Two paradigms have been used to get insight into this controversial issue.

A first line of research has looked at global effects of bilingualism on calculation presented and responded to, in the Arabic code. In this respect, Mägiste (1980) compared weak and strong bilinguals and found a disadvantage of strong bilinguals, increasing with task complexity. This global effect, however, was not replicated by Geary et al. (1993). The only disadvantage that they found in strong bilinguals was a subtle slowerness to perform carry-over. Taken together, the search for global effects did not substantially increase our insight in the controversy.

Other researchers have examined the effects of imposing one of the two languages when performing calculation by manipulating input and/or output modality. The rationale behind this view was that interactions between language and factors stemming from the retrieval stage of arithmetical facts are an indication of language-specific calculation, i.e. language-specific storage and retrieval of arithmetic facts. All of these studies reported a global advantage of the subject's preferred language (e.g. Marsh & Maki, 1976; McClain & Huang, 1982; Frenck-Mestre & Vaid, 1993; Campbell & Kanz, 1999). But, what was more important is the presence or absence of the crucial interaction. This yielded controversial results too. Marsh and Maki did not find an interaction of language with the number of additions per problem. The same failure was reported by McClain and Huang. Similarly, Frenck-Mestre and Vaid did not observe any influence of the language on the split effect in a verification task. However, the same authors reported a greater associative confusion between addition and multiplication in the preferred language than in the non-preferred language. Campbell and Kanz reported the strongest evidence in favour of the dual code hypothesis. They observed effects of input format on the size effect as well as an influence of the output language on the operation effect (i.e. the difference between addition and multiplication) and on the size effect.

In this article, we have tried to unravel the controversy and to propose reasons which make the apparently incompatible data a little more compatible. We have argued that some of the critical interactions between language and calculation might stem from encoding or production stages rather than from retrieval as such (e.g. the differential associative confusion effect obtained by Frenck-Mestre and Vaid or the operation-by-output language interaction obtained by Campbell and Kanz). More importantly, we have pointed out that language may affect other stages than long-term storage, encoding or production. Indeed, some of the STM processes are language-based too. In this respect, we proposed that STM constraints could differentially influence subject's calculation performance according to the language used to deal with the task. Such a proposition was made by Ellis (1992) who showed that children speaking Welsh had both lower digit spans and worse performance in calculation tasks than their English-speaking peers. It is argued that the influence of language on STM spans is mediated by articulation speed and could account for the observation of specific interactions between language and calculation (e.g. the larger effect of calculation complexity reported by Mägiste (1980) in bilinguals or the influence of the output language on the problem-size effect and the operation effects observed by Campbell and Kanz).

By pointing to these alternative interpretations in terms of encoding/production processes or differential STM capacity according to the subject's language, we do not aim at dismissing all the results indicating critical effects of the language on calculation. We do indeed think that these types of evidence cannot all be easily explained by single store models, at least as they are formulated now. For instance, we did not find good alternative interpretations for the differential types of error observed in Chinese and English as reported by Campbell and Kanz. In playing the "devil's advocate", we wanted to examine all the possible factors that could explain these results while keeping the assumption of a single arithmetical-fact store. We indeed think that the debate between a single store for arithmetical facts versus multiple verbal stores according to the subject's languages cannot be closed until such alternative interpretations have been ruled out.

This article has been centred around the issue of the possible influence of a bilingual's language on performing calculation tasks. However, the domain of number processing is much wider than tasks like addition, multiplication, subtraction or division. A wide territory needs to be discovered. Nearly nothing is known about the influence of bilingualism on other tasks such as parity judgment or magnitude comparison (but see Campbell & Kanz for first data on this latter point). Similarly, there has been nearly no interest for studying the acquisition of number (but see Bialystok & Codd, 1997) and number names in bilingual children (but see Mägiste, 1992). The influence of the lexico-syntactic structure of a verbal system in tasks requiring the manipulation of numbers has also received too little attention (but see Miller & Zhu, 1991). For instance, Noël and Seron (1997) have argued that numbers such as *seven* or *VII* do not activate the same intermediate representations. In the former, the quantity <7> would be activated whereas in the latter, the quantities <5> and <2> in a sum relationship would be activated first. Similarly, they have shown that processing *twelve hundred* and *one thousand two hundred* is different: the former would activate an intermediate representation of the type <12> x <100> whereas the latter would activate a representation of the type <1000> + (<2> x <100>). Interesting differences could also be exploited across languages. For instance, French people say "quatre-vingt-trois" (<4> x <20> + <3>) but the Dutch say "drie en tachtig" (<3> + <80>). Do these different lexico-syntactic structures provoke differences when comparing French and Dutch speakers and how do these differences appear in French-Dutch bilinguals?

Such studies are not only interesting in themselves, but might also play a role in solving the single- versus multiple-code controversy in mental arithmetic. We hope that this review will contribute to the development of research in this new and exciting area.

References

Ashcraft, M.H. (1992). Cognitive arithmetic: A review of data and theory. *Cognition, 44*, 75-106.

Ashcraft, M.H., Donley, R.D., Halas, M.A., & Vakali, M. (1992). Working memory, automaticity, and problem difficulty. In J.I.D. Campbell (Ed.), *The nature and origins of mathematical skills* (pp. 301-329). Amsterdam: Elsevier.

Baddeley, A. D. (1990). *Human memory*. Hillsdale, NJ: Erlbaum.

Brysbaert, M. (1995). Arabic number reading: On the nature of the numerical scale and the origin of phonological recoding. *Journal of Experimental Psychology: General, 124*, 434-452.

Byalystok, E., & Codd, J. (1997). Cardinal limits: Evidence from language awareness and bilingualism for developing concepts of number. *Cognitive Development, 12*, 85-106.

Campbell, J.I.D. (1994). Architectures for numerical cognition. *Cognition, 53*, 1-44.

Campbell, J.I.D., & Clark, J.M. (1988). An encoding-complex view of cognitive number processing: Comment on McCloskey, Sokol and Goodman (1986). *Journal of Experimental Psychology: General, 117*, 204-214.

Campbell, J.I.D., & Kanz, C.L. (1999). *Number processing in Chinese-English bilinguals*. Manuscript submitted for publication.

Chincotta, D., & Hoosain, R. (1995). Reading rate, articulatory suppression and bilingual digit span. *European Journal of Cognitive Psychology, 7*, 201-211.

Chincotta, D., & Underwood, G. (1997). Digit span and articulatory suppression: A cross-linguistic comparison. *European Journal of Cognitive Psychology, 9*, 89-96.

Clark, J.M., & Campbell, J.I.D. (1991). Integrated versus modular theories of number skills and acalculia. *Brain and Cognition, 17*, 204-239.

Cornet, J.A., Seron, X., Deloche, G., & Lories, G. (1988). Cognitive models of simple mental arithmetic: A critical review. *Cahiers de Psychologie Cognitive, 8*, 551-571.

da Casta Pinto, A. (1991). Reading rates and digit span in bilinguals: The superiority of mother tongue. *International Journal of Psychology, 26*, 471-483.

Dehaene, S. (1992). Varieties of numerical abilities. *Cognition, 44*, 1-42.

Dehaene, S., & Cohen, L. (1995). Towards an anatomical and functional model of number processing. *Mathematical Cognition, 1*, 83-120.

Elliott, J.M. (1991). *Language characteristics and arithmetic in bilingual children in Singapore*. Poster presented at the Lancaster Conference.

Elliott, J.M. (1992). Forward digit span and articulation speed for Malay, English, and two Chinese dialects. *Perceptual and Motor Skills, 74*, 291-295.

Ellis, N. (1992). Linguistic relativity revised: The bilingual word-length effect in working-memory during counting, remembering numbers, and mental calculations. In R.J. Harris (Ed.), *Cognitive processing in bilinguals, advances in psychology* (pp. 137-155). Amsterdam: Elsevier.

Ellis, N.C., & Hennelly, R.A. (1980). A bilingual word-length effect: Implications for intelligence testing and the relative ease of mental calculation in Welsh and English. *British Journal of Psychology, 71*, 43-51.

Frenck-Mestre, C., & Vaid, J. (1993). Activation of number facts in bilinguals. *Memory and Cognition, 21*, 809-818.

Geary, D.C., Cormier, P., Goggin, J.P., Estrada, P., & Lunn, M.C.E. (1993). Mental arithmetic: A componential analysis of speed-of-processing across monolingual, weak bilingual, and strong bilingual adults. *International Journal of Psychology, 28*, 185-201.

Glanzer, M., & Duarte, A. (1971). Repetition between and within languages in free recall. *Journal of Verbal Learning Verbal Behavior, 10*, 625-630.

Healy, A. F., & Nairne, J. S. (1985). Short-term memory processes in counting. *Cognitive Psychology, 17*, 417-444.

Hoosain, R. (1979). Forward and backward digit span in the languages of the bilingual. *The Journal of Genetic Psychology, 135*, 263-268.

Hoosain, R. (1982). Correlation between pronunciation speed and digit span size. *Perceptual and Motor Skills, 55*, 1128.

Hoosain, R. (1984). Lateralization of bilingual digit span functions. *Perceptual and Motor Skills, 58*, 21-22.

Hoosain, R., & Salili, F. (1987). Language differences in pronunciation speed for numbers, digit span, and mathematical ability. *Psychologia, 30*, 34-38.

Kolers, P.A. (1968). Bilingualism and information processing. *Scientific American, 218*, 78-86.

LeFevre, J., Bisanz, J., Daley, K.E., Buffone, L., Greenham, S.L., & Sadesky, G.S. (1996a). Multiple routes to solution of single-digit multiplication problems. *Journal of Experimental Psychology: General, 125*, 284-306.

LeFevre, J., Sadesky, G.S., & Bisanz, J. (1996b). Selection of procedures in mental addition: Reassessing the problem-size effect in adults. *Journal of Experimental Psychology: Learning, Memory and Cognition, 22*, 216-230.

Lemaire, P., Abdi, H., & Fayol M. (1996). The role of working memory resources in simple cognitive arithmetic. *European Journal of Cognitive Psychology, 8*, 73-103.

Macnamara, J., & Kushnir, S.L. (1971). Linguistic independence in bilinguals: The input switch. *Journal of Verbal Learning and Verbal Behavior, 10*, 480-487.

Mägiste, E. (1980). Arithmetical calculations in monolinguals and bilinguals. *Psychological Research, 42*, 363-373.

Mägiste, E. (1992). Secondly language learning in elementary and high school. *European Journal of Cognitive Psychology, 4*, 355-365.

Marsh, L.G., & Maki, R.H. (1976). Efficiency of arithmetical operations in bilinguals as a function of language. *Memory and Cognition, 4*, 459-464.

McClain, L., & Shih Huang, J.Y. (1982). Speed of simple arithmetic in bilinguals. *Memory and Cognition, 10*, 591-596.

McCloskey, M. (1992). Cognitive mechanisms in numerical processing: Evidence from acquired dyscalculia. *Cognition, 44*, 107-157.

McCloskey, M., Caramazza, A., & Basili, A. (1985). Cognitive mechanisms in number processing and calculation: Evidence from dyscalculia. *Brain and Cognition, 4*, 171-196.

McCloskey, M., Aliminosa, D., & Sokol, S.M. (1991). Facts, rule, and procedures in normal calculation: Evidence from multiple single-patient studies of impaired arithmetical-fact retrieval. *Brain and Cognition, 17*, 154-203.

Miller, K.F., & Zhu, J. (1991). The trouble with teens: Accessing the structure of number names. *Journal of Memory and Language, 30*, 48-68.

Noël, M-P, & Seron, X. (1997). On the existence of intermediate representations in numerical processing. *Journal of Experimental Psychology, Learning, Memory and Cognition, 23*, 697-720.

Shannon, B. (1984). The polyglot mismatch and the monolingual tie. Observations regarding the meaning of numbers and the meaning of words. *New ideas in Psychology, 2*, 75-79.

Tulving, E., & Colotla, V.A. (1970). Free recall of trilingual lists. *Cognitive Psychology, 1*, 86-98.

Received May, 1998