http://dx.doi.org/10.7468/jksmed.2014.18.2.149
J. Korean Soc. Math. Educ., Ser. D, Res. Math. Educ.
Vol. 18, No. 2, June 2014, 149–170

When 5004 is Said "Five Thousand Zero Hundred Remainder Four": The Influence of Language on Natural Number Transcoding: Cross-National Comparison[1]

NGUYEN, Hien Thi-Thu[*]
Faculty of Psychology and Educational Sciences, Catholic University of Louvain,
1348 Louvain-la-Neuve, Belgium; Email: thi.thu.nguyen@uclouvain.be

GREGOIRE, Jacques
Faculty of Psychology and Educational Sciences, Catholic University of Louvain,
1348 Louvain-la-Neuve, Belgium; Email: jacques.gregoire@uclouvain.be

(Received October 10, 2013; Revised June 27, 2014; Accepted June 28, 2014)

The Vietnamese language has a specific property related to the zero in the name-number system. This study was conducted to examine the impact of linguistic differences and of the zero's position in a number on a transcoding task (verbal number into Arabic number). Vietnamese children and French-speaking Belgian children, from grades 3 to 6, participated in the study. The success rate and the type of errors they made varied, depending on their grade and language. At Grade 4, Vietnamese children showed performances equivalent to Grade 6 Belgian children. Our results confirmed the support provided by language to the understanding and performances in a transcoding task. Results also showed that a syntactic zero is easier to manipulate than a lexical zero for Vietnamese children. The relative influence of language and the source of errors are discussed.

Keywords: number transcoding, syntactic zero, lexical zero, cross-cultural, Vietnamese
MESC Classification: D10, E40, C50
MSC2010 Classification: 97C50, 97D10

1. INTRODUCTION

[1] This article is an extended version of the paper (Nguyen & Gregoire, 2011) presented at the International Commission Mathematical Instruction (ICMI) Study 21 Conference: Mathematics Education and Language Diversity, 16–20 Sept. 2011, São Paulo, Brazil.

[*] Corresponding author

In early mathematical activities, such as rote counting of numbers in sequence, children encounter their first code for numbers, the spoken verbal number, for example, “one” or “twenty-three”. Later, mainly in primary school, they start to learn other codes. These are the Arabic number (e.g., 23) and the written verbal number (e.g., twenty-three). When they practice mathematics, children not only use numbers in different codes but they progressively have to translate numbers from one code to another. This is called “number transcoding”. Number transcoding is a basic and parallel process in learning arithmetical operations and other more advanced mathematical skills.

When children transcode a written verbal number or a verbal number into Arabic code, they regularly make mistakes. This happens particularly when the tens and hundreds in the number are equal to zero. For example, when children transcode the written verbal number “five thousand twenty-four” into an Arabic number, incorrect responses such as “500024” or “5240” are observed. Despite their understanding of place-value, children in primary school often make these errors. This study aimed to examine the effect of the zero’s position on a task of transcoding verbal numbers into Arabic numbers.

Additionally, through several studies concerning the influence of language in the conception and understanding of mathematics, researchers found that different languages can lead to facility or difficulty when performing certain tasks. The Vietnamese language has a special case of denomination for numbers when the digit in the tens or hundreds position is equal to zero. For instance, 5004 is said “five thousand zero hundred remainder four”. The purpose of the present study is also to examine the influence of numerical language characteristics on the aforementioned transcoding task.

1.1. The process of number transcoding and the two types of zero

The transcoding process implies the presence of a code input and of an expected code. We consider three different forms of codes for numbers in this study: verbal number, written verbal number and Arabic number. This study specifically focuses on the transcoding of a written verbal number into Arabic code (e.g., “two hundred forty-five” → “245”). We considered two models for number transcoding. The cognitive model of McCloskey, Caramazza & Basili (1985) and an algorithm of number writing to dictation from Power & Dal Martello (1990). For both models, the transcoding process contains two separate stages: the number comprehension stage and the number production stage. In the first stage, a semantic representation is built on the input code (written verbal number). This semantic representation is then converted into the expected code (Arabic number) by the application of lexical and syntactic rules in the production stage. In McCloskey’s model, the semantic representation is a base-ten representation, indicating the quantities of the number (e.g., four hundred twenty five → $[4].10^2,[2].10^1,[5].10^0$). In

Power & Dal Martello's model, it is considered as a decomposition of the number (e.g., four hundred twenty five is (4×100)+(2×10)+5). The semantic representation is an interpretation of a verbal number. It plays an important and intermediate role in moving from written verbal code to Arabic code. Recall that sum and product relationships are used to express the semantic representation of the number and to transcode a verbal number into Arabic code. A concatenation operation (noted by &) is applied for the product relationship, for example "four hundred" gives 4x 100=<4> & 00 = 400 and the overwriting operation (noted by #) is used for the sum relationship, for example "fifty two" gives 50+2=<50>#<2>=52.

According to Deloche & Seron (1982), two types of errors, lexical and syntactic, were distinguished. Lexical errors indicate that wrong lexical elements constitute the number while its syntactic structure (size) is preserved (e.g., "four hundred twenty-five" [425] → 415). Syntactic errors result from an incorrect application of a combination rule that modifies the syntactic structure of the number (e.g., "four hundred five" [405]→4005). Previous studies conducted on children in grade 2 and 3 have shown that most errors were syntactic (Power & Dal Martello, 1990, 1997; Seron, Deloche & Noeel, 1991; Seron & Fayol, 1994). Syntactic errors are made during the production stage and come from a confusion between the application of the relation sum or multiplication (e.g., "hundred and five" →(100+5)→ 100&5→1005 instead of (100+5)→ 100#5→105). These errors can be explained by the fact that children do not master combination rules. In this case, they transcode each word by their corresponding Arabic counterpart. This kind of error would increase when they work with large numbers or numbers with a zero. Seron and his colleagues (1991) suggested that these errors come from a wrong generalization of the transcoding rule (e.g., "one hundred and five"→105 is correct but "one hundred and thirteen"→1013 or "one hundred and thirty"→1030 are incorrect).

The zero is considered to be a source of difficulty for children when they have to produce numbers because of its characteristic of having a masked position. The distinction between the two types of zero was mainly emphasized in the transcoding task (Granà, Lochy, Girellid, Seron, & Semenza, 2003). The first type is the lexical zero, as in the numbers "120" or "50145", which is derived from semantics, into the Arabic number. The second type is the syntactic zero, as in the numbers "508" or "7014", which is a production of syntax. The syntactic zero is inserted into the number by application of a rule, to indicate a missing value in a position. Considering their characteristics, the syntactic zero is obviously more difficult than the lexical zero for a transcoding task. This was shown by previous study conducted on children in grade 2 and 3 of primary school (Censabella, 2000). The length of a number and its frequency of production affect the ability to transcode with a syntactic zero.

1.2. Impact of language characteristics on mathematical thinking and performances

The language, which is an element of the cultural context, plays an important role in the processes of acquisition, understanding and performance in mathematics (Kaput, 1991; Steffe, Cobb & von Glasersfeld, 1988). Indeed, numbers and mathematical activities are linked to language. Mathematical activities for children mainly adhere to language and number-name to explain the mathematical representations in a more concrete way.

Through several studies concerning the influence of language in the conception and understanding of mathematics, researchers found that different languages can lead to aptitudes or difficulties when performing certain tasks, presented below. Asian languages (Chinese, Japanese, Korean and Vietnamese) possess a numerical names system in which the verbal number corresponds exactly to its written form. There is no particular case for verbal numbers from eleven to nineteen as in English, for example, 11 is spoken ten-one and 16 as ten-six. Verbal numbers indicating a decade are constituted by the relation of multiplication, for example, 20 as two-ten and 70 as seven-ten. As a consequence, 46 is spoken four-ten-six. Indeed, due to a more efficient numerical names system, Asian children have shown more advanced abilities than Western children in counting (Miller, Smith, Zhu & Zhang, 1995; Miller & Stigler, 1987) and adding mentally (Geary, Bow-Thomas, Fan & Siegler, 1993; Geary et al., 1996). Several studies of Miura and his colleagues (Miura, Kim, Chang & Okamoto, 1988; Miura & Okamoto,1989; Miura, Okamoto, Kim, Chang, Steere & Fayol, 1994; Miura, Okamoto, Kim, Steere & Fayol, 1993) showed that the effect of language also facilitated the cognitive representation of a number, the understanding of the canonical base-10 system (e.g., four ten-blocks and two unit-blocks for the number 42) and the understanding of place-value (e.g., the meaning of the individual digit in a multidigits number) for children from China, Japan and Korea, compared to English, French and Swedish children. The effect of language has also helped Chinese children to surpass their English and American counterparts in embedded-ten cardinal understanding (Ho & Fuson, 1998) and in the acquisition and use of ordinal numbers corresponding to their ordinal names (Miller, Major, Shu & Zhang, 2000).

Research conducted on the task of transcoding have shown that the verbal structure also has an influence on the success rate and on the nature of errors. Seron and Fayol (1994) showed that with different verbal names for 70 and 90, the French make more errors than Belgian French-speaking children when writing verbal numbers into their Arabic code. (In Belgium, 70 and 90 are spoken “seventy” (septante) and “ninety” (nonante) while in France, 70 is “sixty-ten” (soixante-dix) and 90 is “four twenty-ten” (quatre-vingt-dix). The German language has the particular feature of reversing the decade-unit structure in the verbal form (e.g. 58 “fifty-eight” is pronounced “eight and fifty”). With regard to this aspect, Lochy, Delazer and Seron (2003) studied the process of transcoding among young French-speaking Belgian and German-speaking Austrian children. They observed that the error rate was higher during the first grade of primary school in Austria than in Belgium.

1.3. The Vietnamese verbal number

According to previous cross-national comparison studies between East Asian and Western countries, the differences can be explained mainly by the transparent verbal number system used in East Asian languages (the regularity of number words). Like other verbal systems of numbers in East Asia, there is not, for the Vietnamese lexical primitives, a word for the particular numbers “eleven”, “twelve”, “thirteen”, “fourteen”, “fifteen” and “sixteen”. The designation of numbers from eleven to nineteen is done with an addition relationship (e.g., eleven is expressed as “ten one”). The Vietnamese language names the numbers in the tens using a multiplicative relationship, highlighting the structure by the number's position (e.g., “two ten” for 20). This characteristic is not found for tens in French. The absence of special cases for tens in Vietnamese induces a perfect correspondence between verbal numbers and their decompositions. For instance, the number 265 will be decomposed as $265 = 2\times100 + 6 \times 10 + 5$. In Vietnamese, it is expressed as “two hundred six ten five”, but in French it is expressed as “two hundred and sixty five”. However, the word “sixty” in French does not express “6×10” as clearly and straightly as the word sixty in Vietnamese (six-ten). With the advantage of the correspondence between the verbal number and its decomposition, the Vietnamese language system can help children to have a better understanding of the structure of the number and place value.

However, the Vietnamese language has peculiarities when the digit in the tens or hundreds position is a zero. This is not found in other Asian languages like Chinese, Korean or Japanese. Vietnamese uses the word “zero” as a lexical primitive in the number construction. For example, 3024 is said “three thousand zero hundred two ten four”. This syntactic zero is not masked in its verbal form like in other languages. There is, however, an exception when the zero is in the tens place where it is replaced by the word “remainder”. For example, 309 is said “three hundred remainder nine”. We can understand the word “remainder” as the amount remaining of the division of 309 by 100. Therefore, in Vietnamese, 5004 will be spoken “five thousand zero hundred remainder four”.

The semantic representation plays an important and mediate role in the transcoding process from a verbal number into Arabic code. Children have to interpret the verbal representation into a semantic representation and then apply two production rules, sum or product, to produce the corresponding Arabic number. In the Vietnamese language, the semantic representation is nearly identical to the verbal number. Moreover, the particularity of the syntactic zero can erase the differences between syntactic and lexical zeros, which is convenient for Vietnamese children when they are transcoding with a syntactic zero. According to its verbal number characteristics and especially in the case of zero, the Vietnamese language provides a more economic model and better support than other languages for the task of transcoding.

Our first prediction was that Vietnamese children should have higher performances than Belgian children in the task of transcoding from a verbal number into Arabic number (Hypothesis 1). It was also predicted that the syntactic zero would not be a source of difficulty for Vietnamese children in the task of transcoding, because of the clear presentation of this kind of zero in the Vietnamese number name (Hypothesis 2). Concerning the errors, Belgian children should have the tendency to erase the zero in a transcoding task when they manipulate large numbers more often than Vietnamese children, due to the masked position characteristic of the syntactic zero (Hypothesis 3).

The particular property of the syntactic zero in the Vietnamese language provides an important opportunity to extend previous researches about the impact of language on mathematical understanding and competences. Many of the previous studies, examining the effects of the symbolic structure system, were conducted on preschool and young primary children (Grades 1–2). It would be interesting to examine if the use of that system also has an effect on older children (Grades 3–6) in this study.

The current study was aimed at examining the influence of these linguistic differences in relation to the syntactic zero on a verbal-Arabic transcoding task. We examined whether the difference due to the numerical language characteristics between Asian and Western countries is still valid for older children (Grades 3–6) who had already acquired the numbers system at school. The second purpose of this study was to assess the impact of the zero's position (syntactic and lexical zeros) and to determine the effect of a syntactic zero on older children in both countries, Vietnam and Belgium.

2. METHOD

2.1. Participants

Two groups of children participated to this research. In Belgium, we assessed 121 children of Grade 3 (55 girls and 66 boys; mean age 107 months), 109 children of Grade 4 (62 girls and 47 boys; mean age 121 months), 143 children of Grade 5 (68 girls and 75 boys; mean age 129 months), and 127 children of Grade 6 (62 girls and 65 boys; mean age 143 months). In Vietnam, we assessed 708 children of Grade 3 (338 girls and 368 boys; mean age 106 months), 690 children of Grade 4 (305 girls and 381 boys; mean age 117 months) and 699 children of Grade 5 (357 girls and 341 boys; mean age 128 months).

Our sample included 52 classes from 8 different schools in two developed cities in VietNam, HoChiMinh and BuonMaThuot. In Belgium, we had 28 classes from 14 different schools in the French-speaking Community[2]. The number of participants in Vietnam was higher than in Belgium. Also, the size of each school and the numbers of children in each class were different because of different school organizations. Each class in Vietnam had about 40 children while in Belgium, there were around 15 children. Despite these differences, both the Vietnamese and the Belgian samples were broad enough to represent the profile of each population.

Transcoding numbers is a basic ability in mathematics so from second grade in primary school, this competence has been taught in both countries. In third grade, children learn to transcode with numbers up to 5 digits.

2.2. Task and stimuli

A paper-and-pencil transcoding task was constructed with eighteen items. The task was for participants to write the natural numbers from a verbal form into Arabic code (“write the number in spoken form into a number with digits”, we gave an example “fourteen →14” in order for children to well understand the task). The test included 18 items with numbers from 3 to 6 digits. There were two categories labelled Syntactic Zero and Non Syntactic Zero. The Syntactic Zero category was used for numbers containing a syntactic zero (e.g. 308 and 35007). Numbers without this type of zero, such as 745 and 700126, were used in the Non Syntactic Zero category. Eighteen items were randomly organized in a list and presented in the same order to all the participants. The test was performed collectively in the classroom during the last month of school year. Children worked individually, without time pressure. The average time to achieve the test was between five and ten minutes.

3. RESULTS

3.1. Overall analysis

3.1.1. Age Comparison of both samples

---

[2] There are three official languages taught in Belgium: Dutch, German and French. Consequently, Belgium shelters a French speaking community, a Dutch speaking community, a German speaking community and a mixed French/Dutch speaking community. To avoid potential bias, we chose our samples exclusively in areas where French is the native language of people and the official language used in school. Children of our sample were learning Dutch and English as second languages, just like most French children learn English, Spanish or German in French schools.

Even though the primary school lasts 5 years in Vietnam when it lasts 6 in Belgium, the age of the children was almost identical. Age of both samples for Grade 3 and 5 did not differ significantly $F(1,762) = 0.041$, $p = 0.840 > 0.05$ and $F(1,802) = 1.770$, $p = 0.184 > 0.05$. Although, Belgian children for Grade 4 were older than Grade 4 Vietnamese children $F(1,749) = 58.407$, $p < .001$.

3.1.2. Gender comparison

For each participant, the mean rate of correct responses for all the items was calculated.

Vietnamese boys' and girls' performances did not differ significantly, $F(1,2092) < 1$, ns. In Belgium, the performances of the boys were better than the girls', $F(1,499) = 6.98$, $p < 0.05$. The mean of correct response in percent of Belgium boys and girls were 74 and 67 respectively.

3.2. The impact of language

3.2.1. Vietnamese Grades 3, 4 and 5 with Belgians Grades 3, 4 and 5

The next table presents the means and standard deviations of scores in percent, according to the grade and the country.

Table 1. Means rates (in percent) and standard deviations of scores according to the grade and the country

Grade	Vietnam		Belgium	
	M	SD	M	SD
3	76.5	22.4	35.2	17.0
4	91.2	13.1	65.9	26.9
5	94.1	9.1	85.8	17.0
6			91.3	13.4

We compared the Vietnamese and Belgian samples according to grade (3, 4 and 5). The comparison was conducted by a two-way analysis of variance (ANOVA) with the Grade (3–5) and Country (Belgium and Vietnam) as variables between-subject. The difference between the three grades level was significant, $F(2, 2464) = 472.512$, $p < 0.001$, partial $\eta^2 = 0.28$, showing that correct response rates increased with grades. Post hoc t-tests showed that all grades differed significantly from each other ($p < 0.001$ for all comparisons). The Country effect was also significant, $F(1, 2464) = 705$, $p < 0.001$, partial $\eta^2 = 0.2$ indicating that the response rates of Vietnamese children were higher than those of Belgians. Finally, the interaction between the Country and the Grade was significant, $F(2, 2464) = 109.234$, $p < 0.001$, partial $\eta^2 = 0.08$, which means that the advancement across grades differed according to the country. In fact, the difference between grades 3 to 5 was less stressed in Vietnam than in Belgium. As illustrated in Fig.1, there are striking differences between Grade 3 Vietnamese and Belgians children, but this difference is less important in Grade 5. At the same age, Vietnamese children in the last three years of primary school have a higher ability than Belgians to write numbers from a verbal form in the Arabic form.

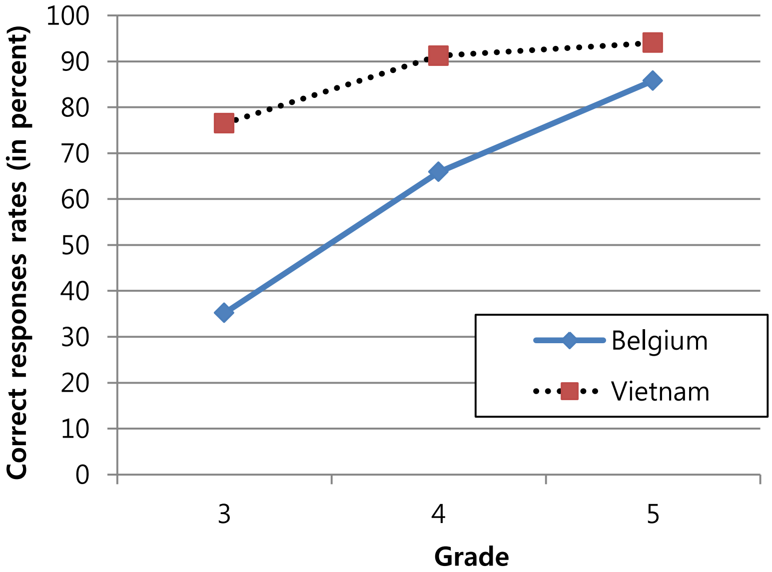

Figure 1. Means of children's scores in percentages according to their Grades and Countries (Grades 3–5 Vietnamese vs. Grades 3–5 Belgian).

3.2.2. Vietnamese Grades 3, 4 and 5 with Belgian Grades 4, 5 and 6

Recall that primary school in Belgium lasts 6 years instead of 5 years in Vietnam. Mathematical program for Vietnamese children in Grade 3 is more advanced than in Belgium concerning large numbers.

Differences in the instruction school between the two countries could explain the significant differences between children of similar age from Grades 3 to 5 as reported above. Consequently, we made a comparison between Vietnamese Grades 3 to 5 and Belgian Grades 4 to 6. Noticed that only in this section of Results, Belgian Grades 3, 4 and 5 used for this analysis (statistic results and in Fig. 2) relate to Belgian Grades 4, 5 and 6 in reality.

There is a one-year lag between Belgian and Vietnamese children, but all effects are still significant according to the two-way analysis of variance with the Grade (3–5) and Country (Belgium and Vietnam) as variables between-subject. The main effect of Grade was significant, $F(2, 2470) = 189.51$, $p < 0.001$, partial $\eta^2 = 0.13$. In addition, performances of Vietnamese children were higher than those of Belgian children. They differed significantly with the effect of Country, $F(2, 2470) = 45.94$, $p < .001$, partial $\eta^2 = 0.02$. The interaction between Grade and Country was also significant, $F(2, 2470) = 5.83$, $p < .05$, partial $\eta^2 = 0.005$.

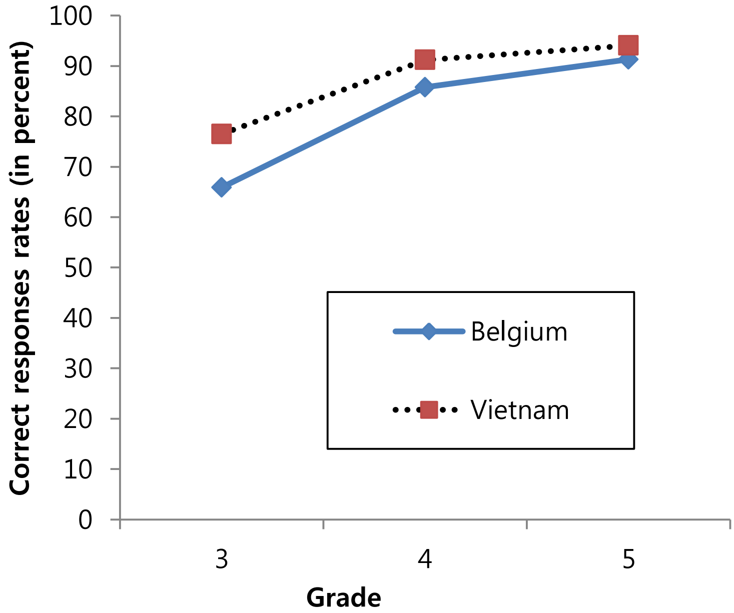

Figure 2. Means of children's scores in percentage according to their Grades and Countries (Grades 3–5 Vietnamese vs. Grades 4–6 Belgian).

The advancement of the Vietnamese children still existed across each Grade even with a one–year lag between Belgian and Vietnamese children, as illustrated in Fig.2. For the last three years of primary school, the transcoding performances among Vietnamese children were better than those of the Belgian children. The scores of Grade 4 Vietnamese children (9 years and 9 months) were equal to the scores of Grade 6 Belgian children (11 years 11 months).

3.3. The impact of syntactic zero

The following table presents the means and standard deviations of scores in percentage depending on the category: Syntactic Zero and Non Syntactic Zero.

Table 2. Means rates (in percent) and standard deviations of scores in percentage depending on the type of zero

Grade	Vietnam		Belgium	
	Syntactic Zero	Non Syntactic Zero	Syntactic Zero	Non Syntactic Zero
3	77.5 (23.9)	74.5 (24.1)	35.0 (18.2)	35.4 (19.8)
4	91.7 (14.2)	90.3 (15.4)	64.9 (28.0)	67.7 (28.0)
5	94.8 (9.3)	92.5 (13.2)	86.4 (17.9)	84.4 (19.1)
6			91.6 (13.6)	90.6 (16.9)

We performed a repeated measure ANOVA on the correct response rates with Category (Non Syntactic Zero and Syntactic Zero) as a within-subject variable and Country and Grade (3–5 for both countries) as between-subject variables. The Category main effect was significant, $F(1, 2464) = 4.56$, $p < .05$, showing the higher success rates for the category Syntactic Zero. The main effect of Grade was also significant, $F(2, 2464) = 468.9$, $p<0.001$. There is a remarkable dissimilarity of the Country main effect as seen in results above. We did find a clear difference between the performances of Vietnamese and Belgians children, $F(1, 2464) = 679.5$, $p < 0.001$. Indeed, the performances of Vietnamese children were higher than those of Belgian children. The interaction between Category and Grade was also significant $F(2, 2464) = 3.7$, $p <0.05$, which means the effect of Syntactic Zero was unlike across Grade. The interaction between Category and Country was significant, $F(1, 2464) = 8.4$, $p <.001$ showing that the effect of the type of zero also differed according to the country. Although, the triple interaction Syntactic Zero × Country × Grade was not significant ($p > 0.05$).

To understand more clearly the interactions described above and to examine the effect of the syntactic zero across grade in each country, two analyses of two–way repeated measures ANOVA were conducted separately, with Category (Non Syntactic Zero and Syntactic Zero) as a within-subject variable and Grade as a between-subject variable. In both countries, Belgium and Vietnam, the main effect of Grade was significant (in all cases $p < 0.001$). In Belgium, the Syntactic Zero effect did not affect the transcoding task. Performances of Belgian children in two categories (Syntactic Zero and Non Syntactic Zero) did not differ significantly, $F(1,496) = .001$, ns. In contrast, this main effect was significant for Vietnamese children, $F(1,2094) = 43.21$, $p < .001$, showing that it was easier for Vietnamese children to write verbal numbers in their Arabic form when they had a syntactic zero.

In both countries, the interaction between Category and Grade did not differ (in all cases $p < 0.05$). The same pattern was observed across grades for each country. Indeed, the syntactic zero was a factor facilitating performances for Vietnamese grades 3, 4 and (p < 0.001 for grade 3 and 5, p < 0.05 for grade 4). Among Belgian children, the syntactic zero wasn't a source of influence for grades 3 and 4 and it was a slight source a difficulty for grades 5 and 6. The Syntactic Zero factor was not significant for every class in Belgium. These results supported Hypothesis 2 (Fig. 3 and Fig. 4).

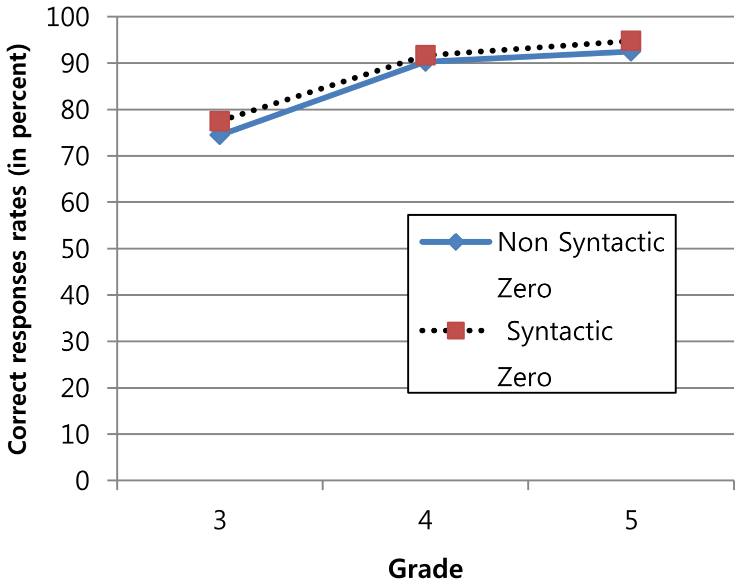

Figure 3. Means of children's scores in percentage by grade in categories Syntactic Zero and Non Syntactic Zero in Vietnam.

3.4. Nature of errors and cluster analyses

To explore errors for each sample of children, Belgian and Vietnamese, cluster analyses were conducted to group children into homogeneous patterns of responses. Three clusters analyses were performed for three samples: Belgian children, Vietnamese children and a group containing participants in both countries. Four types of responses were determined: the correct response and three types of errors: lexical error (response has the same number of digits than the correct response, e.g., "three hundred eight" (308) → 380), omission error (response has less digits than the correct response, e.g., "twenty-five thousand fourteen"(25014) → 2514) and insertion error (response has more digits than the correct response, e.g., "one thousand two hundred fifty three" (1253) → 1000253).

In the first cluster analysis for Belgian children, three groups of children were identified based on the type of response. The clusters were labeled high-omission (n = 100), omission (n = 95) and expert (n = 305). Children in the cluster "high-omission" had a tendency to provide a response with less digits than the correct one. This tendency was valid for all the items with 5 and 6 digits including 1 or 2 zeros (e.g. 607 029). Cluster "omission" included children who had the tendency to erase some digits when they transcoded a verbal number into Arabic code. They showed this tendency only for four items that had 6 digits with 2 digits equal to zero. Participants in the cluster "expert" responded almost correctly to all of the items (86% to 99%).

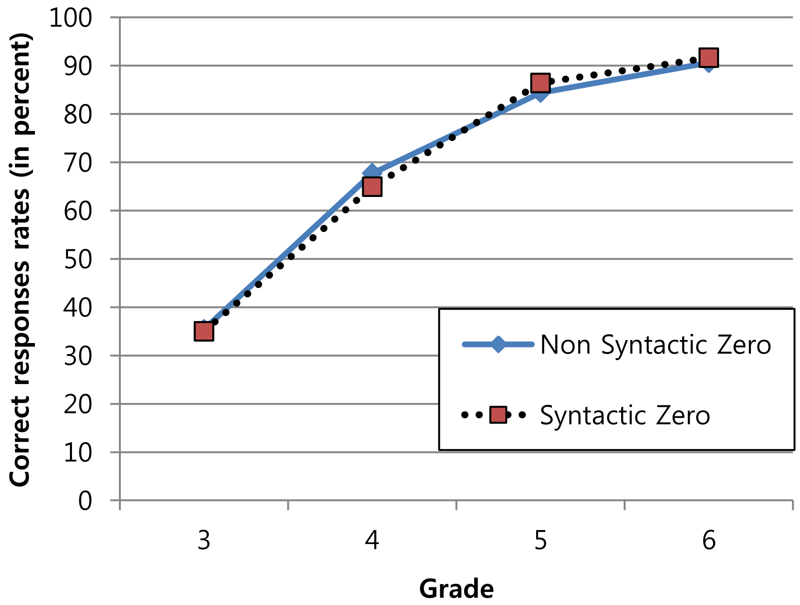

Figure 4. Means of children's scores in percentage by grade in categories Syntactic Zero and Non Syntactic Zero in Belgium.

Table 3. Percent of participants per grade of each cluster Belgian

	Grade 3	Grade 4	Grade 5	Grade 6
High-omission	54	22	6	2
Omission	35	28	10	6
Expert	11	50	84	92

The cluster membership distribution in each grade differed significantly from the expected distribution as shown by a chi-square analysis. In Grade 3, $\chi^2(2, N= 121) = 32.34$, $p < 0.001$, the great majority of children belonged to clusters "high-omission" (54%) and "omission"(35%) which means they had a tendency to erase the zero, mainly the syntactic zero, when they transcoded a verbal number into Arabic code. Few children were into the cluster “expert” (11%). In Grade 4, $\chi^2(2, N= 109) = 13.56$, $p< 0.001$, half of the children belonged to the cluster “expert”, but the rates of children who belonged to clusters “high-omission” and “omission” were still high (22% and 28% respectively). An evolution was found in Grade 5, χ^2 (2, N= 143) =1 65.16, $p<0.001$ and Grade 6, χ^2 (2, N = 127) = 197.73, $p < 0.001$. Indeed, the majority of children belonged to cluster “expert” (84% for Grade 5 and 92% for Grade 6) and few of them were classed into cluster “high-omission”. Noticed that there were children in Grade 5 and Grade 6 who belonged to the cluster “omission” which means they also had difficulties to manipulate a six digits number, with two digits equal to zero. These results supported hypothesis 3, Belgian children have a tendency to erase the syntactic zero in the transcoding task. This tendency is more prevalent among younger children (Grade 3).

The second cluster analysis was run for Vietnamese children. Three clusters were also determined in this sample. The clusters were named “omission” (n = 246), “expert” (n = 978) and “perfect” (n = 876). Children in the cluster “omission” had a tendency to provide a response with less digits than the correct response. This tendency was observed for the five items with 6 digits, 2 of them (digit) equal to zero. Participants in the cluster “expert” responded correctly to almost every item (85% to 98%). Cluster “perfect” included children who gave correct responses to every item (100%).

Table 4. Percent of participants per grade of each cluster Vietnamese

	Grade 3	Grade 4	Grade 5
Perfect	27	45	54
Expert	48	48	43
Omission	25	7	3

The cluster membership distribution in each grade differed significantly from the expected distribution as shown by a chi-square analysis.

In Grade 3, $\chi^2(2, N = 708) = 70.43$, $p < 0.001$, the majority of children belonged to the cluster “expert” (48%), which means they nearly mastered the competence to transcode from a verbal number into Arabic code . A quarter of children in Grade 3 were into the cluster “omission”.

In Grade 4, $\chi^2(2, N= 690) = 226.67$, $p < 0.001$, there was the same rate of children who belonged to the cluster “expert” compared to Grade 3. However there were more children in the cluster “perfect” (45%) and fewer children in the cluster “omission” (only 7%). In Grade 5, $\chi^2(2, N = 699) = 297.43$, $p < 0.001$, almost all children belonged to clusters Expert or Perfect (97%).

According to the nature of the separate clusters identified for Belgians and Vietnamese participants and to the distributions in each cluster for each country, Vietnamese children showed more advanced abilities than Belgian children. To examine directly the advance of Vietnamese over Belgian children referring to their errors, a third cluster analysis was run for all participants of both countries. Despite the difference between Belgian and Vietnamese participants, similar errors were observed. Children used to erase some digits, mainly the syntactic zero digit, when they transcoded a verbal number into Arabic code.

Three clusters of children were also identified, based on the type of response in this case. The clusters were labeled "omission" (n = 350), "expert" (n = 1268) and "perfect" (n=979). Children in the cluster "omission" had a tendency to provide a response with less digits than the correct response. This tendency was observed for items with six digits, two of them equal to zero. Participants in cluster "expert" responded almost correctly to all items (85% to 98% for all items, except 71% for one item). Cluster "perfect" comprised children who gave a correct response to all items (100 %).

The distribution of clusters was influenced significantly by the differences between the two countries and the difference across grades. The chi-square test was the following: χ^2 (2, N = 2597) = 355.63, $p < 0.001$; χ^2 (2, N = 2597) = 256.81, $p<0.001$ respectively for Country x Cluster and Grade x Cluster. It means that the delivering of children to each cluster depends on the country and the grade level.

The evolution of children across grade and the evolution of Vietnamese children over Belgian children are also illustrated in Table 5.

Table 5. Percent of participants per grade of each cluster according to the country

Grade	Country	Clusters		
		Omission	Expert	Perfect
Grade3	Belgium	88	12	0
	Vietnam	17	56	27
Grade4	Belgium	49	39	12
	Vietnam	4	51	45
Grade5	Belgium	16	56	28
	Vietnam	2	45	53
Grade6	Belgium	10	50	40

4. DISCUSSION

The present study was motivated by the particular number name feature present in the Vietnamese language and its relation to the syntactic zero. Previous comparison studies between Asian and Occidental children did not use a transcoding task directly related to language. Our study, using this transcoding task on older children, confirmed that different languages lead to variations in the numerical performances between Asian and non-Asian children. The results regarding the impact of a syntactic zero on a verbal-Arabic transcoding task and the nature of errors committed are discussed. Factors influencing performances such as the mathematical program, the numerical experience of children and the perspectives are also discussed below in more detail.

4.1. Impact of structural characteristics of numerical languages on a transcoding task

In a verbal-Arabic transcoding task, Vietnamese children in grade 4 achieved performances equivalent to older Belgian children in grade 6.

Several previous studies on preschool and younger children showed advanced performances for Asian children compared to children from Western countries. Our study presents results from older children at primary school. The findings are consistent with our first hypothesis; the better performances of Vietnamese children are related to a greater transparency of the Vietnamese language compared to the French language. This could be explained by the following two reasons:

First, in Vietnamese, more words are needed than in French to produce a verbal number, particularly for numbers with a syntactic zero. A good example is the number 208004. In French, five words are used to write it (two hundred eight thousand four), while nine words are needed in Vietnamese (two hundred remainder eight thousand zero hundred remainder four). When transcoding a verbal number into an Arabic number, Vietnamese children have to read and translate the words using a term to term correspondence. The semantic representation is embedded in its expression. Belgian children work with fewer words. It seems easier but they have to apply syntactic rules to produce the corresponding Arabic code. The syntactic structure is generally more difficult than the lexical structure to produce an Arabic number from a verbal number. For a transcoding task, Vietnamese children use the "lexical pattern" and Belgian children use the "syntactic pattern". This can explain results from comparisons between Vietnamese grades 3, 4 and 5 with Belgian grades 4, 5 and 6. Even when the difference in the program between the two countries becomes smaller, the outperformance of Vietnamese children is still highlighted.

Second, according to the Dehaene three codes model (1992), the verbal representation of a number is dependent on the base-10 structure. As the Vietnamese system number name is very transparent with the base-10 structure, it will consequently be a support for Vietnamese children in a verbal-Arabic transcoding task.

Specific errors may emerge from the specific property of a number-name system (Zuber, Pixner, Moeller & Nuerk, 2009). Due to the structure of verbal numbers in French, the French-speaking Belgian children in our study committed a typical error which was not found in the Vietnamese sample. For example, with the number: 542 183 "five hundred forty-two thousand one hundred four-twenty three" (French verbal expression), Belgian children showed many more insertion errors. The percentage of failure was 41% for Belgian children and 9% for Vietnamese. Interestingly, 29% of Belgian children (46% at Grade 3) produced responses ending with "423" (instead of "183") such as 542 100 423, 542 423 or 50 042 423. How could we explain this kind of error? In French, the expression "four twenty" is used to indicate the number 80. In this particular case, it is necessary to combine lexical and syntactic rules to produce the Arabic number from the verbal number. When manipulating large numbers containing a sub-syntactic structure such as "four twenty" (80 in the French language), children have the tendency to provide 423 instead of 83 (four twenty three). When working with large numbers, Belgian children use lexical relations more often than syntactic ones. This error was not observed with Vietnamese children, they also used the lexical relation rather than the syntactic relation or confused the two relations when manipulating large numbers, but their errors were different, for example 50 042 183 or 54 200 183.

4.2. The impact of the syntactic zero and the nature of errors

Our second aim for the current study concerned the influence of the syntactic zero in a transcoding task. By having access to a large sample of children at different grade levels (Grade 3–6), we observed that this influence is less important or non-existent compared to previous results (Censabella, 2000) based on younger samples (Grades 2–3). Indeed, the syntactic zero factor was not a source of influence for the Belgian children from grades 3 to 6. Examining this factor from grade 3 to 6 allowed us to verify the development of children across grades and to correctly understand the location of exertion of the syntactic zero. Unlike previous studies, the syntactic zero was a support for Vietnamese children during a transcoding task. The reasonable explanation for these findings is that the Vietnamese numerical language is transparent for a number with a syntactic zero.

According to the cluster analysis, 89% of Belgian children in grade 3 belonged to clusters "high-omission" and "omission", while 50% in grade 4 belonged to these clusters. It means that the syntactic zero would pose a typical difficulty for children in the transcoding process. Notice that we did not find content referring to the syntactic zero in mathematical curricula and lessons of both countries. In the Vietnamese language, the difference between syntactic and lexical zeros is not explicit because the notion of syntactic zero is represented in spoken numbers. This is, however, not valid for other languages, such as French. It could be necessary to include the notion of syntactic zero in mathematics lessons. Teachers could emphasis the existence of the syntactic zero in verbal numbers in order to help children to recognize and understand its role better, at least in verbal representations and transcoding tasks.

Two meanings can be differentiated regarding the function of the zero for the construction of a number. The first relates to the syntactic rule in the production of an Arabic number. When a power of ten is missing in a number, the insertion of a zero is required in order to conserve the positional system of other digits (e.g., “Four thousand twenty-five” is 4**0**25). The second is the opposite. When a digit in a number is equal to zero, even if it is worth null, this digit always occupies a position in the number. For example, in the number 4025, the value of the hundreds position is null but there is always a digit indicating the hundreds, which is 0. From their tendency to omit the digit equal to zero while transcoding a number with a syntactic zero, we can predict that children will encounter difficulties when they manipulate a zero in other skills, such as place-value understanding (e.g., “Determine the digit indicating hundreds in the number 4025”). We expect that for this task, children will respond that “there is no digit indicating hundreds”. They might confuse the value of the digit 0 with its role of conservation. In the current or in previous studies, the syntactic zero was analyzed only in transcoding from a verbal number into an Arabic number. The syntactic zero properties might lead to difficulties in other types of presentations, such as symbolic or analog, as well as in other advanced mathematical activities. It would be necessary to explore more precisely the role of the zero in these tasks in a future study.

Previous studies suggested that the most frequent errors in a verbal-Arabic transcoding task were syntactic ones. Children had the tendency to produce a response with more digits than the correct response by inserting extra zeros (e.g. “three hundred and sixty-five” (365) was transcoded by 30065 or 3065). This is called an insertion tendency. Noel and Turconi (1999) explained these errors by the fact that the product relation is normally mastered before the sum relation, children applying the product relation intead of sum relation when they produce an Arabic number (e.g., “four hundred two” → (4x100)+2) → 400+2 → 400&2 → 4002 instead of (4x100)+2) → 400+2 → 400#2 → 402). According to Seron and his colleagues (Seron, Deloche & Noel, 1991; Seron & Fayol, 1994), errors of insertion of extra zeros are common in the second and third grade. Children transcoded each word by its Arabic counterpart (e.g., one thousand nineteen was written 100019). They applied the transcoding rules of addition structures of Units for Decades (e.g., “one hundred three” becomes 103 (correct) but “one hundred thirty” becomes 1030(fault)). In the present study, children were older in comparison with previous study samples, but they still transcoded each word by their Arabic counterpart (e.g., “one thousand two hundred fifty-three” 1253 → 1000253; 120053 and 100020053). For numbers with a syntactic zero, they mainly erased the zero not present in the verbal form of the number (e.g., the common response for 25014 is 2514 and 503008 is 5308). The tendency to produce a response with fewer digits than the correct response (omission tendency) in this study is explained by the particular characteristic of the syntactic zero. It suggests that children have not yet master the understanding of place-value and the role of a zero in the production phase of an Arabic number.

In short, errors in manipulation with the syntactic zero made by the children were similar across countries. They showed a tendency to erase the syntactic zero when they transcoded a verbal number into an Arabic number. This type of error was more common in Belgium than in Vietnam. This is explained by the fact that the syntactic zero is highly masked in French numerical language, while it is specifically named in the Vietnamese language.

4.3. Perspectives and future studies

In previous cross-national studies and mathematical achievement studies, besides language factor, other factors like parents' behavior, initial arithmetical knowledge, learning context and mathematical methodology could account for the advancement of Asian children over Western children (Ho & Fuson, 1998; Miller, Smith, Zhu, & Zhang, 1995; Saxton & Towse, 1998; Towse & Saxton, 1997). In the present study, the difference between mathematical program and arithmetical experience between Belgium and Vietnam can explain part of the progress of Vietnamese children. In fact, in Belgium, transcoding with numbers up to 5 digits is taught in grades 3 and 4 and with larger numbers in grades 5 and 6. In Vietnam, children have already learnt to manipulate numbers with 6 digits at grade 3. Numbers with 7 digits are introduced at grade 4. In addition, Vietnamese children also have a frequent use of large numbers (starting with 5 digits) in everyday life, in relation to the monetary system whose unit is a hundred.

The task instructions also accounted for the variations in performances in cross-national studies, as reported by Saxton and Towse (1998). We observed that performances in a transcoding task for children at grades 3 and 4 with numbers of 3 and 4 digits seemed less advanced in comparison with the results of previous studies (Seron & Fayol, 1994; Censabella, 2000). Indeed, the test included a large number of items with 5 and 6 digits, including one or two syntactic zeros. So the difficult items may have influenced global performances, mainly for younger Belgian children (grades 3 and 4) who do not have a lot of experience with large numbers. The effect of task instructions was considered as a factor affecting outperformances by Vietnamese children. In order to avoid confounding factors and to focus on the language impact, it would be more useful to select appropriate items for the task instructions.

The influence of language was demonstrated mainly on abilities related to the canonical base-10 system and on linguistic tasks such as counting or transcoding. According to the three codes model of Dehaene (1992), each representation of a number (verbal, Arabic and Analog) is associated with specific manipulations. The Analog (magnitude) representation is independent of the base-10 structure. It could be interesting to examine whether the transparent property of the syntactic zero also emerges in place-value understanding (e.g., identify a digit in a multidigits number) and in non-linguistic tasks, such as Analog-Arabic transcoding. Children would be asked to write down an Arabic number from its analog-representation. The number can be represented with cubes, for example, a big cube represents a hundred and a small cube represents ten. Children would have to give the Arabic number (240) from a presentation with 2 big cubes and 4 small cubes. Therefore, children in both countries would have the same input code. Vietnamese children could no longer rely on the support of the transparent denomination of the syntactic zero. Using non-linguistic tasks is also an efficient way to evaluate the impact of language (Casasanto, Fotakopoulou, Pita & Boroditsky). A future study, not only with verbal-Arabic transcoding but also with Analog representation-Arabic transcoding and place-value understanding is currently being conducted. It will allow a precise understanding of the effect of language on the interpretation and performance of children in a more complete way.

REFERENCES

Casasanto, D.; Fotakopoulu, O.; Pita, R. & Boroditsky, L. (In Revision). How deep are effects of language on thought? Time estimation in speakers of English and Greek. *Cognition*.

Censabella, S. (2000). *L'écriture des nombres sous dictée chez l'enfant. Etude qualitative et implication de la mémoire de travail*. Unpublished master thesis. Liège, Belgique: Université de Liège.

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

Deloche, G., & Seron, X. (1982). From one to 1: An analysis of a transcoding process by means of neuropsychological data. *Cognition* **12**, 119–149.

Geary, D. C.; Bow-Thomas, C.C.; Fan, L. & Siegler, R.S. (1993). Even before formal instruction, Chinese children outperform American children in mental addition. *Cognitive Development* **8(4)**, 517–529.

_____ (1996). Development of arithmetical competencies in Chinese and American children: Influence of age, language, and schooling. *Child Development* **67(5)**, 2022–2044.

Granà, A.; Lochy, A.; Girellid, L.; Seron, X. & Semenza, C. (2003). Transcoding zeros within complex numerals. *Neuropsychologia* **41**, 1611–1718.

Ho, C. S.-H., & Fuson, K. C. (1998). Children's knowledge of teen quantities as tens and ones: Comparisons of Chinese, British, and American kindergartners. *Journal of Educational Psychology* **90(3)**, 536–544.

Lochy, A.; Delazer, M. & Seron, X. (2003). *Influence of language in the early acquisition of num-*

bers writing: A comparison of French and German. Paper presented at the Third Aachen–Gent Brain and Number Workshop, Aachen, Germany.

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

Miller, K.; Major, S. M.; Shu, H. & Zhang, H. (2000). Ordinal knowledge: Number names and number concepts in Chinese and English. *Canadian Journal of Experimental Psychology* **54(2)**, 129–140.

Miller, K. F.; Smith, C. M.; Zhu, J. &Zhang, H. (1995). Preschool origins of cross-national differences in mathematical competence: The role of number-naming systems. *Psychological Science* 6, 56–60.

Miller, K. F. & Stigler, J. W. (1987). Counting in Chinese: Cultural variation in a basic cognitive skill. *Cognitive Development* **2(3)**, 279–305.

Miura, I. T.; Kim, C. C.; Chang, C. M. & Okamoto, Y. (1988). Effects of language characteristics on children's cognitive representation of number: Cross-national comparisons. *Child Development* **59**, 1445–1450.

Miura, I. T. & Okamoto, Y. (1989). Comparisons of American and Japanese first graders' cognitive representation of number and understanding of place value. *Journal of Educational Psychology* **81**, 109–113.

Miura, I. T.; Okamoto, Y.; Kim, C. C.; Steere, M. & Fayol, M. (1993). First Graders' Cognitive Representation of Number and Understanding of Place Value: Cross-National Comparisons—France, Japan, Korea, Sweden, and the United States. *Journal of Educational Psychology* **85(1)**, 24–30.

Miura, I. T.; Okamoto, Y.; Kim, C. C.; Chang, C. M.; Steere, M. & Fayol, M. (1994). Comparisons of children's cognitive representation of number: China, France, Japan, Korea, Sweden and the United States. *International Journal of Behavioral Development* **17**, 401–411.

Noël, M-P. & Turconi. E. (1999). Assessing number transcoding in Children. European Review of Applied Psychology **49(4)**, 295–302.

Power, R. J. D. & Dal Martello, M. F. (1990). The dictation of Italian numerals. *Language and Cognitive Processes* **5**, 237–254.

Power, R. J. D. & Dal Martello, M. F. (1997). From 834 to eighty thirty four: The reading of Arabic numerals by 7-year-old children. *Mathematical Cognition* **3(1)**, 63–85.

Saxton, M. & Towse, J.W. (1998). Linguistic Relativity: The case of place value in muiti-digit numbers. *Journal of Experimental Child Psychology* **69**, 66–79.

Seron, X.; Deloche, G. & Noël, M-P. (1991). Number transcoding by children: Writing Arabic numbers under dictation. In: J. Bideaud, C. Melijac & J. P. Fisher (Eds.), *Pathway to Number* (pp. 245–264). Hillsdale New Jersey: Lawrence Erlbaum Associates.

Seron, X. & Fayol, M. (1994). Number transcoding in children: A functional analysis. *British Journal of Developmental Psychology* **12**, 281–300.

Steffe, L. P.; Cobb, P. & von Glasersfeld, E. (1988). *Construction of arithmetical meanings and strategies*. New York: Springer-Verlag.

Towse, J. W. & Saxton, M.(1997).Linguistic influences on children's number concepts: Methodological and theoretical considerations. *Journal of Experimental Child Psychology* **66**, 362–374.

Zuber, J.; Pixner, S.; Moeller, K. & Nuerk, H.-C. (2009). On the language specificity of basic number processing: Transcoding in a language with inversion and its relation to working memory capacity. *Journal of Experimental Child Psychology* **102**, 60–77.