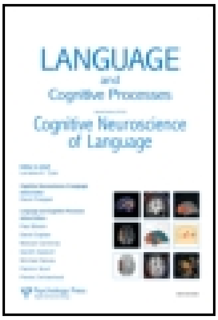

Language and Cognitive Processes
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/plcp20

The dictation of Italian numerals

R. J. D. Power [a] & M. F. Dal Martello [a]
[a] Dipartimento di Psicologia, Facoltà di Magistero, Padova, Italy
Version of record first published: 13 Dec 2007.

To cite this article: R. J. D. Power & M. F. Dal Martello (1990): The dictation of Italian numerals, Language and Cognitive Processes, 5:3, 237-254

To link to this article: http://dx.doi.org/10.1080/01690969008402106

LANGUAGE AND COGNITIVE PROCESSES, 1990, *5* (3) 237-254

The Dictation of Italian Numerals

R. J. D. Power and M. F. Dal Martello
Dipartimento di Psicologia, Facoltà di Magistero, Padova, Italy

Some recent studies on acalculia raise the question of how people translate verbal numerals (e.g. three hundred and sixty five) into Arabic numerals (365). This ability can conveniently be tested by a dictation task: The experimenter presents a spoken verbal numeral which the subject must write down in Arabic digits. Two broad categories of dictation error have been observed in acalculia studies. First, the patient may select a wrong digit ("lexical" error, e.g. 364); secondly, he may produce a numeral with the wrong overall structure, in particular by inserting the wrong number of zeros ("syntactic" error, e.g. 30065). The present study examines another source of evidence, the performance of children who are in the process of acquiring this ability (the crucial age is around 6–8 years). A total of 15 Italian children were given a graded dictation test for numbers below 1 million. In comparison with acalculia patients, the most striking result was the overwhelming preponderance of syntactic errors—of 128 errors scored, 111 were classified as syntactic and only 3 as lexical. The most common syntactic error was the insertion of extra zeros (e.g. 30065 or 3065 instead of 365). A formal theory explaining such errors is proposed. According to the theory, the production of an Arabic numeral like 365 requires the combination of 300 and 65 by a string operation which we call "over-writing"; children who have not yet learned this operation tend to fall back on concatenation.

INTRODUCTION

People in our culture are familiar with two notations for naming the natural numbers—verbal numerals, such as three hundred and sixty five, and Arabic numerals, such as 365. Verbal numerals are used mainly in conversation; Arabic numerals are preferred for performing calculations and for writing down large numbers. By the age of 10, most children have mastered both the Arabic system and a verbal system, and can translate between them.

Investigations of numeral transcoding usually employ two tasks, referred to as "reading" and "dictation". In a reading test, the experimenter

Requests for reprints should be addressed to R. J. D. Power, Dipartimento di Psicologia, Facoltà di Magistero, Piazza Capitaniato 3, 35100 Padova, Italy.

presents a written Arabic numeral (e.g. 24) and the subject responds by producing a spoken verbal numeral ("twenty four"). Conversely, in a dictation test, the experimenter presents a spoken verbal numeral which the subject tries to write down in Arabic digits.

The present article investigates the performance of 7-year-old Italian children on the dictation task. The subjects were at the beginning of their second year of elementary school, at which time most children start to form ideas about the structure of complex numerals, and hence to produce revealing errors. We hoped that such errors would throw light on two issues: first, the nature of the adult transcoding mechanisms and, secondly, the process by which they are acquired.

So far as we are aware, no study of this kind has previously been carried out on children. The only data available on numeral transcoding come from neuropsychological studies of patients suffering from acalculia (see especially Deloche & Seron, 1982; McCloskey, Caramazza, & Basili, 1985). In these studies, two main categories of dictation error have been observed: first, "lexical" errors and, secondly, "syntactic" errors. The following are some examples of each category:

Lexical errors (Deloche & Seron, 1982, French)

Stimulus	*Response*
cinquante huit	59
quatre cent cinquante neuf	469
neuf cent quarante trois	843

Syntactic errors (Singer & Low, 1933, American)

Stimulus	*Response*
two hundred forty two	20042
two thousand five hundred	2000500

As can be seen, with lexical errors the overall structure of the response is correct, but the subject slips in selecting one (or more) of the digits. With syntactic errors the digits are correctly selected but wrongly arranged (e.g. too many zeros are inserted).

Some idea of the relative frequencies of these errors can be obtained from the data on the 32 patients studied by Deloche and Seron (1982), each of which had suffered brain damage impairing his or her ability to read numerals and to write them down under dictation. A corpus of 887 dictation errors was collected, of which 242 (roughly 25%) were classified as lexical, and 445 (roughly 50%) as syntactic. A subsidiary aim of the present study is to find out whether the errors made by young children are similarly distributed.

METHOD

Materials

The stimuli were Italian numerals spoken by the experimenter. These were presented in groups, beginning with single-digit numbers (uno, due, tre, etc.) and progressing in gradual stages to numbers just below 1 million. The groups were defined according to the length and composition of the correct response: group 1 covered the numbers 1–9, group 2 the numbers 11–99, and so forth. All of the groups except group 1 were split into three subgroups labelled a, b, and c. The criteria for this subdivision, and the detailed contents of each subgroup, are shown below:

1. Single-digit numerals:
   1 2 3 4 5 6 7 8 9

2a. Two-digit numerals in the range 10–19:
   10 11 12 13 14 15 16 17 18 19

2b. Two-digit numerals in the sequence 20, 30, . . ., 90:
   20 30 40 50 60 70 80 90

2c. Other two-digit numerals:
   26 34 42 59 67 75 83 91

3a. Three-digit numerals ending in 00:
   100 200 300 400 500 600 700 800 900

3b. Other three-digit numerals without internal zero:
   125 281 312 488 530 674 753 820 965

3c. Three-digit numerals with internal zero:
   106 205 303 402 509 604 708 806 908

4a. Four-digit numerals ending in 000:
   1000 2000 3000 4000 5000 6000 7000 8000 9000

4b. Other four-digit numerals without internal zero:
   1700 2610 3194 4388 5777 6235 7426 8888 9351

4c. Four-digit numerals with internal zero:
   1002 2074 3508 4006 5409 6013 7007 8305 9072

Test materials were also prepared for five- and six-digit numbers; however, almost all the subjects found these too difficult.

A zero is regarded as "internal" if it has at least one non-zero digit both in the string on its left and the string on its right. In the numeral 10200300, for example, the first three zeros are internal and the final two zeros are not. Numerals containing internal zeros were assigned to a separate category because they were expected to pose special problems.

Subjects

The sample was made up of nine boys and six girls from the second-year class of the Manin elementary school in Padua. The average age of the children was 7 years 3 months.

Procedure

The experiment was carried out during school hours. After having first been introduced to the class, the experimenter took the children one at a time to an adjoining room where the test was administered. Each subject was seated at a table and given a pen and paper. The experimenter then read out a series of Italian numerals which the subject was asked to write down in digits. The children were not told whether their answers were right or wrong; general encouragement was given, but no corrections were made.

The test items were presented group by group, starting with 1 and ending with 6c (if the child managed to get that far). The experimenter did not present every item in every group. If the child answered the first three or four items in a group correctly, the experimenter proceeded to the next group: the aim was to discover the zone where the child began to make errors, and to explore that zone thoroughly. When the experimenter judged that the problems were so far beyond the child's ability that the errors were no longer enlightening, the test was halted. Some of the children wanted to stop before this point; others became enraptured with the larger numbers and insisted on continuing to the very end. In either case, the subject's wishes were respected.

RESULTS

Overall Performance

The overall level of performance is shown in Table 1. The names assigned to the subjects are pseudonyms of the appropriate gender; boys' names end in -o (e.g. Alberto) and girls' names in -a (e.g. Clara). For convenience, the names have been selected in such a way that alphabetical order corresponds roughly to ability on the task, i.e. Alberto performed best and Vittorio worst. Table 1 reveals the following:

1. The consistency of responses within groups was remarkably high. Out of a possible 90 (15 × 6) groups in the critical range from 3a to 4c, the subjects attempted 80. For 35 of these, all responses were correct; for

TABLE 1
Overall Performance

Name	*1*	*2a*	*2b*	*2c*	*3a*	*3b*	*3c*	*4a*	*4b*	*4c*	*5a*	*5b*	*5c*	*6a*	*6b*	*6c*
Alberto	A	A	A	A	A	A	A	A	A	A	A	A	S	A	S	N
Bruno	A	A	A	A	A	A	A	A	N	S	A	S	S	S	N	N
Clara	A	A	A	A	A	A	A	A	N	S	S	N	S	A	N	
Daria	A	A	A	A	A	S	A	A	N	N	A	N	N	N	N	N
Emilio	A	A	A	A	A	N	A	A	N	N	N	N	N	N	N	N
Franco	A	A	A	A	A	N	N	A	N	N						
Guido	A	A	A	A	A	N	N	A	N	N						
Lisa	A	A	A	A	A	N	N	A	N							
Marco	A	A	A	A	A	S	S	N	N	N	N	N	N			
Nino	A	A	A	A	A	N	N	N	N	N	N	N	N	N	N	N
Paola	A	A	A	A	A	N	N	A*	N							
Rita	A	A	A	A	A	N										
Silvic	A	A	A	A	S	N	S	N	N							
Teresa	A	A	A	A	N	N	S*	N	N	N	N	N				
Vittorio	A	A	A	A	A*		N									

Key: A, All correct; N, none correct; S, some correct; *, help given; empty slot, not attempted.

37, all were mistaken; for 7, some were correct and some mistaken. Thus responses were consistent (i.e. all correct or all incorrect) for 72 (90%) of the groups attempted. If errors had been evenly distributed across these groups, the percentage of groups with consistent responses would have been 25%, an enormous difference. To confirm this result statistically, expected and observed frequencies were calculated for the eight possible permutations of right (R) and wrong (W) responses to the first three items in each group: RRR (all correct), RRW, RWR, WRR, RWW, WRW, WWR, and WWW (all mistaken). The resulting distributions were compared by the chi-squared test. As one would expect, the difference between the expected and obtained frequencies was highly significant ($\chi^2 = 190.2$, d.f. $= 7$, $P < 0.00001$). For this test of within-group consistency, we thought it best to omit the easier groups (1–2c), in which there were no errors, and the harder groups (5a–6c), which many subjects were unable even to attempt. However, it is worth remarking that the overall trend was virtually the same: out of 179 groups attempted, responses were consistent in 162 (90.5%).

2. All numerals below 100 were correctly dictated: not a single slip occurred. This result confirms our impression that the subjects were taking the experiment seriously and trying to perform well; it also shows that they had mastered the dictation of numerals in this range.

3. Few of the subjects were able to transcribe five- or six-digit numerals. Most of them began to make errors in groups 3b, 3c, 4b, and 4c; consequently, most of the analyses that follow concentrate on these groups.

Lexical and Syntactic Errors

Tables 2 and 3 show the frequencies of lexical and syntactic errors for the first three responses by each subject to the stimuli in groups 3a, 3b, 3c, 4a, 4b, and 4c. There are thus at most $3 \times 6 = 18$ responses for each subject—fewer than 18 if he or she gave up during this phase of the test. The responses were placed in four categories:

+L +S	Lexically and syntactically correct
+L −S	Lexically correct, syntactically incorrect
−L +S	Lexically incorrect, syntactically correct
−L −S	Lexically and syntactically incorrect

A response was regarded as lexically correct if it contained all the non-zero digits of the correct answer. It was regarded as syntactically correct if the digits were correctly ordered and if zeros were added in the correct positions. Here is an example:

Stimulus:	cento venti cinque	
Correct response:	125	
Guido's response:	10025	(+L −S)
Teresa's response:	195	(−L +S)
Silvio's response:	1505	(−L −S)

Guido's response was classified as lexically correct (it contains the digits 1, 2, and 5) but syntactically incorrect (it has two extra zeros). Teresa's response was classified as syntactically correct (the digits are correctly arranged) but lexically incorrect (9 is used instead of 2). Silvio's response was classified as both lexically incorrect (5 instead of 2) and syntactically incorrect (one extra zero). Responses like 521 or 251, in which the digits are misordered without the insertion of extra zeros, would count as syntactic errors; however, no responses of this kind were obtained.

The main result to emerge from Tables 2 and 3 is the overwhelming preponderance of syntactic errors over lexical errors. Of the 128 errors considered, 111 (86.7%) were purely syntactic, 3 (2.3%) were purely lexical, and 14 (10.9%) were mixed. The only child that clearly departed from this trend was Teresa, who was responsible for 10 of the 17 lexical errors observed. The difference between the +L −S and −L +S columns of Table 2 was significant ($P < 0.0005$, Wilcoxon).

Table 3 shows that the trend was obtained for all the groups considered. As one would expect, there were more errors in groups 3b, 3c, 4b, and 4c

TABLE 2
Lexical and Syntactic Responses by Each Subject to Stimulus Groups 3a–4c

Name	+*L* +*S*	+*L* −*S*	−*L* +*S*	−*L* −*S*
Alberto	18	0	0	0
Bruno	12	6	0	0
Clara	12	6	0	0
Daria	10	8	0	0
Emilio	9	8	0	1
Franco	6	12	0	0
Guido	6	11	0	1
Lisa	6	9	0	0
Marco	8	10	0	0
Nino	3	15	0	0
Paola	6	8	0	1
Rita	3	0	1	1
Silvio	2	11	0	2
Teresa	4	4	2	8
Vittorio	3	3	0	0
Total	108	111	3	14

than in 3a and 4a; there were also rather more errors for four-digit numerals than for three-digit numerals. The significance of these differences was confirmed by a two-way analysis of variance on the number of correct responses given by each subject for each group. The two factors were the length of the correct response (3 or 4) and its syntactic type (a, b, or c). Both main effects were significant [for length, $F(1,84) = 10.97$, $P < 0.001$; for syntactic type, $F(2,84) = 21.05$, $P < 0.0001$]. No interaction effect was obtained [$F(2,84) = 0.17$].

TABLE 3
Lexical and Syntactic Errors for Each Stimulus Group

Group	+*L* +*S*	+*L* −*S*	−*L* +*S*	−*L* −*S*
3a	40	4	0	1
3b	13	22	2	3
3c	19	22	1	1
4a	29	9	1	1
4b	3	31	0	5
4c	4	24	0	2

Common Syntactic Errors

Appendix 1 reports the first three responses of each subject for each group in the range 3a–4c. The most common syntactic errors in each group were as follows.

3a. Only 4 syntactic errors were obtained; no pattern emerged.
3b. 22 syntactic errors were obtained. 14 of these were of type X00XX (10025, 20081, 30012) and 6 of type X0XX (1025, 2081, 3012). (X here denotes any non-zero digit.)
3c. 22 syntactic errors were obtained, of which 18 were of type X00X (1006, 2005, 3003).
4a. 9 syntactic errors were obtained. 5 of these were due to the addition of an extra zero (10000, 20000, 30000).
4b, 4c. The errors in these groups were less consistent. Except for Alberto, all the children introduced unwarranted zeros for mille or mila (e.g. mille sette cento 1000700), but the number of zeros introduced varied from one to five:

Stimulus:	tre mila cento novanta quattro	
Responses:	30194	(Bruno)
	30010094	(Clara)
	300010094	(Daria)
	30000194	(Marco)
	300000100904	(Emilio)

For many of the errors categorised above, the child seems to be forming constituents of the Arabic numeral separately and then concatenating them. For example, cento venti cinque is composed of cento (100) and venti cinque (25); concatenate these and you obtain 10025. This type of error was notably absent for components below 100—venti cinque was not rendered as 205. In the whole of the experiment, only a single instance of such an error was observed: for the stimulus tre mila cento novanta quattro, Emilio produced the numeral 300000100904.

Bizarre Errors

Paola, Silvio, and Teresa produced systematic errors which were not of the above types. The following responses by Paola seem to result from misinterpretation of the verbal numeral:

Stimulus	*Response*
cento venti cinque	2500
cento sei	600

The same applies to these errors by Silvio:

Stimulus	*Response*
due cento	102
tre cento	103
quattro cento	104

However, we have no explanation for his response to group 4a:

Stimulus	*Response*	*Stimulus*	*Response*
mille	1000	sei mila	16000
due mila	1200	sette mila	7000
tre mila	1300	otto mila	8000
quattro mila	1400	nove mila	9000
cinque mila	1500		

Most remarkable of all was Teresa's treatment of group 3a:

Stimulus	*Response*	*Stimulus*	*Response*
cento	2000	sei cento	6000000
due cento	20000	sette cento	70000000
tre cento	3000	otto cento	800000000
quattro cento	40000	nove cento	9000000000
cinque cento	500000		

It seems here that after two or three bemused responses, Teresa decides to impose a rule of her own: cento is translated by a string of zeros equal in length to the number preceding it. When tackling group 4a, Teresa began by copying some responses from 3c (she had been helped here by the experimenter), and then invented a new rule for translating mila:

Stimulus	*Response*	*Stimulus*	*Response*
mille	1000	sei mila	606
due mila	205	sette mila	707
tre mila	303	otto mila	808
quattro mila	404	nove mila	909
cinque mila	505		

DISCUSSION

Theoretical Background

A formal theory of the perception, production, and learning of numerals has been proposed by Power and Longuet-Higgins (1978). The theory takes the form of a computer program that can learn natural numeral systems from examples. Before discussing the results of the present study, it is convenient to review the main implications of this earlier work:

1. When people translate verbal numerals into Arabic numerals (or vice versa) they construct an intermediate semantic representation. Some evidence from studies of acalculia supports this view (McCloskey et al., 1985; Sartori, Roncato, Rumiati, & Maso, 1985). Moreover, there are strong general grounds for supposing that some kind of intermediate representation is created. A complex numeral such as one thousand two hundred and thirty four has a clearly identifiable syntactic structure which can be represented by a tree or by an equivalent bracket notation:

((one thousand) ((two hundred) and (thirty four)))

This syntactic structure is articulated when the numeral is read aloud (e.g. by marki ıg the boundary after "thousand"); it can also be linked by precise rules to a semantic structure based on the operations of addition and multiplication (Hurford, 1975; Stampe, 1976). These linguistic results are not in doubt. It therefore seems natural to assume the construction of some kind of intermediate representation when a verbal numeral is perceived during the transcoding task. We should mention, however, that not all researchers share this assumption: Deloche and Seron (1987) review some acalculia studies suggesting that transcoding and calculation abilities are independent; they also describe two transcoding procedures which are "asemantic" in the sense that they do not imply the construction of an intermediate representation.

2. The form of the semantic representation reflects the structure of the subject's verbal numeral system. For Italian (and for most other European languages), the primitive numerical concepts are C1, C2, . . ., C10, C100, C1000 (for numbers up to 1 million). The prefix C here serves to distinguish semantic concepts from Arabic numerals. Every non-primitive number is represented as the sum or product of two unequal numbers. Following Power and Longuet-Higgins (1978), we shall refer to the larger of these numbers as the "major term" and the smaller as the "minor term". We shall also adopt the convention that the major term of an arithmetical expression is placed before the minor term. Thus the number sixty is represented by the expression:

C10 * C6

where C10 is the major term and C6 the minor term. Sixty five is represented by:

(C10 * C6) + C5

where C10 * C6 is the major term and C5 the minor term. And the number of days in a year is represented by:

(C100 * C3) + ((C10 * C6) + C5)

Note that this method of distinguishing the major and minor terms is employed for reasons of typographical convenience; in a computer model one would probably prefer a record structure with slots for major term, minor term, and arithmetical operation. Such a structure, with no implications of serial order, was used in Power and Longuet-Higgins' program.

An alternative semantic representation, based on the Arabic numeral system, has been suggested by McCloskey, Sokol, and Goodman (1986). The representation consists of a series of terms of the form A * (10 ^ N) where A is a number in the range 0–9 and the symbol ^ denotes exponentiation. By assigning to N the values 0, 1, 2, 3..., we obtain units, tens, hundreds, thousands, etc. The two theories yield identical representations for numerals up to 9999, but thereafter they diverge: according to our view, the semantic representation of twenty three thousand is C1000 * ((C10 * C2) + C3); according to McCloskey et al. it is effectively (C10000 * C2) + (C1000 * C3). At present, we know of no empirical evidence which might help to choose between these representations. On general grounds we prefer a representation based on verbal numeral systems because of their greater universality. A small additional point is that the method in common use for punctuating Arabic numerals gives priority to the lexicalised multiplicands C1000 and C1000000: For example, the numeral 23000 is written as 23,000 so that the thousands (not the ten thousands) are separated from the lower orders.

3. The translation of a verbal numeral into an Arabic numeral proceeds in two stages. First, the subject interprets the verbal numeral, constructing the corresponding semantic expression. For example:

cento venti cinque => C100 + ((C10 * C2) + C5)

Next, starting from this semantic expression, the subject produces the Arabic numeral:

C100 + ((C10 * C2) + C5) => 125

Although it is convenient to regard these stages as sequential, they could in practice be interleaved. For example, if a numeral began with the words twenty thousand, the transcoder could safely write down the digits 20, which are compatible with all possible continuations; such a procedure is presumably essential if the number exceeds our short-term memory span. However, we think it is useful to distinguish the rules which underlie linguistic competence from those which allow special performance skills such as on-line transcoding. Analogously, we can distinguish a person's knowledge of two languages, English and Italian perhaps, from the special skills employed when performing a simultaneous translation from one to the other.

4. Arabic numerals are produced by applying a set of rules of the kind shown in Appendix 2. We will assume henceforth that the reader is familiar with the notational conventions introduced in this appendix.

Explanation of Errors

The clearest error patterns observed were those of groups 3b and 3c. For convenience, these are recapitulated below:

1.	*Correct response*:	XXX	(e.g. 125)
	Subject's response:	X00XX	(e.g. 10025)
2.	*Correct response*:	XXX	(e.g. 125)
	Subject's response:	X0XX	(e.g. 1025)
3.	*Correct response*:	X0X	(e.g. 106)
	Subject's response:	X00X	(e.g. 1006)

In all these cases, the subject seems to interpret the Italian numeral correctly but then to go astray in producing the Arabic numeral. Subjects who produced such errors could dictate numerals below 100 and numerals of type 3a (100, 200, 300, etc.); we may therefore assume that they had acquired the non-recursive rules R1–R11 of Appendix 2 together with the product rules R13 and R14. The rule they lack is R16, which realises sums:

<A + B> = <A> # <B>

The subjects who dictated cento venti cinque as 10025, and cento sei as 1006, seem to be applying the following alternative to R16:

<A + B> = <A> & <B>

in which the (familiar) concatenation operator "&" replaces the (peculiar) overwriting operator "#".

Let us spell out this explanation in full by considering the dictation of cento sei. Assuming that this Italian numeral is interpreted correctly, it will be assigned the meaning C100 + C6. This semantic expression matches the alternative version of R16:

<C100 + C6> = <C100> & <C6>

By rules R11 and R6, <C100> = 100 and <C6> = 6. Therefore,

<C100> & <C6> = 100 & 6 = 1006.

So far so good. However, if we attempt a similar derivation for the error 10025, we encounter what a computer programmer would call a "bug". Assuming that cento venti cinque is correctly interpreted, the semantic form to be realised is

<C100 + ((C10 * C2) + C5)>

On applying the variant version of R16, we obtain:

<C100> & <(C10 * C2) + C5>

Now comes the bug. Re-application of the erroneous sum rule to the expression <(C10 * C2) + C5> will ultimately yield the numeral 100205, not 10025. The results of the present study show that errors of the type *venti cinque* => 205 almost never occur. Apparently, then, a single sum rule is insufficient. The subjects must possess at least two sum rules, one for numbers below a hundred, and one for numbers above a hundred.

How can the former of these rules (let us call it R16a) be formulated? One possibility is as follows:

<(C10 * A) + B> = <C10 * A> # <B>

However, this is implausible for two reasons: first, it contains the difficult operator #; secondly, if the subjects already possessed such a rule, they should have little trouble generalising it to numbers above a hundred. Fortunately, there is an alternative:

<(C10 * A) + B> = <A> & <B>

This rule works, it employs concatenation and, moreover, it is simpler than the rule previously suggested. We may therefore hypothesise that some of the children were applying two sum rules, both based on concatenation, in place of R16:

R16a: <(C10 * A) + B> = <A> & <B>
R16b: <A + B> = <A> & <B>

Because some expressions—e.g. <(C10 * C2) + C5>—match both these rules, we must also stipulate that R16a has precedence over R16b. A complete derivation of the error 10025 can now be given:

<C100 + ((C10 * C2) + C5)>	
= <C100> & <(C10 * C2) + C5>	(rule R16b)
= <C100> & (<C2> & <C5>)	(rule R16a)
= 100 & (2 & 5)	(rules R11, R2, R5)
= 100 & 25	(concatenation)
= 10025	(concatenation)

There are indications that a few of the children had invented overwriting operations of their own which worked in some cases but not in others. Daria, for example, responded correctly to 3c (e.g. *due cento cinque* => 205) but made errors of type X0XX for 3b (*due cento ottantuno* => 2081).

Conceivably, she had invented an operator (call it ##) that overwrites only the *final* zero:

1000 ## 2 = 1002
1000 ## 23 = 10023
1000 ## 234 = 100234

With such an operator, Daria could formulate the following sum rule:

<A + B> = <A> ## <B>

This works correctly for numbers below a hundred:

<(C10 * C2) + C5>
= <C10 * C2> ## <C5>
= 20 ## 5
= 25

and for numbers in group 3c:

<C100 + C6> = <C100> ## <C6> = 100 ## 6 = 106

but not for numbers in group 3b:

<C100 + ((C10 * C2) + C5)>
= <C100> ## <(C10 * C2) + C5>
= 100 ## 25
= 1025

Leaving aside such speculations, our main conclusion is that the children erred because they lacked a correct rule for combining the components of a sum. For numbers below a hundred, a special rule based on concatenation can be formulated (R16a); for numbers above a hundred, however, the overwriting operator is needed. This operator is difficult to learn because it is specific to the Arabic numeral system; natural languages rely for the most part on concatenation. To cite just one example,

S = NP & VP

is a concatenation rule. So far as we know, it is only at the morphological level that rules similar to overwriting occur. For instance, the rule for forming the plural of English nouns ending in -y (preceded by a consonant) requires the replacement of the terminal -y by -ies (e.g. city => cities). If our assumption is correct, such rules should take longer to learn than those employing straightforward concatenation (e.g. cat & s = cats).

Comparison with Studies of Acalculia

The most striking difference between our data and those obtained in studies of acalculia concerns the relative frequency of syntactic and lexical errors. Of the 887 dictation errors collected by Deloche and Seron (1982), 50% were syntactic and 27% were lexical; of our corpus of 128 errors, 87% were syntactic and less than 3% were lexical (the remaining 10% were mixed). For 14 of our 15 subjects, the number of lexical errors was negligible: We do not believe that normal adult subjects would have performed any better. The remaining subject, Teresa, obviously has special difficulties.

Despite this difference in relative frequency, the *kinds* of syntactic error made by children and acalculia patients are similar—so similar, in fact, that we cannot see any features that distinguish them. In each case, the problem lies in the insertion of zeros rather than in the misordering of digits: the numeral cento venti cinque is dictated as 10025 or 1025 but not as 521 or 251. In each case, moreover, the error of inserting extra zeros does not occur for numbers below a hundred: The apparently consistent mistake of dictating venti cinque as 205 is conspicuously absent.

An interesting issue raised by this comparison can be crudely stated thus: Is acalculia second childhood? In general, of course, the answer to this question must be no, as the errors made by some acalculia patients are lexical rather than syntactic. In the case of a patient who produces only syntactic errors, however, we might conjecture that his or her transcoding rules have simply been wiped out by damage to a specific area of the brain, and that he or she has been obliged to learn them again from scratch more or less as a child would. Because the overwriting operation # is specific to the Arabic numeral system, it too would have to be learned from scratch, and there does not seem to be any reason why its counter-intuitive nature should not trouble a 57-year-old just as much as a 7-year-old.

Manuscript received October 1986
Revised manuscript accepted January 1990

REFERENCES

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

Deloche, G. & Seron, X. (1987). Numerical transcoding: A general production model. In G. Deloche & X. Seron, *Mathematical disabilities: A cognitive neuropsychological perspective*. Hillsdale, N.J.: Lawrence Erlbaum Associates Inc.

Hurford, J. R. (1975). *The linguistic theory of numerals*. Cambridge: Cambridge University Press.

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., Sokol, S. M., & Goodman, R. A. (1986). Cognitive processes in verbal-number production: Inferences from the performance of brain-damaged subjects. *Journal of Experimental Psychology: General*, *115*, 307–330.

Power, R. J. D. & Longuet-Higgins, H. C. (1978). Learning to count: A computational model of language acquisition. *Proceedings of the Royal Society of London*, *B200*, 391–417.

Sartori, G., Roncato, S., Rumiati, R., & Maso, A. (1985). *Number processing and dyscalculia*. Unpublished ms, Dipartimento di Psicologia Generale, Universita di Padova, Italy.
Singer, H. D. & Low, A. A. (1933). Acalculia (Henschen): A clinical study. *Archives of Neurology and Psychology*, *29*, 476–498.
Stampe, D. (1976). Cardinal number systems. In *Proceedings of the 12th annual conference of the Chicago Linguistics Society*. Chicago: Chicago University Press.

APPENDIX 1: RESPONSES FOR GROUPS 3a–4c

We reproduce here the first three responses of each subject for the groups 3a–4c. The stimuli were as follows:

3a.	cento due cento tre cento	4a.	mille due mila tre mila
3b.	cento venti cinque due cento ottantuno tre cento dodici	4b.	mille sette cento due mila sei cento dieci tre mila cento novanta quattro
3c.	cento sei due cento cinque tre cento tre	4c.	mille due due mila settanta quattro tre mila cinque cento otto

The responses are tabulated below. Gaps in the table indicate that the subject gave up. Asterisks indicate that the subject was helped. The correct responses are those given by Alberto.

Name	*3a*	*3b*	*3c*	*4a*	*4b*	*4c*
Alberto	100	125	106	1000	1700	1002
	200	281	205	2000	2610	2074
	300	312	303	3000	3194	3508
Bruno	100	125	106	1000	10700	10002
	200	281	205	2000	2060010	20074
	300	312	303	3000	30194	30508
Clara	100	125	106	1000	100700	1002
	200	281	205	2000	20060010	20074
	300	312	303	3000	30010094	3005008
Daria	100	125	106	1000	1007100	10002
	200	2081	205	2000	2000610010	200074
	300	3012	303	3000	300010094	30005008
Emilio	100	10025	106	1000	1000700	1002
	200	20081	205	2000	20000070010	2000074
	300	30012	303	3000	300000100904	300005008
Franco	100	10025	1006	1000	10007100	10002
	200	20081	2005	2000	200060010	200074
	300	30012	3003	3000	300010094	30005008

Name	*3a*	*3b*	*3c*	*4a*	*4b*	*4c*
Guido	100	10025	1006	1000	100077	100002
	200	20081	2005	2000	2000610	2000074
	300	30012	3003	3000	300010094	300005008
Lisa	100	10025	1006	1000	10007100	
	200	20081	2005	2000	20006100010	
	300	30012	3003	3000	3000100094	
Marco	100	125	106	10000	10000700	100002
	200	281	205	20000	2000060010	2000074
	300	312	303	3000	30000194	300005008
Nino	100	1025	1006	10000	10700	100002
	200	2081	2005	20000	2000610	200074
	300	3012	3003	30000	301094	305008
Paola	100	2500	600	1000*	000700	
	200	20081	2005	2000	200060010	
	300	30012	3003	3000	3000004	
Rita	100	5000				
	200	261				
	300					
Silvio	100	1505	1006	1000	10071	
	102	210011	21005	1200	2100610010	
	103	310012	31003	1300	310010094	
Teresa	2000	195	10000	1000	0720	02
	20000	201	205*	205	2710	2074
	3000	3012	303	303	306094	20708
Vittorio	100*		1006			
	200		2005			
	300		3003			

APPENDIX 2: RULES FOR PRODUCING ARABIC NUMERALS

We give here a complete set of rules for producing Arabic numerals for numbers below 1 million. The rules are labelled R1, R2, etc., for purposes of reference. Angle brackets are employed in order to represent the numeral corresponding to a given semantic expression: thus <C100> denotes the Arabic numeral corresponding to C100, namely 100. When considering rules like R11, it is crucial to keep in mind the distinction between semantic concepts and Arabic numerals. Anyone who thinks that the formula <C100> = 100 is tautologous is forgetting that C100 is a primitive symbol.

R1	<C1>	= 1
R2	<C2>	= 2
R3	<C3>	= 3
R4	<C4>	= 4
R5	<C5>	= 5

R6	<C6>	= 6
R7	<C7>	= 7
R8	<C8>	= 8
R9	<C9>	= 9
R10	<C10>	= 10
R11	<C100>	= 100
R12	<C1000>	= 1000
R13	<C10 * A>	= <A> & 0
R14	<C100 * A>	= <A> & 00
R15	<C1000 * A>	= <A> & 000
R16	<A + B>	= <A> # <B>

The recursive rules R13–R16 employ two string operators, & and #. The former represents concatenation:

3 & 00 = 300
120 & 000 = 120000

The latter represents an operator which we will call "overwriting". The meaning of this operator is shown by the following examples:

1000 # 2 = 1002
1000 # 23 = 1023
1000 # 234 = 1234

As can be seen, the # operator takes two strings, which we may call X and Y, and produces a new string by overwriting the terminal zeros of X with Y. A fuller name for the operator would thus be "overwriting zeros from the right". The operator cannot be applied unless X contains a substring of terminal zeros at least as long as the whole of string Y.

To show how the rules are used, we will derive the Arabic numeral for the number two hundred thousand and thirty four, an example which tests whether they can cope with a plethora of zeros:

<(C1000 * (C100 * C2)) + ((C10 * C3) + C4)>	
= <C1000 * (C100 * C2)> # <(C10 * C3) + C4>	(rule R16)
= <(C100 * C2> & 000) # <(C10 * C3) + C4>	(rule R15)
= ((<C2> & 00) & 000) # <(C10 * C3) + C4>	(rule R14)
= ((2 & 00) & 000) # <(C10 * C3) + C4>	(rule R2)
= ((2 & 00) & 000) # (<C10 * C3> # <C4>)	(rule R16)
= ((2 & 00) & 000) # ((<C3> & 0) # <C4>)	(rule R13)
= ((2 & 00) & 000) # ((3 & 0) # 4)	(rules R3, R4)
= (200 & 000) # (30 # 4)	(concatenation)
= 200000 # (30 # 4)	(concatenation)
= 200000 # 34	(overwriting)
= 200034	(overwriting)