137

Linguistic Relativity Revisited: The Bilingual Word-length Effect In Working Memory
During Counting, Remembering Numbers, and Mental Calculation.

Nick Ellis
University College of North Wales

ABSTRACT

In different languages the names of numbers take different times to articulate. This chapter considers the role of language and representation in arithmetic. It reviews studies which demonstrate that digit word-length limits the short-term memory for digit sequences (such as telephone numbers or digit span as used in many intelligence tests). Three experiments are reported which show that a language's number name word-lengths have a determinative influence upon the ease of mental calculation and counting in that language - some languages are more conducive to mental arithmetic than others. More general aspects of the effects of language word-length are also considered.

INTRODUCTION

The facility of human cognitive processes depends on the internalisation of effective representational systems and the greatest of all such systems is language. "Human beings do not live in the objective world alone, nor alone in the world of social activity as ordinarily understood, but are very much at the mercy of the particular language that has become the medium of expression for their society" (Sapir, in Spier, 1941). "A change in language can transform our appreciation of the Cosmos." (Whorf, 1956). "'This way,' says the word, 'is an interesting thought: come and find it.' And so we are led on to rediscover old knowledge." (Cooley, 1962). "Speech is the best show man puts on. It is his own 'act' on the stage of evolution, in which he comes before the cosmic backdrop and really 'does his stuff'." (Whorf, 1942).

"The mathematical formula that enables a physicist to adjust some coils of wire, tinfoil plates, diaphragms, and other quite inert and innocent gadgets into a configuration in which they can project music to a far country puts the physicist's consciousness on to a level strange to the untrained man. ... We do not think of the designing of a radio station or a power plant as a linguistic process, but it is one nonetheless. The necessary mathematics is a linguistic apparatus, and, without its correct specification of essential patterning, the assembled gadgets would be out of proportion and adjustment, and would remain inert. ... the mathematics used in such a case is a *specialised* formula-language, contrived for making available a specialised type of force manifestation through metallic bodies only, namely, *electricity*, as we today define what we call by that name." (Whorf, 1942).

The history of mathematics, of logic, of computation, even of thought, is marked with the milestones of the developments of representational systems.

In the case of mathematics the invention of positional notation was of enormous significance for civilisation. Early systems of numeration used by the Egyptians, Hebrews, Greeks and Romans were based on a purely additive system. Thus in the Roman symbolism, for example, one wrote: CXVIII = one hundred + ten + five + one + one + one.

A disadvantage of such additive notations was that more and more new symbols were needed as numbers got larger. But the chief problem was that computation with numbers was so difficult that only the specialist could handle any but the simplest problems. To realise the efficiency of our present positional notation, we have only to try to perform an addition by means of Roman numerals, for example:

CCLXVI	266
MDCCCVII	1807
DCL	650
MLXXX	1080
MMMDCCCIII	3803

Without converting the Roman numerals into our modern system the problem is difficult, if not impossible to solve. And this is only an addition - multiplication or division would be even worse. Such systems do not lend themselves to calculation because of the static nature of their basic numerals, which are essentially only abbreviations for recording the results of calculations already done by means of a counting board or abacus. "That is why, from the beginning of history until the advent of our modern positional numeration, so little progress was made in the art of reckoning." (Dantzig, 1930).

Additive systems are quite different from place-value (positional) systems which were independently conceived only four times in history. Three of these conceptions were by the Babylonians (in the early second millennium B.C.), the Mayas (probably in the Classic Period, third to ninth centuries A.D.), and the Chinese (shortly before the beginning of the Christian era). But these place-value systems were defective in comparison with the numeration developed by the Hindus that is still in use. This positional system has the agreeable property that all numbers, however large or small, can be represented by the use of a small set of different digit symbols (in the decimal system these are the Arabic numerals 0, 1, 2, ..., 9) and the place-value principle is used consistently with powers of the base 10. In conjunction with the place-value principle, discovery of the zero made the decisive stage in a process of development without which we cannot imagine the progress of modern mathematics, science, and technology. The zero freed human intelligence from the counting board that had held it prisoner for thousands of years, eliminated all ambiguity in the written expression of numbers, revolutionized the art of reckoning, and made it accessible to everyone (Ifrah, 1987). The most important advantage is that of ease of computation. The rules of reckoning with numbers represented in positional notation can be stated in the form of addition and multiplication tables for the digits, and these can be memorised once and for all. As Courant & Robbins (1941) extol: "The ancient art of computation, once confined to a few adepts, is now taught in elementary school. There are not many instances where scientific progress has so deeply affected and facilitated everyday life."

Another key example of the crucial importance of representation is to be found later in the development of mathematics where we have retained the notation of Leibniz, dx/dy for the derivative, $f'(x)$ and $\int f(x)\,dx$ for the integral because it is extremely useful, allowing the limits of quotients and sums to be handled 'as if' they were actual quotients or sums, notwithstanding the fact that Leibniz' explanations and theory was clearly surpassed by Newton's - "Leibniz' notation is at least an excellent notation for the limit process; as a matter of fact, it is almost indispensable in the more advanced parts of the theory." (Courant & Robbins, 1941, p. 435).

As written representations vary in their efficiency, so do mental representations. Our entry into mathematics, the very beginnings which may one day allow us to consider the sublime calculus and beyond, lies in our being taught about number, counting and simple arithmetic. And each language has its own names for the digits and the operations thereon. These surface

features of language are, at first sight, an unlikely locus for cognitive constraint. Yet this chapter will demonstrate that such a simple feature as the time it takes to pronounce the names of the digits affects the ability of a native speaker of a language to remember numbers, to count, and to perform mental calculations.

DIGIT LENGTH AND SHORT-TERM MEMORY

BILINGUAL DIGIT NAMING RATE

In Gwynedd in North Wales over 60% of the population speak Welsh (Baker, 1985). In 1978 I was attempting to learn the language. Casual observation suggested that it mostly takes longer to articulate the names for digits in the Welsh language (*dim, un, dau, tri, pedwar, pump, chwech, saith, wyth, naw, deg*) than their English equivalents (*nought, one, two, three, four, five, six, seven, eight, nine, ten*). Ellis & Hennelly (1980) therefore tested this in 12 bilingual subjects who were required to read aloud as fast as possible 200 instances of randomly ordered digits in English, and in Welsh. There was a highly significant difference in reading time for the two languages: even though only one-third of the subjects rated themselves more competent in English than in Welsh, every subject read the digits faster in English. It took on average 385 ms to read a Welsh digit compared with 321 ms to read an English digit. That is, on average, a subject would read six digits in English in the time taken to read five in Welsh.

These cross-language digit name length differences may affect performance in tasks where vocal or subvocal articulation of digit names is involved, i.e. in short-term memory (STM) for digits, in counting, and in mental arithmetic.

BILINGUAL STM SPAN

Baddeley, Thomson & Buchanan (1975) demonstrated that immediate memory span for short words is greater than that for long words. This effect cannot be solely attributed to the number of syllables or phonemes in the stimulus. Rather the effect is truly one of word-length: even when the number of syllables and phonemes is held constant, the memory span for words which take a short time to articulate (e.g. wicket, phallic) is greater than that for words which take a long time to articulate (e.g. zygote, coerce). In general the span could be predicted on the basis of the number of words which the subject could read in approximately 2 s. Baddeley (1986) interprets such word-length effects in terms of the Working Memory model. In the original formulation (Baddeley & Hitch, 1974) items are encoded in STM in an articulatory code. Loss of information occurs by passive decay, but this can be countered by rehearsing the traces of decaying items. As long as all the items in a sequence can be refreshed within the decay time of the store, they can be maintained more or less indefinitely. If, however, the length of a sequence of spoken items exceeds the decay time, errors in recall will occur. Thus the rehearsal process is limited by temporal duration, and the articulatory loop is seen to be analogous to a tape loop of specific length which can hold a message which fits onto that length of tape. Thus subjects' STM capacity is limited to roughly the amount of material that can be rehearsed sub-vocally in about 2 seconds (Baddeley, 1986; but see Gordon Brown [this volume] for some qualifications of this).

In combination with the bilingual digit-name length differences, these findings led to the prediction that the immediate memory span for English digits will be greater than for Welsh digits, even in subjects who consider themselves more competent in the Welsh language. Ellis & Hennelly (1980) therefore tested the same 12 subjects for their STM for Welsh and English digits.

For each condition the stimuli consisted of three trials at each length of string from two to ten digits and these were presented in ascending order of length. The subject listened to the string and, upon a cue to respond, tried to repeat the digits in the correct order and the same language of presentation, continuing in this fashion until there had been incorrect responses on three consecutive trials. STM span was calculated as 1 + (number of trials correct/3).

The mean STM span for English was 6.55 which was significantly greater than that of 5.77 for Welsh digits (p< 0.01). This was the case even though the majority of subjects were more proficient in the Welsh language and the difference is consistent with an explanation that it results from the word-length effect whereby Welsh digits take longer to articulate than English digits.

Baddeley *et al.* (1975) demonstrated that a subject's span could be predicted to be the number of words that could be read in approximately 2 s, and concomitantly demonstrated a significant correlation between subjects' reading speed and memory span. Both of these findings were confirmed in Ellis & Hennelly (1980): (i) A Spearman rank-order correlation between the 12 subjects' digit reading speeds and their STM spans was significant at rho = 0.47 (p< 0.05); (ii) The mean time taken to read a Welsh digit was 385 ms, the mean Welsh span was 5.77; this number of digits could thus be read in 2.2 s. Comparable figures for the English language are a digit span of 6.55 items which at a reading rate of 321 ms/digit could be read in 2.1 s.

BILINGUAL STM SPAN UNDER ARTICULATORY SUPPRESSION

Digit span measured in the Welsh language is thus smaller than that measured in English. It is not possible to conclude, however, that this is necessarily an effect of word-length: both the span and reading rate differences might be attributable either to word-length differentials or to differences in degree of familiarity. This latter possibility must be considered as it seems that Welsh speakers do on occasion preferentially use English number names. For example, the year of the Ellis & Hennelly experiments was often referred to as 'nineteen seventy-eight' in preference to 'mil naw saith wyth' or the more clumsy 'un mil naw cant saith deg wyth'. It is thus possible that numbers are a special case of language usage, and therefore the language competence self-ratings obtained for our bilingual subjects may not represent their language of preference when dealing with numbers.

Effects of word-length and familiarity can be distinguished if articulatory suppression is used as an interference task. The word-length effect, which Baddeley *et al.* (1975) attribute to the functioning of the articulatory loop, is much reduced with visual stimulus presentation if the subject's articulation is simultaneously suppressed by their repeatedly whispering an irrelevant phrase such as 'the the the ...'. Therefore if the difference between English and Welsh digit spans is the result of the differential articulation time of the digit names, i.e. if it is a word-length effect, this difference should be either absent under articulatory suppression, or, if present, present in a much reduced form.

Eight bilingual subjects were therefore tested for their digit spans in Welsh and English with visual presentation and articulatory suppression. The digit strings were presented sequentially on a memory drum at a rate of one item per second. To ensure that the stimuli were processed in the required language, digit words were presented, e.g. 'pedwar' or 'four', as opposed to the digit figures. The subjects were again required to report the component digits of the strings in the correct order at the end of string presentation. The major difference between this and the prior procedure was that throughout the period of digit string presentation the subject was to whisper the sequence 'a-b-c-d' in a continuous cycle at the fastest rate compatible with clarity

of pronunciation. The subjects were tested on both conditions with order of presentation counterbalanced.

The mean digit spans under articulatory suppression in Welsh and English were 3.75 and 4.00 respectively, a non- significant difference. These figures are to be compared with those of 5.77 and 6.55 where no suppression was used and stimulus presentation was auditory.

It must therefore be concluded that the bilingual digit span differential is a word-length effect. Even for subjects who consider themselves more proficient in Welsh, the structure of the Welsh digit names necessitates that it is easier to remember lists of numbers in English. This effect, albeit relatively small (the English span being 114 per cent that of the Welsh span) must be assumed to be operative in everyday situations such as the short-term remembering of telephone numbers.

INTELLIGENCE TESTING

Individual differences in the span of immediate memory, as measured using strings of random digits as stimuli, have commonly been utilized as sub-components of intelligence tests. In the Terman-Merrill (1974), for example, a 10 year old child is tested on their ability to repeat six-digit strings in the correct order. Similarly, in the Wechsler Intelligence Scale for Children (WISC, 1949) the same age child is tested for their ability to repeat digit strings both in their original and reversed order. The sum of forwards and reversed spans measured on this test are compared with the norm score of 9 for a child of this age.

The development or modification of intelligence tests for use with different languages or dialects must be accompanied by re-normalisation. As Burt (1939) stated in reference to the use of the WISC in England: testers in England 'should be supplied with a standardised procedure and with standardised norms - a procedure which has been experimentally adjusted to English idioms and to English customs, norms which have been statistically deduced from extensive trials with English children, trained in English homes and taught in English schools.' Norms for different adaptations of an intelligence test should not be directly compared with an aim to deducing intellectual differences between the populations from which these norms were derived. Our demonstrations reinforce this claim - cross-lingual differences in word-length result in different magnitudes of digit span as measured in those languages and this entails that digit span norms cannot be compared across languages as an indicator of cultural intellectual differences.

Table 1									
Digit span scores (sum of digit spans forwards and reversed) for the American population tested in English on the WISC procedure, and the Welsh population tested in Welsh on the WCIS translation of WISC digit span procedure.									
Subject age (years)	6.10	7.10	8.10	9.10	10.10	11.10	12.10	13.10	14.10	15.10
WISC digit span score	7	8	8	9	9	10	10	10	11	11
WCIS digit span score	7	7	8	8	8	9	9	9	9	10

William & Roberts (1972) developed a Welsh language Children's Intelligence Scale (WCIS) by modifying and translating the Wechsler Intelligence Scale for Children (WISC).

The WISC was experimentally adjusted to Welsh idioms and to Welsh customs and norms were statistically deduced from extensive trials with Welsh-speaking children taught in Welsh schools. The digit span sub-test of the WCIS, was in effect, a direct translation of that of the WISC, the same digit strings are used. The norms on this test are compared to those of the original WISC in Table 1 where the digit span figures represent the sum of digit span forwards and digit span reversed. It can be seen that the norms for the Welsh sample are reliably less than those of the American sample. However this cannot be taken to imply intellectual differences between the two populations, rather they are the result of the differing languages - English digits are easier to remember than Welsh digits as a consequence of their word-length.

RECENT CONFIRMATIONS IN OTHER LANGUAGES

Ellis & Hennelly (1980) suggested that this digit name length effect would also operate in other languages, i.e. languages would be more or less conducive to number memorability and manipulation/calculation as a function of the word-length of the languages' number names. They called for a survey of the word-lengths of the digit names in a wide variety of languages and there have since been a number of replications in other languages.

Stigler, Lee & Stevenson (1986) demonstrated that Mandarin Chinese number words (*i, er, san, si, wu, liu, chi, ba, jiou, shi*) were of a significantly shorter pronunciation (0.40 s per digit for university students) than English number words (0.53 s per digit; see also Liu & Shen, 1977). {Note that here and hereafter, mean digit naming times are not comparable *across* studies because of differences in procedures and subjects in the different experiments}. Associated with these differences the mean digit span for the Chinese subjects was 9.2 whilst that of the Americans was 7.2. {We might speculate that, had George Miller (1956) been of Chinese extraction, the magical number would have been 9!}. As in the Working Memory model, Stigler *et al.* (1986) interpret the finding that the total pronunciation duration for a subject's maximum span did not differ between Chinese and Americans as evidence for a temporally limited store.

Naveh-Benjamin & Ayres (1988) investigated digit word-length and memory span in English, Spanish, Hebrew and Arabic. The mean number of syllables per word for the digits used (0-6, 8-9) were 1.0, 1.6, 1.9, and 2.3 respectively and this led to reliable differences in digit naming time with English fastest at 0.26 s per digit and Arabic slowest at 0.37 s per digit. As a result the digit STM spans in these languages were 7.2, 6.4, 6.5 and 5.8 respectively. Again digit span was approximately predicted by the number of digits that could be read in 2 s.

As Baddeley (1990) observes, the record so far for speed of articulation goes to Cantonese speakers of Chinese residing in Hong Kong. Hoosain (1979) had demonstrated that the digit span for such undergraduates was 9.9 and Hoosain (1986, 1987) and Hoosain and Salili (1987) showed this to be a result of the pronunciation speed and sound duration (0.31 s per digit) of Cantonese number names compared to English (0.38 s per digit). Hoosain and Salili used the identical procedure as Ellis & Hennelly to measure digit reading speed in Chinese and this resulted in an estimate of 0.26 s per digit compared with 0.32 s for English and 0.39 s for Welsh in Ellis & Hennelly.

In summary, it is clear that there are differences between languages in the lengths of their digit names and these affect the time it takes a native speaker to articulate them. Material in STM decays rapidly unless it is refreshed by the use of the articulatory loop for rehearsal. Thus bilingual digit-name length differences affect performance in tasks such as short-term memory for digits which involve vocal or subvocal articulation of digit names. Other potential tasks where these surface features of language might play a determining role include counting and mental arithmetic.

DIGIT LENGTH AND MENTAL CALCULATION

Mental arithmetic falls into two distinct classes: associative and procedural. Some answers (e.g. 5 x 9 = ?) we just 'know' - the answer is stored in long-term memory and the association is recalled directly. Other problems (e.g. 254 x 187 = ?) have not been learned, but most people do know *how* to compute them and, by following the rules of multiplication, the appropriate answers can be produced. In this type of problem the procedural routines applicable to its solution are stored, the numerical product per se is not. These types of sum have been considered (Hunter, 1957) to involve short-term storage of (a) the original problem (if e.g. presented auditorily), (b) the results of interim calculation stages or routines, (c) the particular stage the subject is at in the calculation as a whole. The similarity between the short-term storage involved in digit span tasks and that in mental calculation is illuminated in the following stream of consciousness from Joyce:

> "- Bill, sir? she said, halting. Well, it's seven mornings a pint at twopence is seven twos is a shilling and twopence over and these three mornings a quart at fourpence is three quarts is a shilling and one and two is two and two, sir."

In working out a complex problem the fundamental difficulty is not a lack of number facts, but rather it is trying to remember where we are in the problem and what has been achieved at each stage. As the complexity of the problem increases, the amount of temporary information to be kept track of also increases and this can defeat our short-term memory capacity leading to resort to pencil and paper for external scratch-pad memory.

It is clear from introspection that we use internal speech in keeping track during reconstructive arithmetic problems. If you are not convinced, try doing the following sum under Articulatory Suppression (i.e. whilst repeatedly saying 'the the the ...'): 4798 x 7. It is likely that you find it very difficult, if not impossible, when you are denied the short-term storage facility afforded by internal speech. But you can probably do this sum quite easily under normal conditions. There is also experimental confirmation. Groen & Parkman (1972) demonstrated that 6-7 year old children counted to themselves when doing simple additions - they would start with the larger of the two digits presented (whatever the order of presentation) and increment the counter a number of times equivalent to the smaller of the two digits with a time of about 0.4 seconds per increment, a rate consistent with that of their counting aloud. In adults simple addition has become a highly over-learned skill and the slope of the regression line is now a mere 20 milliseconds, much faster than the rate of internal speech counting which is at best one number every 150 msec (Landauer, 1962; Restle, 1970; Groen & Parkman, 1972; Ashcraft & Fierman, 1982). However, adults still resort to short-term memory and articulatory rehearsal to keep track during more complicated sums as is shown by Sokolov (1972) who recorded the electromyogram activity of the speech musculature of adults performing mental arithmetic and thus revealed their considerable sub-vocal articulatory activity. Sokolov (1972) also demonstrates that suppressing covert articulation by having subjects pronounce irrelevant speech sounds during mental calculation impairs performance.

Lindsay and Norman (1972) and Kahneman (1973) similarly argue that such mental calculations are limited by the need to hold information temporarily in a transient working store and Hitch (1978a,b) demonstrates that forgetting increases during the course of calculation as a function of the number of calculational stages intervening between initial presentation and subsequent utilisation of information.

On the basis of the findings of Ellis & Hennelly (1980) it may be predicted that bilingual differences will also be found in mental calculation tasks which involve short-term storage. If this storage in any way involves articulatory encoding (the level at which word-length effects

are thought to operate, see Baddeley *et al.*, 1975) then bilingual differences in efficiency should be found. We should note, however, that the analysis given here is the normative one. People can choose or learn different ways of doing mental arithmetic. For example learning to use a mental abacus as a calculational tool affects the mathematical competence and digit spans of those who acquire this skill (Hatano, Miyake, & Binks, 1977; Stigler, 1984), and expert calculators have a wide range of idiosyncratic number knowledge and routines (Hunter, 1977).

As bilingual number-name word-length differences show their effects at a short-term storage level, bilingual differences are not expected in associative mental arithmetic problems (e.g. 5 x 9) where little or no short-term storage or manipulation is involved and the answer is directly retrieved from long-term memory. Similarly, bilingual differences are less to be expected in written calculation where the child can (but may not) use the visible page as a permanent working store which provides an efficient substitute for human working storage (Hitch, 1978; Lindsay and Norman, 1972).

The prediction is therefore that Welsh/English number-name word-length differences will result in slower and less accurate calculation in Welsh for problems which involve an appreciable short-term working storage load.

EXPERIMENT 1

Subjects

25 bilingual 'Welsh' and 25 'English' children between the ages 9-12 years were tested individually. The 'Welsh' children were drawn from 3 schools, the 'English' children from 4. The criteria for 'Welsh' and 'English' were (a) attendance at a predominantly Welsh/English speaking school, and (b) the same main language had to be used both at home and at school. The 'Welsh' children performed the mental calculations in their preferred Welsh, the 'English' children in English.

The two groups were matched for age, intelligence as determined using the Deeside intelligence test, and, as far as possible, socio-economic background. All the children were of average or above average intelligence as determined using the intelligence tests. The 'Welsh' children attended schools in Gwynedd, as did some of the 'English' children, the remainder being from Wolverhampton.

Apparatus and procedure

Practice sums and 24 test sums were presented on a Commodore Pet 2001 Personal Computer. The children were individually instructed, in their own language, that they could start a trial by pressing a key, and that, using the numeric keys, they were to type in the answer to the sum which appeared on the screen as soon as they had worked it out. This was to be followed by pressing the 'return' button. The completion times accurate to 1/60s were recorded for each sum, as were the responses.

There were 6 examples of four sum types in the test trials:

Type 1   Simple multiplication   e.g. 5 x 3 =
Type 2   Simple multiple-figure (3, 2) addition   e.g. 305 + 42 =
Type 3   Complex multiple-figure (3, 2) addition with carrying   e.g. 134 + 88 =
Type 4   Multiple figure (9) addition   e.g. 5 + 3+ 7 + 4 + 9 + 8 + 6 + 5 + 3 =

These were presented in standard format with addens aligned vertically. The sum remained displayed throughout the trial. No interim workings (either written or keyed) were allowed,

and the answer was to be input in left to right order (i.e. 305 + 42 requires a '347' response rather than that in the typical order of calculation '743').

It can be seen that sum type 1 is associative (should the child have learnt his tables), sum type 2 requires use of reconstructive strategies with little associated short-term memory involvement, and sum types 3 and 4 require reconstructive strategies with carrying and a considerably greater short-term memory load is incurred.

Results

The response time data were analysed as a 3 factor ANOVA (2 Groups x 4 Sum types x 6 Sums) with subjects nested within groups. The Groups factor ($F(1,48) = 6.75$, $p < 0.05$) demonstrates that on average the 'English' children solved the sums faster than the 'Welsh' children (mean response times 18.9s and 24.1s respectively). The Sum type factor ($F(3,144) = 153.5$, $p < 0.01$) is significant, a Duncan's Multiple Range Test demonstrates that Type 1 sums produced the fastest responses, Type 2 the next fastest, and Types 3 and 4, which did not differ from each other significantly, produced the slowest responses. The most interesting finding is the significance of the Group x Sum type interaction ($F(3,144) = 3.84$, $p < 0.05$). The relevant interaction means can be seen in Table 2. A Duncan's Multiple Range Test shows that whilst the 'Welsh' children do not differ significantly from the 'English' children in the average speed at which they answer Type 1 or Type 2 sums, the 'Welsh' children are significantly ($p < 0.01$) slower at answering Type 3 and 4 sums which involve carrying and many interim stages.

Table 2				
Mean Response Latencies (s) for the 'Welsh' and 'English' Children on the Four types of Sum.				
Sum Type	'Welsh'	'English'	difference	p
1 Simple multiplication	8.77	7.25	1.52	n.s.
2 Simple multiple-figure addition	14.45	12.43	2.02	n.s.
3 Complex multiple-figure addition	35.86	27.73	8.13	p < .01
4 Multiple figure addition	37.15	28.02	9.13	p < .01

Table 3				
Mean Errors out of 6 for the 'Welsh' and 'English' Children on the Four types of Sum.				
Sum Type	'Welsh'	'English'	d	p
1 Simple multiplication	0.24	0.16	0.08	n.s.
2 Simple multiple-figure addition	0.64	0.20	0.44	n.s.
3 Complex multiple-figure addition	1.96	1.56	0.40	n.s.
4 Multiple figure addition	1.60	0.96	0.64	p=.06
Total	4.44	2.88		p< .01

The 'Welsh' children also differed from the 'English' children in that they made significantly more errors (Group means 4.44 and 2.88 respectively). A 3 factor ANOVA

demonstrates a significant Group difference ($F(1,48)=7.2$, $p<.01$), a significant effect of Sum type ($F(3,144)=27.28$, $p<.01$) but the Group by Sum type interaction failed to reach significance ($F(3,144)=0.69$, n.s.). In Table 3 there is a numerical trend whereby there is a greater difference between the Groups on the non-associative sums, and independent samples t tests assessing group differences on each of the Sum types fail to reach significance except in the case of the multiple figure addition sums which was marginally significant at $p=.06$. These data are illustrated in the left-hand graphs of Figure 1.

EXPERIMENT 2

Experiment 1 demonstrates an interaction whereby Welsh children are slower and more error-prone on sums which involve considerable working storage. However, these are also the sums which involve more calculational steps and so, notwithstanding the matching of the children for intelligence and SES, it might be argued that it is calculational complexity rather than the temporary storage demands that underlie these effects.

In order to clarify this issue we therefore ran a second study where any need for calculation was removed and the dependent variable was simply the time to pass through typical interim numerical solutions. We asked a subject to do aloud the mental calculations for the sums in Experiment 1 and we transcribed the interim numbers that he generated. Thus a Type 1 sum '7 x 3' transcribes as '7 3 21' corresponding to '7 times 3 is 21'; a Type 2 sum '204 + 41' as '1 4 5 4 0 4 2 245' corresponding to '1 and 4 is 5, 4 and 0 is 4, 2 answer 245'; a Type 3 sum '688 + 75' as '8 5 13 3 1 7 8 8 16 6 1 6 7 763' corresponding to '8 and 5 is 13, 3 down carry 1 and 7 is 8 and 8 is 16, 6 down carry 1 and 6 is 7, answer 763'; a Type 4 sum '9 + 4 + 3 + 6 + 7 + 4 + 3 + 7 + 6' as '9 4 13 3 16 6 22 7 29 4 33 3 36 7 43 6 49'. The transcriptions for the six sums of each type were written on cards, with one card for each sum type.

Subjects

Three groups of 25 subjects each were used. These were all shop-keepers or assistants in the Bangor-Caernarfon area of North Wales. Their ages ranged from 17 to 67. Two of the groups were bilingual in that they claimed their proficiency in Welsh was better than that in English but both languages had been acquired in childhood. The third group comprised monoglot English speakers.
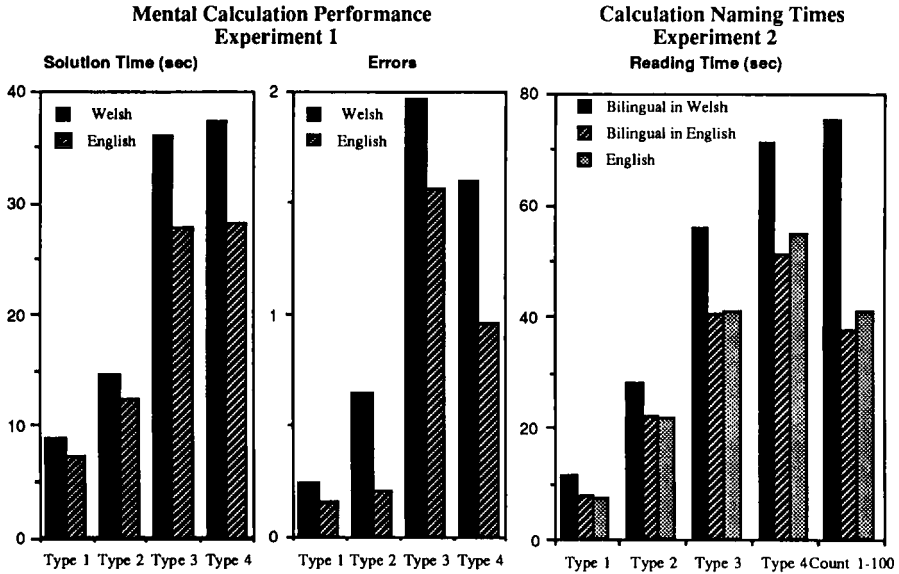
Procedure

The subjects were approached in their shops early in the morning whilst trading was quiet. They were asked to read the numbers on the cards at a comfortable fast rate making as few mistakes as possible. Twenty five of the bilingual subjects were randomly chosen and asked to read the numbers in Welsh and 25 in English. The English group read in English. The subjects were timed for each card and also for the time it took them to count from 1 to 100 as quickly as possible.

Results

The reading times for the 6 sums of each type and the counting times are shown in the right-hand graph of Figure 1 where it can be seen that there is no difference between the bilingual subjects performing in English and the English subjects, but bilingual subjects are slower reading out the interim numbers in Welsh, and this is more pronounced with sum types 3 and 4. This is confirmed by ANOVA with a significant group effect ($F(2,72)=40.3$, $p<.001$), a significant sum type effect ($F(2,288)=567.2$, $p<.001$) and a significant group by sum type interaction ($F(8,288)=25.1$, $p<.001$). Figure 1 also shows the solution times and the error rates for the subjects actually doing these sums in Experiment 1. The scale differences on the

two graphs involving time reflect (i) the differing subjects in the two experiments, and (ii) reading times for Experiment 2 are for 6 sums but involve no calculations.

FIGURE 1

Mental Calculation Performance
Experiment 1

Solution Time (sec)   Errors

Calculation Naming Times
Experiment 2

Reading Time (sec)

There is a clear correspondence between the time it takes here to name the numbers involved in the interim calculations and the actual times taken and the errors made when people to do the sums in the two languages.

Discussion

It is clear that the relative time differences between the languages on the different types of sum are preserved even when we take out the calculation components and simply record the articulation times for the interim 'workings'. The close relationship between these interim 'workings' naming latencies and the actual times taken to do the sums suggest that people do go through these stages of interim calculation and that the latency effects found in Experiment 1 really are a result of the differing articulation rates for numbers in Welsh and English. The fact that the differential error rates on the sums closely parallel these latencies is consistent with the notion of the involvement of a temporally limited STM store. Loss from this store affects computational accuracy and is greater with longer names for the digits involved.

EXPERIMENT 3

The procedure of the first experiment was strictly controlled, and the sum types, with prevention of interim written working, perhaps artificial. To determine whether these results are of any significance in the classroom a field study was undertaken.

Method

74 Primary Schools in Gwynedd and Clwyd were randomly selected from the telephone directory and the head-teachers of these schools were asked whether they would arrange for their 9, 10 and 11 year old children to attempt the sums on a test sheet. This sheet of 60 sums contained 6 examples of the following types of sum: $3 + 4$; 7×6; $9 + 2 + 8 + 4 + 6 + 3$; $754 - 231$; $384 - 197$; 563×2; $84990 + 52529$; $36 + 59 + 42 + 19 + 36 + 54$; $224 \div 4$; 29×24. Those head-teachers who responded were sent the necessary number of forms, and were asked to administer the test ensuring that the children noted both the language(s) of math teaching and the language they used in mental calculation. No other instructions were given. Children were assigned to the Welsh or the English group on the basis of the language used in mental calculation.

As a result, error data for 88 nine year olds, 118 ten year olds and 43 eleven year olds were obtained for each group, i.e. 249 Welsh and 249 English children.

The 'English' children used in this study represented a random sample by age from a considerably larger pool of respondents since it was much more difficult to find children performing the test in Welsh even though the mathematics instruction in the vast majority of the schools was bilingual.

Results

The average number of errors was 9.7 for the 'Welsh' children and 7.0 for the 'English' children (Mann-Whitney U Test, $z = 2.34$, $p < 0.01$, 1 tailed). This difference, although statistically significant, is small, being of the order of 3 sums out of a possible 60.

Discussion

The results of Experiment 1 confirm the speculation that the longer word-length of Welsh digit-names which result in smaller Welsh digit span (Ellis and Hennelly, 1980) also result in relative slowness and increased errors in reconstructive mental calculation in Welsh.

It might be claimed that the procedure of Experiment 1 was artificial in that no written workings were allowed. However, bilingual differences were found at their largest in long addition, sum type 4, where typically no interim workings are used. These sums were designed to represent analogues of the calculations performed traditionally by shop-keepers, albeit the case that the mental solution of this type of problem has now been made unnecessary by the advent of the electronic till.

In a less rigourously controlled environment, with subjects under no time stress and where memory loads are reduced by the use of written scratchpad memory (Experiment 3), bilingual differences can be seen to a lesser extent in the significant difference between the number of sums which the 'Welsh' and 'English' children could attempt successfully. Such differences were also perhaps reflected by the fact that in these bilingual schools surveyed it was considerably easier to find children who perform mental calculation in English. One headmaster, totally unprompted, enclosed a note with the return of the test sheets. This reads: "...I would like to add that they are Welsh speaking and taught in their mother tongue, but one and all prefer to calculate in English. The answer apparently is 'It's easier in English'!"

Although the differences obtained in this study are statistically significant, they are fairly small in Experiment 3 where written calculation is used. These bilingual differences, although interesting, are therefore considered of less significance in the real-world educational setting

where few situations demand calculation without the use of external memory of paper, calculator or fingers.

Relative Welsh/English number name word-length has been shown to affect both the memorability of digit strings (e.g. telephone numbers) (Ellis and Hennelly, 1980), and the ease of mental calculation in these languages. There is no reason to doubt that this effect also operates in other languages. Languages will be more or less conducive to number memorability and calculation, and this will be dependent upon the word-length of the languages' number names.

COUNTING

Figure 1 illustrates the difference in time taken for bilingual individuals to count from 1 to 100 in Welsh and English. It is apparent that this difference is greater than that found for the reading of interim (one or two digit) calculations for sum types 1 through 4. It is likely that this effect is a result of the even greater redundancies in Welsh counting above ten. Thus there is no equivalent to the '-ty' suffix in English, but rather each decade must be expressed as two words (10 *deg*, 20 *dau ddeg*, 30 *tri deg*, 40 *pedwar deg*, etc.), a contrast clearly illustrated by the economy of, e.g. *forty four* in comparison with *pedwar deg pedwar*. Welsh, like other Celtic languages, Breton and Irish, also allows a more traditional, and even longer, form of counting in a system centring on twenties (and, to a lesser extent fifteens). Thus 15 is *pymtheg*, 19 *pedwar ar bymtheg*, 20 *ugain*, 31 *un ar ddeg ar hugain*, 40 *deugain*, etc. Welsh is by no means the only language which uses bases other than ten. Thus in French *vingt* and *quatre-vingts* for 20 and 80 suggest that for some purposes a system with base 20 might have been used. In Danish the word for 70, *halvfirsindstyve* means half way (from three times) to four times twenty. The Eskimos of Greenland, the Tamanas of Venezuela, and the Ainus of Japan are three of the many other people who count by the scores, showing a universal tendency for people to take off their socks as well as their gloves in order to count. Thus for 53, for example, the Greenland Eskimos use the expression *inup pinga-jugsane arkanek-pingasut*, "of the third man, three on the first foot" (Ifrah, 1987). The Babylonian astronomers took a system of notation that was partly sexagesimal (base 60) from their predecessors, the Sumerians, and this is believed to account for the customary division of the hour and the angular degree into 60 minutes (Courant & Robbins, 1941).

There are various processes underlying the counting of the number of objects in an array. Kaufman, Lord, Reese & Volkmann (1949) demonstrate that adults can provide a rapid, confident, and accurate report of the numerosity of small arrays of elements (up to six or seven items) presented for short durations - they name this phenomenon *subitizing*. With larger arrays subjects become increasingly inaccurate unless display times are lengthened to allow actual counting. Logie & Baddeley (1987) show that there is a linear increase in time taken for subjects to count the number of items (between 8 and 25) in arrays, with a slope of approximately 0.32 s per additional item. This latency increase is close to that which one would expect for subvocal counting. Furthermore, the fact that articulatory suppression affects both counting time and accuracy suggests that internal articulation (i.e. subvocalisation of a running total) is required for the accurate counting of arrays of elements. Although we have not directly investigated effects of language digit word-length on subjects' latency of counting the number of elements in an array, we have demonstrated clear differences in the time taken to count from 1->100 in different languages, and this taken together with the Logie & Baddeley findings, provides clear reason to suggest that language digit word-length effects will also operate here.

Hurford (1975) has produced a linguistic theory of numbering systems, distinguishing between a set of primitive numbers (e.g. in English and typically 0 to 9 and units such as the '-ty' in twenty) combined according to base rules to form compound numbers such as 29.

Compound numbers cannot be directly combined - thus 20 + 9 is *twenty-nine*, but 20 + 11 is not *twenty-eleven*. Combining compound numbers such as 20 + 11 involves first their unpacking into primitives (e.g. 20 + 10 + 1) and then repacking them so that the larger leftmost units take on the largest possible value. Thus *twenty-eleven* is ill-formed but *thirty-one* is not. One problem of the English counting system is the idiosyncratic way in which names are formed for numbers in the teens. Number names from 20 to 99 are formed by suffixing a decade name with a unit value. But number names such as *eleven* and *twelve* do not preserve the decade name, and the later teens must be addressed by a 'Switch' rule (Hurford, 1975) whereby the unit value precedes the decade name, with the decade name being the special term *teen*. The reported examples of non-standard numbers being produced by American children consist of the improper concatenation of legitimate number names, and Miller & Stigler (1987) demonstrate that American children have a particular problem with the non-standard teens (e.g. treating them as primitives and generating non-standard numbers like *forty-twelve* for 52) whereas Chinese children never make such errors, Mandarin having no equivalent to the teens, but rather forming these by the regular compounding of a decade value with a unit value. Chinese also lacks the special (and slightly confusing) decade names such as twenty, decades are the compound of the value and *ten* (*twenty seven* is, in effect, *two-ten-seven : er shi qi*).

Here again we see clear effects of linguistic legacy on mental arithmetic. The transparent Chinese counting system makes it easier for Chinese children to induce the difference between primitive and compound numbers. This induction is more difficult in English, and children frequently show confusion over what is or is not a primitive number that can be combined in forming compound numbers. Thus there are large country differences generally favouring Chinese over American children in counting (Miller & Stigler, 1987).

More important is the realisation that counting is the entrée to mathematics. The child's ability to reason arithmetically rests on their representations of numerosity. Developmental investigations make it clear that the young child obtains such representations of by counting - "the judgement of equivalence or order, the application of the operations of addition, subtraction, and identity, and the process of solving all depend on counting" (Gelman & Gallistel, 1978; p. 244; Miller & Gelman, 1983). Furthermore, counting provides an important source of feedback for the learning of arithmetical relationships (Siegler & Robinson, 1982). This view is confirmed by studies of children with specific arithmetical learning disability (ALD). ALD children have a specific working memory deficit in relation to processing numerical information - they are particularly poor at working memory tasks involving counting but not those involving more general language processing (Siegel & Ryan, 1989). Performance on working memory tasks which involve counting increases as a function of speed of counting, and asking adults to count in an unfamiliar language causes a drop in their performance on counting-working-memory tasks to levels of 6 years old children (Case, Kurland & Goldberg, 1982). In line with this, Hitch & McAuley (in press) demonstrate that it is impaired counting that affects the acquisition of arithmetical skills as well as their execution in ALD children.

The counting systems of a language, and the names that it has for its numbers, affect the ease of counting in that language. This, in turn, affects the development of arithmetical competence in its speakers.

MORE GENERAL ASPECTS OF LANGUAGES' WORD-LENGTHS

Languages differ on many dimensions relevant to efficient communication. Cherry (1966) states "The relationship between the whole structure of a language (the morphemic, syntactic, grammatical formalism) and the outside world associations (its semantic functioning) is extremely complicated; it is essentially empirical and, above all, varies between different

languages. Again, redundancy is built into the structural forms of different languages in diverse ways. No general laws exist."

In considering word-length we are addressing just one of the factors contributing to the redundancy of a language. Within this limited area, however, there do appear general laws. Cherry (1966) observes "... under the natural stress of human economising; the most frequently used words are the shortest; when a word comes into frequent and popular use we tend to abbreviate it (UNESCO, NATO, gas)." Zipf (1935) formulated the law of abbreviation: whenever a long word or phrase suddenly becomes common, we tend to shorten it. Similarly Miller and Newman (1958) suggest that an evolutionary process of selection has been working in favour of short words and demonstrated empirically that the average frequency of words of *i* letters in length is a reciprocal function of their average rank with respect to increasing length.

If the value of some redundancy in allowing detection of errors in transmission or reception under non-optimal conditions is for the present ignored (cf. e.g. van Amerongen, 1975), then it can be seen that having a minimum-redundancy code is desirable. Miller and Chomsky (1963) demonstrate this from an economic viewpoint: there is a cost to communication and the average length of the message is an appropriate measure of this cost since it takes either more time or more equipment to transmit more symbols. In a given period of time, more information can be transmitted using a low redundancy short code. High frequency words, by definition, are those commonly used for communication. It is these which have apparently evolved to be of short word-length and low redundancy.

If languages are compared for average word-length, moreover, gross differences can be seen. Fuchs (1968) devised the following mathematical relationship for the mean frequency distribution h_i of i-syllabic words when the mean number of syllables is $\bar{i}$:

$$h_i = \frac{(\bar{i} - 1)^{i-1}}{(i - 1)!} \ e^{-(\bar{i}-1)}$$

This relationship appears valid for all languages. The only criterion that varies from one language to another is the value $\bar{i}$ for the average number of syllables. For example, in nine languages investigated by Fuchs, this value was as small as 1.41 for English, ranging through 2.10 for Arabic and Greek, and as large as 2.46 for Turkish. The English language appeared to contain the highest proportion of monosyllabic words i.e. it requires the least number of syllables to convey a given amount of information, and proponents of English as a *lingua franca* have suggested (van Amerongen, 1975) that this may be one reason why English has to so great an extent become adopted as an international language.

These considerations of language efficiency as a function of word-length have arisen from the viewpoint of interpersonal communication. In addition, however, word-length has been shown to affect a number of functions involved in intrapersonal information processing and manipulation, e.g.. reading and short term memory span (Baddeley *et al.* 1975) and mental calculation (Ellis and Hennelly, 1980 and the experiments reported here).

Word-length effects operate at an articulatory encoding level. It must therefore be concluded that efficiency at any task which involves articulation will to some extent be a function of the length of the words to be so encoded. The generation of speech is the most obvious example of such a task. In addition, Kleiman (1975) speculates that any condition in which information enters the system more rapidly than it can be semantically processed may cause the subject to use articulatory encoding as a back-up store. Given the large cross-lingual differences in average word-length it can be concluded that, as a function of word-length, languages differ in the efficiency at which they can be used to communicate and manipulate information.

CONCLUSIONS AND BROADER CONSIDERATIONS

We have demonstrated a range of effects of languages' numbering systems, their word length and their transparencies in forming number names from primitives, that affect the facility of speakers of that language in remembering numbers, in counting, and in mental calculation.

There are very large national and cultural differences in mathematics ability (Husén, 1967; Stevenson, Lee & Stigler, 1986; Stigler, Lee & Stevenson, 1987). Thus American children lag behind Japanese and Chinese children in mathematics ability (Stevenson et al., 1986) and Israeli children clearly surpass English who in turn are better than Swedish children at 13 years old (Husén, 1967). The linguistic relativity effects reviewed here play but one role in determining these. We must remember that counting and mental arithmetic are merely the portals of mathematics and play little role in the abstractions of algebra, geometry, sets, calculus, proofs, logic, ..., and mathematical creativity. Also important in determining this skill-base in the populace are the children's schooling, the attitudes of their parents and their culture towards mathematics, the involvement of parents and children in school-work, teacher training and competence, and the child's expectations and aspirations (Husén, 1967; Stevenson et al., 1986; Stigler et al., 1987).

ACKNOWLEDGEMENTS

REFERENCES

Ashcraft, M.H. & Fierman, B.A. (1982) Mental addition in third, fourth, and sixth graders. *Journal of Experimental Child Psychology*, 33, 216-234.

Baddeley, A.D. (1986) *Working Memory*. Oxford: Clarendon Press.

Baddeley, A.D. (1990) *Human Memory: Theory and Practice*. Hove and London: Lawrence Erlbaum Associates.

Baddeley, A. D. & Hitch, G. J. (1974). Working memory. In G. A. Bower (ed.), *The Psychology of Learning and Motivation: Advances in Research and Theory*, 8, pp. 47-89. New York: Academic Press.

Baddeley, A.D., Thomson, N. & Buchanan, M. (1975) Word-length and the structure of short-term memory. *Journal of Verbal Learning and Verbal Behaviour*, 14, 575-589.

Baker, C. (1985) *Aspects of Bilingualism in Wales*. Clevedon: Multilingual Matters.

Burt, C. (1939). *Mental and Scholastic Tests*. London: King.

Case, R., Kurland, D.M., & Goldberg, J. (1982) Operational efficiency and the growth of short-term memory span. *Journal of Experimental Child Psychology*, 33, 386-404.

Cherry, C. (1966) *On Human Communication*. Cambridge, Mass.: M.I.T. Press. 2nd Edition.

Cooley, C.H. (1962) *Social Organisation*. New York: Schoken Books.

Courant, R. & Robbins, H. (1941) *What is Mathematics?* Oxford: Oxford University Press.

Dantzig, T. (1930) *Number: The Language of Science*. New York: Macmillan Company.

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.

Fuchs, W. (1968) *Nach allen Regeln der Kunst*. Stuttgart: Deutsche Verlags-Anstalt.

Gelman, R. & Gallistel, C.R. (1978) *The Child's Understanding of Number*. Cambridge, Mass.: Harvard University Press.

Groen, G.J. & Parkman, J.M. (1972) A chronometric analysis of simple addition. *Psychological Review*, **79**, 329-343.

Hatano, G., Miyake, Y. & Binks, M.G. (1977) Performance of expert abacus operators. *Cognition*, **5**, 47-55.

Hitch, G.J. (1978a) Mental arithmetic: short-term storage and information processing in a cognitive skill. In A.M. Lesgold, J.W. Pellegrino, S.D. Fokkema & R. Glaser (Eds.) *Cognitive Psychology and Instruction*. New York: Plenum Publishing Corporation.

Hitch, G.J. (1978b) The role of short-term working memory in mental arithmetic. *Cognitive Psychology*, **10**, 302-323.

Hitch, G.J. & McAuley, E. (In press). Working memory in children with specific arithmetical difficulties. *British Journal of Psychology*. In press.

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

Hoosain, R. (1986) Language, orthography and cognitive processes: Chinese perspectives for the Sapir-Whorf hypothesis. *International Journal of Behavioural Development*, **9**, 507-525.

Hoosain, R. (1987) Correlation between pronunciation speed and digit span size. *Perceptual and Motor Skills*, **55**, 1128.

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

Hunter, I.M.L. (1964) *Memory*. Harmondsworth: Penguin.

Hunter, I.M.L. (1977) Mental calculation. In P.N. Johnson-Laird & P.C. Wason (Eds.) *Thinking: Readings in Cognitive Science*. Cambridge: Cambridge University Press.

Hurford, J.R. (1975) *The Linguistic Theory of Numerals*. Cambridge: Cambridge University Press.

Husén, T. (1967) *International Study of Achievement in Mathematics*. Volumes I and II. New York: John Wiley & Sons.

Ifrah, G. (1987) *From One to Zero: A Universal History of Numbers*. New York: Viking Penguin.

Joyce, J. (1969) *Ulysses*. Harmondsworth: Penguin.

Kahneman, D. (1973) *Attention and Effort*. Englewood Cliffs, N.J.: Prentice-Hall.

Kaufman, E.L., Lord, M.W., Reese, T.W. & Volkman, J. (1949) The discrimination of visual number. *American Journal of Psychology, 62*, 498-525.

Kleiman, G.M. (1975) Speech Recoding in Reading. *Journal of Verbal Learning and Verbal Behaviour, 14*, 323-339.

Landauer, T.K. (1962) Rate of implicit speech. *Perceptual and Motor Skills, 15*, 646.

Lindsay, P.H. & Norman, D.A. (1972) *Human information processing: An introduction to psychology*. New York: Academic Press.

Liu, K.S. & Shen, H.M. (1977) Effects of language structure and intertrial interval on the speed of silent reading. *Acta Psychologica Taiwanica, 19*, 57-59.

Logie, R.H. & Baddeley. A.D. (1987) Cognitive processes in counting. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 13*, 310-326.

Miller, G.A. (1956) The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review, 63*, 81-97.

Miller, G.A. (1965) Introduction to Zipf, G.K. *The Psycho-Biology of Language*. Cambridge, Mass: M.I.T. Press.

Miller, G.A. & Chomsky, N. (1963) Finitary Models of Language Users. In Luce, R.D., Bush, R.R. and Galanter, E. (eds) *Handbook of Mathematical Psychology*, Volume II. New York: John Wiley.

Miller, G.A. & Newman, E.B. (1958) Tests of a statistical explanation of the rank- frequency relation for words in written English. *American Journal of Psychology, 71*, 209-258.

Miller, K. & Gelman, R. (1983) The child's representation of number: a multidimensional scaling analysis. *Child Development, 54*, 1470-1479.

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

Naveh-Benjamin, M. & Ayres, T.J. (1986) Digit span, reading rate, and linguistic relativity. *Quarterly Journal of Experimental Psychology, 38*, 739-751.

Restle, F. (1970) Speed of adding and comparing numbers. *Journal of Experimental Psychology, 83*, 274-278.

Spier, L. (Ed.) (1941) *Language, Culture and Personality: Essays in Memory of Edward Sapir*. Menasha, Wis.: Sapir Memorial Publication Fund.

Siegel, L.S. & Ryan, E.B. (1989) The development of working memory in normally achieving and subtypes of learning disabled children. *Child Development, 60*, 973-980.

Siegler, R.S. & Robinson, M. (1982) The development of numerical understandings. In H. Reese & L.P. Lipsitt (Eds.) *Advances in Child Development and Behaviour, Vol. 16*. New York: Academic Press.

Sokolov, A.N. (1972) *Inner Speech and Thought*. Translated by G. T. Onischenko. New York: Plenum Publishing Co.

Stevenson, H.W., Lee, S.-Y., & Stigler, J.W. (1986) Mathematics achievement of Chinese, Japanese, and American children. *Science, 231*, 693-699.

Stigler, J.W. (1984) "Mental abacus": The effects of abacus training on Chinese children's mental calculation. *Cognitive Psychology, 16*, 145-176.

Stigler, J.W., Lee, S.-Y. & Stevenson, H.W. (1986) Digit memory in Chinese and English: Evidence for a temporally limited store. *Cognition, 23*, 1-20.

Stigler, J.W., Lee, S.-Y. & Stevenson, H.W. (1987) Mathematics classrooms in Japan, Taiwan, and the United States. *Child Development, 58*, 1272-1285.

Terman, L. M. & Merrill, M. A. (1960). *Stanford-Binet Intelligence Scale: Manual for the Third Revision*. Boston: Houghton Mifflin.

Van Amerongen, C. (1975) *The Way Things Work Book of the Computer*. London: George Allen and Unwin Ltd.

Wechsler, D. (1949). *Manual for the Wechsler Intelligence Scale for Children*. New York: The Psychological Corporation.

Whorf, B.J. (1942) Language, Mind and Reality. January and April Issues, *The Theosophist*.

Whorf, B.J. (1956) *Language, Thought and Reality. Selected Writings of Benjamin Lee Whorf*. New York: John Wiley.

William, U. & Roberts, G. (1972). *The Welsh Children's Intelligence Scale*. Windsor: NFER.

Zipf, G.K. (1949) *Human Behaviour and the Principle of Least Effort*. Cambridge, Mass: Addison-Wesley.