Why are the favourite numbers favourite? Regularities of the sizes of gifts donated to charity organisations by postal cheque, and some implications.

István DIENES
Homputer
homputers@yahoo.com

Abstract

Sizes of 440 thousand gifts was studied, which have been given by 140 thousand donors in the past decade to two charity organizations. In 99,95% of the cases, donors chose one of 199 favourite numbers. Donors both in poor and well-to-do neighbourhoods sticked to the same amounts, despite of an inflation of more than 100% in the period, while only the chance, with which a favourite was chosen was subject to change. In various years, 40-99% of the favourites proved to be a round number of 1—4 valuable digits in the decimal, 1-2-5 and 1-2-3-5 systems. More than 95% of the amounts can be paid by less than three banknotes. The favourites differ from Dehaene's referents and the referents can not easily be redefined by favourites. The frequency distrinbution of gift sizes is sparse and is different from the distributions that have been applied earlier in psychology of numbers. The difference can be explained by the situation, in which donors generate the size of the gift. Frequency distribution can not be approximated by either standard probability distributions or referents deduced from standard distributions. However, the data can be described by a Markovian model, in which donors generate the digits of gift sizes sequentially from the left in the 1-2-5 system, always in dependence from the precedent digit. The results indicate that donors adopt a „paying" situation with banknotes rather, than assuming an abstract numerosity or number along a hypothetical number line. Thin-structure of the frequency distribution of gift sizes and extended Koch-Crick principle allows to hypothesize a representation and mechanism from which digits of gift-sizes emerge.

Introduction

On behalf of its clients, CID Cég-INFO Ltd., as a direct marketing agency, regularly has launched charity campaigns. After several campaigns a data processing plan was designed so as to understand the behaviour of donors and to maximize rate of the investment into the campaigns. As soon as the preliminary analysis proved that favourites are independent of sex, age income of neighbourhood of the donor and the date of donation, we understood that those are cognitive processes that define gift size. Hence we tried to create a model, which is able to reproduce the set of favourites and frequency distribution of gift sizes. Sociodemographic background and motives of donors was surveyed by phone and questionnaire techniques. Last but not least an attempt was made to define functional units in the brain, a representation, whose operation the phenomena experienced can be explained in accordance with known facts of psychology of numbers.

Method

Subjects

Subjects were recruited from more than 2 million adults who were mailed in Hungary by CID in the period 1993-2002. Prior to data processing the database was anonimized. The majority of donors were female, of age 35—70 at the moment of mailing. Donors can not be viewed as a sample, which would represent Hungarian population, beacuse they constitute a specific subset. Our approach is observational rather than experimental, because even if the situation of subjects and the results are clearly defined and reproducible, the situation can be influenced to a limited extent only.

Data

The study extended to 440 thousand gifts which were given by 140 thousand subjects to two charity organizations. The exact number of donors is not known, because this can not eb identified from the anonymized database. The items „Amount sent", „ZIP code" and „Date" were recorded from the postal cheque forms.

Terminology, methods of data processing

Digits of gift sizes were double indexed. The first index characterized the decimal position of the digit when starting from the right (Ones). The second index identified the position of the digit from the left, starting from the first valuable decimal digit. The digit of the number x , indexed by ij was denoted by p_{ij} and, absolute frequency by $F_{i.k} = F(p_{i.}(x) = k)$ módon. A point as an index stands for summation taken for all values of the index.

Table1. Illustration of the terminology: The number *„13 657"*

	Place of millions	Place of Hundred Thousands	Place of Ten Thousands	Place of Thousands	Place of Hundreds	Place of Tens	Place of Ones	Numeral	Elements of Numeral	Morphemes of Numeral
Name of the digit wehn counting from the left			First valuable decimal of the example	Second valuable decimal of the sample			Last decimal in the example			
Naem of the digit when counting from the right	Highest decimal in the study		Highest decimal in the example				Lowest decimal in the example			
Digits of the example	0	0	1	3	6	5	7	Tizenháromezer-hatszázötven hét	Tizenhárom , ezer, hatszáz, ötven, hét	Tíz, en, három, ezer, hat, száz, öt, , ven, hét

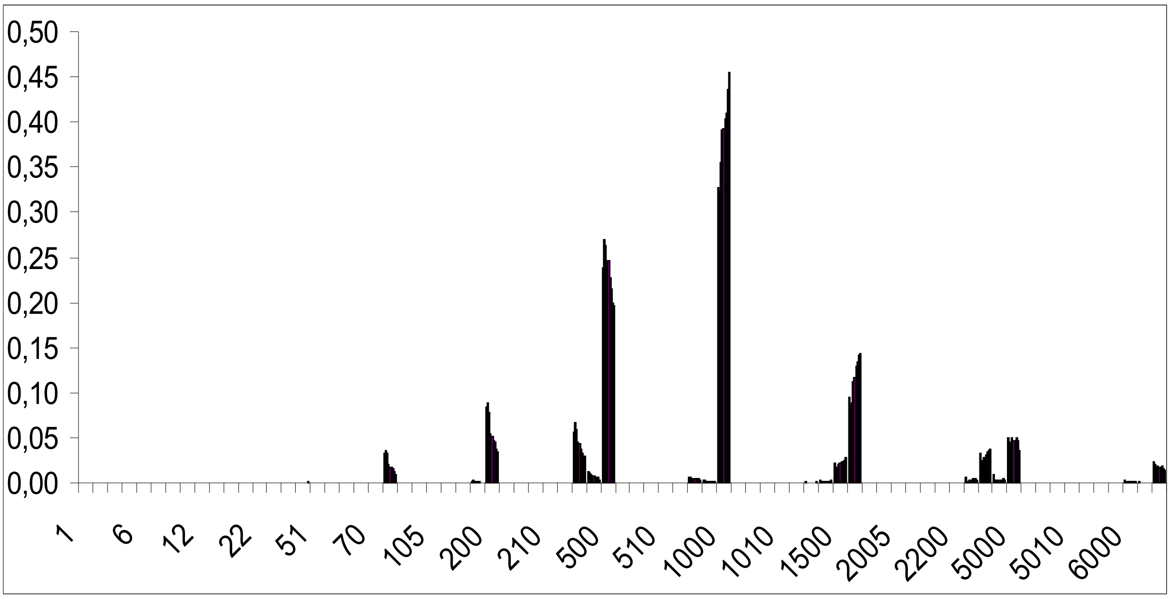
1, 3, 6, 5, 7 are the digits of decimal 13 657, of which 1 is the first and highest and 7 is the last and lowest. Digit 3 in 13 657-ben is denoted by p_{24} (13 657). Mutatis mutandi the same terminology will be used later for non-decimal numbers. The preprocessing was made in a standard ACCESS database, after which standard hypothesis testing and variance analyses were performed, even if almost everything is significant due to the extraordinary size of the sample.

Results

Reproduction of the set of favourites

Altogether 1 537 numbers were identified as size of at least one of 446 870 gifts. This makes 0,03 % of 5 119 600 integers between greatest and lowest gift size. Of the 1 537 integers 199, 13,0% of integers occurring and 0,0039% of all integers between the greatest and smallest gift size have been identified as a favourite. The gifts of size of a favourite, give 99,55% of all gifts.

Figure 1. *Frequency distribution of gifts by size and yearof donation Gifts greater than HUF 6000 have been omitted.*



A gift size have been defined as a favourite if it occurred mroe frequently, than its left and right neighbours together. Annually less and less people gave gifts below HUF 1 000 and more and more pople gave gifts over HUF 500, but favourites themselves remained as favourites.

We are going to reconstruct the set of gif sizes and favourites, so below first we shall characterize gift sizes by the number of decimal and valuable decimal digits they consist of, the number of elements and morphemes in the numeral and number of banknotes needed to pay the amount.

Though most favourite is a round decimal, favourite numbers are rare among round decimals.

Table 2. *Number of decimals and valuiable decimals in gift sizes.*

Number of decimals (Length of the number)	Number of valuable decimals/All decimals	Number of valuable decimals 1	2	3	4	5	6	7	Total
1	1,00	5	0	0	0	0	0	0	5
2	1,11	180	22	0	0	0	0	0	202
3	1,03	156 228	1 976	1 041	0	0	0	0	159 245
4	1,06	262 392	13 369	303	468	0	0	0	276 532
5	1,10	9 788	838	35	25	14	0	0	10 700
6	1,56	122	35	5	2	4	6	0	174
7	3,50	2	1	1	1	2	0	1	8
Relative frequency									
1		100,0%							
2		89,1%	10,9%						
3		98,1%	1,2%	0,7%					
4		94,9%	4,8%	0,1%	0,2%				
5		91,5%	7,8%	0,3%	0,2%	0,1%			
6		70,1%	20,1%	2,9%	1,1%	2,3%	3,4%		
7		25,0%	12,5%	12,5%	12,5%	25,0%	0,0%	12,5%	

Hungarian compound numerals consist of elementary numerals and operational morphemes which refer to the mental explication of the number according to an Euclidean algorithm. The superessival „–on/-en/-an" (English preposition „on"), possessive „-ú/-ű": negyven=négy-ű-en (40 = on the one, which has four) manifest in the compound numeral, multiplicative „–szor/-szer/-ször” (English „times”) and the operator „meg” (= plus) are not marked.

Table 3. *Favourites and non-favourites among gift sizes whose numeral consists of many and few morphemes.*

Number of different gift sizes				Number of gifts			
Number of morphemes in numeral	Number of non-favourites	Number of favourites	Number of favourites/Number of all different gift sizes %	Number of morphemes in numeral	Number of non-favourites	Number of favourites	Number of favourites/Number of all different gift sizes %
Total	1 334	199	13%	Összesen	2021	444 870	100%
1	2	4	67%	1	5	181 202	100%
3	19	29	60%	3	27	246 799	100%
4	2	3	60%	4	5	329	99%
5	123	52	30%	5	222	12 850	98%
6	23	4	15%	6	45	27	38%
7	247	67	21%	7	435	3 380	89%
8	59	6	9%	8	110	28	20%
9	399	29	7%	9	651	201	24%
10	53	1	2%	10	60	17	22%
11	223	4	2%	11	268	12	4%
12	36	0	0%	12	39	0	0%
13	116	0	0%	13	122	0	0%
14	3	0	0%	14	3	0	0%
15	13	0	0%	15	13	0	0%
16	1	0	0%	16	1	0	0%
17	2	0	0%	17	2	0	0%
18	4	0	0%	18	4	0	0%
19	3	0	0%	19	3	0	0%
20	1	0	0%	20	1	0	0%
21	4	0	0%	21	4	0	0%
24	1	0	0%	24	1	0	0%

The numeral of the frequent gift sizes consist of few morphemes, however, the number of morphemes does not determine the frequency of a gift size, since as well 50, which occurred 150 times, as 60, which occured only three times, contains three morphemes.

Dehaene and Mechler 1992 hypothesize that an „analog” representation plays the central role in manipulations with numbers Analog values would manifest through „referents” which are eager to manifest and which have a „span”. Their definition for „span” has been built on the concept of rounding, but is not strict: For instance, 526 acn be rounded to 500, 530 or 600. Neither give these authors an explanation, why are diameters of spans different, e.g. span of 15 is much greater than those of 14 and 16. The span of 10 covers 9 and 11 but this contradicts to usual rules of rounding.

It would be more straughtforward to call Y the span of the number X, if and only if Y is an interval whose numbers manifest as X with a probability (density) p(y) greater than a fixed k for each y inside the Y. Dehaene does not succeed in reconstruction of spans: 10, 12, 15 and 20 are more frequent than it would follow from the model. This theory does not explain the scarcity of minor gifts, dependence of digits, which will be discussed later and even the frequency of gift sizes is not monotone decreasing. Furthermore, Dehaenenan referents are definitely different in each decimal order of magnitude, while the valuable digits of favourites are almost the same through Hundreds, Thousands, Ten Thousand and Hundred Thousands. Favourites must not be equated with referents.

One also might hypothesize that the „intended” gift sizes have a smooth f(y), e.g. Gaussian distribution, there are referents indeed, but Dehaene did not succeed in identifying them, and they are the favourite numbers indeed. Then each favourite should have a frequency distribution $p_i(y)$, and f(y) could be reproduced. However, we could not produce a smooth bell-like f(y) by standard methods and probabibility distributions, so the hypothesis was rejected. On the other hand, the lack of any provable relation of favourites to referents raises doubts concerning the psychological reality of referents in the situation of gift-giving.

In various years, 85,7-87,3% of the amounts given by donors of neighbourhoods of various residential income, can be paid by one benknote or coin, 2,2-2,8% by two banklnotes or coins. Similar conclusion can be drawn from the Barna Research Group's data for the U.S. for 1995 in Table 4.. Frequency distribution of gift sizes is similar to those in Hungary.

Table 4. *Distribution of gift-sizes in the United States in 1995. Assuming 1 USD= 200 HUF*

Gift size (US)	Currency	Gift-size in Hungarian currency (HUF)	Number of denominations in the US	Relative frequency of the gift size in the US in 1995
1	cents	2	1	0
5	cents	10	1	0
10	cents	20	1	0
25	cents	50	1	0
1	dollars	200	1	10%
2	dollars	400	1	0,8%
5	dollars	1 000	1	12%
10	dollars	2 000	1	13,5%
15	dollars	3 000	2	15%
20	dollars	4 000	1	10,8%
25	dollars	5 000	1	20%
30	dollars	6 000	2	1,5%
35	dollars	7 000	3	0,5%
40	dollars	8 000	2	0,5%
45	dollars	9 000	2	0,3%
50	dollars	10 000	1	7%
75	dollars	15 000	3	0,8%
100	dollars	20 000	1	4%
150	dollars	30 000	2	0,8%
200 or more	dollars	40 000 or more		2%

The amounts that can be given by paying two banknotes with receiving back one or by other two-way transactions are not favourites, i.e. the amounts of form x + y are more frequent, than the amounts of form x – y. Surprisingly, no all amounts of shape x + y are frequent. HUF 300 = 100 + 200 is frequent, but HUF 400 = 200 + 200 is very sparse, and HUF 600 = 500 + 100 is also sparse even though is more frequent, then HUF 400. HUF 2500 = 2000 + 500 is a favourite, but 2 200 = 2000 + 200, which stands nearer to the average HUF 2 100 is rare.

In the decimal system every integer N can be written in the following form:

$$N = k_n*10^n + \cdots + k_1*10^1 + k_0*10^0 \quad (1).$$

The value of coefficients k_n can be calculated by the Euclidean algorithm, and the numbers 10^n are the bases of the decimal system. Steps of the Euclidean algorithm are as follows:

(i) Identify the greatest base of form 10^n , that is smaller than N

(ii) Identify the greatest coefficient k_n such, that $k_n*10^n < N$.

(iii) Calculate residuum m_n

(iv) The procedure will be repeated with residuum m_n and base n'=n-1

In Euclidean representation each integer will be represented by *exact* magnitudes and coefficients. Damian concluded that digits gain rapid access to exact numerical magnitude representation, but access is slower to lexical codes.

Both numerals and decimal Arabic numbers can be deduced from the Euclidean algorithm. When generating numerals, the small bases of round numbers as well as bases and coefficients of number-inner zero digits are not mentioned. When generating decimal Arab numbers, bases are not written.

The same Euclidean algorithm can be adopted to define the $N = 2^n + \ldots + 2^1 + 2^0$ binary form of a number. Keeping in mind that favourites are similar in different decimal magnitudes, we recalculated the gift sizes in various systems of which the favourites are the base: the systems 1-2, 1-2-5 and 1-2-3-5. Such systems can be defined by taking 1,2,3 and 5 and their decimal multiples as bases as follows:

$$N^{1\text{-}2} \doteq \Sigma_m (L_m*2 + N_m*1) *10^m,$$

$$N^{1\text{-}2\text{-}5} \doteq \Sigma_m (K_m*5 + L_m*2 + N_m*1) *10^m$$

$$N^{1\text{-}2\text{-}3\text{-}5} \doteq \Sigma_m (K_m*5 + L_m*3 + M_m*2 + N_m*1) *10^m.$$

These systems differ from simple binary, three or five based number systems, their legal coefficients (digits) are shown in Table 5.

Table 5. *Number and sum of coefficients of values of 440 thousand gifts reproduced in various systems of numbers.*

System	Coefficients allowed (digits)	Number of bases neded to reproduce all gift sizes	Sum of all coefficients of all gifts	Number of occurrences of bases altogether (Number of non-zero coefficients)	Average of the non-zero coefficients
1	0, 1, 2, 3, 4, 5, 6, 7, 8, 9	7	1 222 475	467 476	2,62
1 – 2	0, 1, 2, 3, 4	15	798 452	639 261	1,25
1 – 3	0, 1, 2, 3	15	840 643	612 914	1,37
1 – 5	0, 1, 2, 3, 4	15	648 783	474 986	1,37
Denominaions (incomplete 1-2-5 system)	0, 1, 2	14	519 204	510 286	1,02
1 – 2 – 5	0, 1, 2	21	516 629	510 031	1,01
1 – 2 – 3 - 5	0, 1	27	481 584	481 584	1,00

Table 6. Correctly classified gifts when using the number of morphemes, round numbers in the system 1, 1-2, 1-5, 1-2-5, 1-2-3-5 and number of denominations,

K =	F: Number of morphemes	F: System 1-2	F: Number of denominations	F: System 1-2-5	F: System 1-2-3-5	F: System 1-5	F: System 1
1	40,55%	61,46%	87,73%	87,80%	93,82%	94,93%	95,22%
2	95,78%	96,30%	99,00%	99,00%	99,27%	98,99%	99,57%
3	98,78%	99,50%	99,55%	99,55%	99,58%	99,65%	99,88%
4	99,65%	99,78%	99,73%	99,73%	99,75%	99,82%	99,99%

System F:1, the standard decimal system provides less errors of first and second kind, but this is the system that goes with the largest computing workload.

Reproduction of the distribution of gift sizes

First the distribution of favourites was compared to Benford's (1938), Banks és Hill's (1974), Baird és Noma's (1975), Dehaene and Mehler's (1992), and Dehaene and Marques' (2002) and standard statistical distributions but they differed so much that we had to reject them. Hence we attempted to estimate the frequency of gift sizes from the frequency of their decimal digits.

We have found that the distribution of gift sizes can be reproduced even as a simple product of overall frequencies of decimal digits-constituents $f_{ij}^{k} = f_{..}^{k}$, even though the fit is poor, but the distribution of the product of frequencies of digits is sparse and the most outstanding maxima can be observed.

The frequency of digits in various from-the-right, from-the-left and double-indexed positions. This is illustrated in Table 7 with data for the decimal digits in the first from-the-left (highest) positions..

Table 7. *Frequency distribution of decimal digits in the first-from-left (highest) decimal positions*

	Highest decimal position							
Digit in decimal position k	Tens $f_{.2k}$	Hundreds $f_{.3k}$	Thousands $f_{.4k}$	Ten Thousands $f_{.5k}$	Hundred Thousands $f_{.6k}$	Millions $f_{.7k}$	Total	Benford's law
1	2%	6%	67%	78%	70%	88%	45%	30%
2	3%	15%	20%	15%	11%	0%	18%	18%
3	4%	11%	5%	3%	8%	0%	7%	12%
4	1%	2%	1%	1%	2%	0%	1%	10%
5	75%	63%	7%	3%	6%	13%	27%	8%
6	4%	1%	0%	0%	1%	0%	1%	7%
7	6%	1%	0%	0%	1%	0%	0%	6%
8	1%	1%	0%	0%	1%	0%	0%	5%
9	1%	0%	0%	0%	0%	0%	0%	5%

For this reason the distribution of gift sizes can be reproduced by the product of decimals $f_{ij}^{k} = f_{.j}^{k}$ or $f_{ij}^{k} = f_{i.}^{k}$. Digits of little value are sparser in high positions, as it can be seen in Table 8.

Table 8. *Frequency $f_{.j}^{k}$ of digits k by decimal positions Position from-the-left was not considered but number-initial zeros were included in the statistics.*

Digit in decimal position: k	Ones $f_{.1}^{k}$	Tens $f_{.2}^{k}$	Hundreds $f_{.3}^{k}$	Thousands $f_{.4}^{k}$	Ten thousands $f_{.5}^{k}$
No digit in the position (number initial zeroes	0,0%	0,0%	0,0%	35,7%	97,6%
0	99,6%	99,1%	61,2%	2,2%	0,0%
1	0,0%	0,1%	2,1%	41,3%	1,9%
2	0,1%	0,1%	5,6%	12,3%	0,3%
3	0,0%	0,1%	4,2%	3,1%	0,1%
4	0,0%	0,1%	0,8%	0,4%	0,0%
5	0,1%	0,4%	25,1%	4,8%	0,1%
6	0,0%	0,1%	0,5%	0,1%	0,0%
7	0,0%	0,1%	0,3%	0,0%	0,0%
8	0,0%	0,1%	0,2%	0,0%	0,0%
9	0,0%	0,0%	0,0%	0,0%	0,0%

Table 9. and similar other tables suggest that valuable digits tend to group: if there is another non-zero digit somewhere after a non-zero digit, it follows immediately and mostly they are not separated by zeroes. This suggests that dependence of digits may be a critical factor in the generation of a gift size. Are the digits of gist sizes independent and if dependent, is this dependence a (from-the-left) forward or (from-the-right) forward dependence?

Table 9. *Freuency of gift sizes starting with 1, containing two zeroes and consisiting of four digits*

Gift size (HUF)	Absolute frequency	Gift size (HUF	Absolute frequency	Gift size (HUF	Absolute frequency
1 001	2	1 010	0	1 100	139
1 002	0	1 020	3	1 200	644
1 003	3	1 030	4	1 300	175
1 004	0	1 040	1	1 400	97
1 005	0	1 050	21	1 500	9 660
1 006	1	1 060	2	1 600	101
1 007	1	1 070	0	1 700	78
1 008	3	1 080	0	1 800	56
1 009	1	1 090	2	1 900	17

Let P (A) the probability measure of an event A. The event A is stochastically independent from an event B if and only if

$$P (A \mid B) = P (A). \quad \text{(i)}$$

Events A and B are mutually independent if and only if

$$P (A B) = P (A) * P (B). \quad \text{(ii)}$$

When studying the independence of digits in a position the following events will be investigated: A = $p_{i.}^{k}$, or $p_{.j}^{k}$, or p_{ij}^{k}.

The *independence of positions* can also be studied. The „i-th from-the-left", „j-th from-the-right" and the „i-th from-the-left and j-th from the right" position is not dependent from „m-th from-the-left", „n-th from-the-right" and „m-th from-the-left and n-th from-the-right" position if the following equations are true for each k, k=0, 1, …, 9, respectively

$$P (p_{i.}^{k} \mid p_{m.}^{k}) = P (p_{i.}^{k}) \quad \text{(iii)}$$
$$P (p_{.j}^{k} \mid p_{.n}^{k}) = P (p_{.j}^{k}) \quad \text{(iv)}$$
$$P (p_{ij}^{k} \mid p_{mn}^{k}) = P (p_{ij}^{k}) \quad \text{(v)}$$

Two positions are mutually independent if and only if, for k_1, k_2=0, 1…, 9 comes true that:

$$P (p_{i.}^{k1} , p_{m.}^{k2}) = P (p_{i.}^{k1}) * P (p_{m.}^{k2}) \quad \text{(vi)}$$
$$P (p_{.j}^{k1} , p_{.n}^{k2}) = P (p_{.j}^{k1}) * P (p_{.n}^{k2}) \quad \text{(vii)}$$
$$P (p_{ij}^{k1} , p_{mn}^{k2}) = P (p_{ij}^{k1}) * P (p_{mn}^{k2}) \quad \text{(viii)}$$

Table 10. Joint distribution of Tens and Hundreds calculated from the data $P\,(p_{.1}^{k} , p_{.2}^{k})$ *(A) Joint distribution calculated assuming independence (B).*

A

k_1	K_2									
Hundreds Tens	0	1	2	3	4	5	6	7	8	9
0	61,1%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%
1	1,9%	0,0%	0,0%	0,0%	0,0%	0,1%	0,0%	0,0%	0,0%	0,0%
2	5,3%	0,0%	0,0%	0,0%	0,0%	0,2%	0,0%	0,0%	0,0%	0,0%
3	4,1%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%
4	0,7%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%
5	25,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%
6	0,5%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%
7	0,2%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%
8	0,2%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%
9	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%

B

k_1	K_2									
Hundreds Tens	0	1	2	3	4	5	6	7	8	9
0	60,6%	0,0%	0,0%	0,0%	0,0%	0,3%	0,0%	0,0%	0,0%	0,0%
1	2,1%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%
2	5,5%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%
3	4,1%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%
4	0,8%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%
5	24,9%	0,0%	0,0%	0,0%	0,0%	0,1%	0,0%	0,0%	0,0%	0,0%
6	0,5%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%
7	0,3%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%
8	0,2%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%
9	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%

The joint distribution, as expected, is not symmetrical. Round decimals attract round decimals to follow. The investigation of dependencies can be summarized so that several decimal digits depend on various from-the-right and from-the-left neighbours. Tschuprow coefficients were adopted to characterize dependence of positions which is illustrated in Table 11 and 12.

Table 11. *Tschuprow coefficients between from-the right decimal positions.*

	from				
Dependence of	Ones	Tens	Hundreds	Thousands	Ten Thousands
Ones	100%	6,5%			
Tens	6,5%	100,0%	1,9%		
Hundreds		1,4%	100,0%	8,8%	
Thousands			8,8%	100,0%	9,2%
				9,2%	100%

Table 12.. *Tschuprow coefficients between from-the-left positions*

	position from				
Dependence of	First from-the-left position	Second from-the-left position	Third from-the-left position	Fourth from-the-left position	Fifth from-the-left position
First from-the left	100%	0,4%			
Second from-the-left	0,4%	100,0%	4,1%		
Third from-the-left		5,9%	100,0%	8,0%	
Fourth from-the-left			8,5%	100,0%	16,8%
Fifth from-the-left				0,6%	100%

Several significant or even strong dependencies were identified between digits but between-positions. Dependencies are not strong in the decimal system. Thus dependencies of digits and positions of gift sizes recalculated in systems 1-2-5 and 1-2-3-5 were investigated.

This has lead to unexpectedly good results. Table 13. proves that in the system 1-2-5, for each number-intitial digit there is a to-the-left neighbour whose chance to follow is high. For each banknote there is another of smaller value to give usually. For each banknote whose value is ten times, hundred times greater than that, donors give a ten times, hundred times greater second banknote.

Table 13. *Frequency distrbution of gifts that can be payed with two banknotes/coins by the size of the greater and smaller denomination.*

	Greater of the two denominations HUF								
Smaller of the two denominations HUF	100	1 000	10 000	200	2 000	20 000	500	5 000	50 000
2									
5	1,10%			0,02%			0,00%		
10	0,92%			0,04%			30,00%	0,25%	
20	8,09%	0,03%		0,05%	0,00%		70,00%	0,28%	
50	89,89%	0,20%		2,91%	0,02%		0,00%	1,66%	
100	0,00%	1,33%		82,36%	0,19%		0,00%	69,22%	0,15%
200	0,00%	6,16%	0,18%	14,62%	0,35%	0,00%	0,00%	28,60%	0,44%
500		92,29%	0,35%		9,96%	0,20%		0,00%	3,11%
1000		0,00%	4,06%		79,45%	1,18%		0,00%	70,22%
2000		0,00%	8,66%		10,02%	0,99%		0,00%	26,07%
5000			86,75%			34,32%			0,00%
10000			0,00%			63,31%			0,00%
Total	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%

Figure 2. *Frequency distrbution of gifts that can eb payed with two banknotes/coins by the size of the majorr (greater) and minor (smaller) denomination. The value of the smasler denomination was plotted on axis x, the size of the greater denomination is denoted by various colouring..*

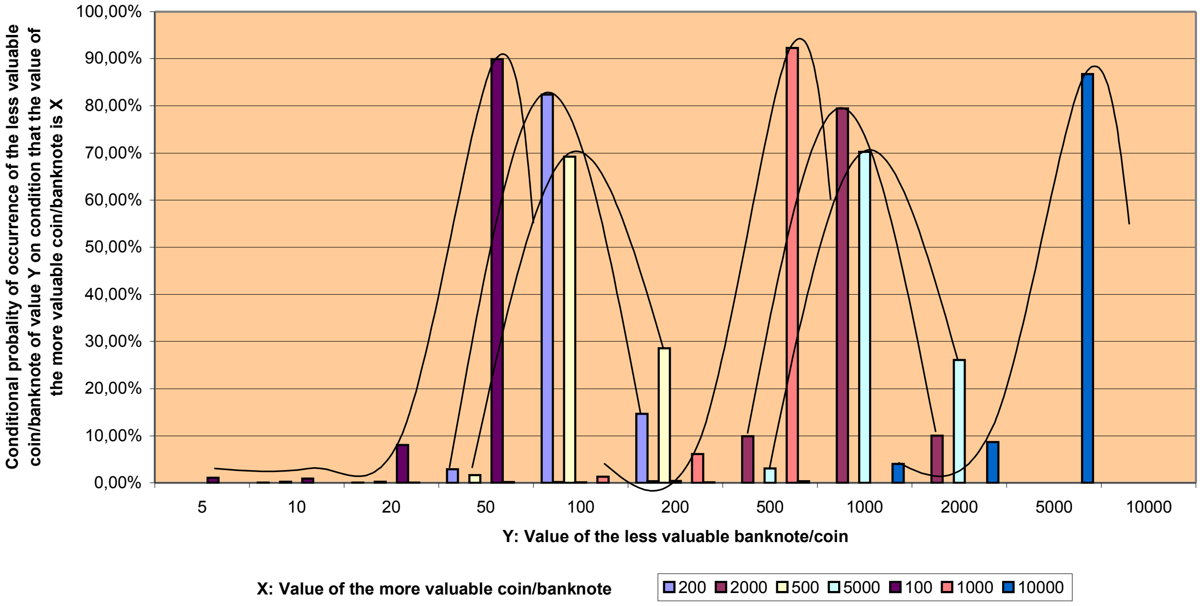


Table 13. and Figure 2. demonstrate that if the value of greater denomination x of two banknotes/coins is x=HUF 10/20/100/200/1 000/2 000/10 000/20 000, then the value of smaller denomination is x/2. Similarly, if the value of greater denomination x of two banknotes/coins is x= HUF 50/500/5 000 then the value of the smaller denomination is mostly x/5. This can be formulated generally in the systems 1-2-5 and 1-2-3-5.

Discussion

Specificity of the distribution of gift sizes rootes in specificity of the situation

We cooncluded that the frequency distribution of gift sizes – and favourites - can be well reproduced by a special sparse distribution, where the numbers of non-zero probability can be generated by a Markovian model, which works in an 1-2-5 system through several decimal magnitudes. Magnitude-invariancy can partly be observed on Baird and Noma (1975)'s distributions, too.

Specificity of the distribution of gift sizes is a consequence of the specific situation in which donors generate the size of their gift. Bestowers do give neither too big (like millions of Euros), nor too small (like Eurocents) gifts. The bell-shaped roof reflects in some way their economic situation, attitudes and values. Furthermore we have no serial numbers. This situation, the mental processes obvously differ from the generation of random numbers, and from the sitation, when the Ss estimate the price of a good or service (Dehaene 2002). In Bock's time-telling experience Ss read out a number, which is fixed. The frequencies of digits, numbers or numerals in various corpora should depend on the content of the corpus: the corpora of laws contains several numbers of the past years, months and days, and small numbers for numbering of paragraphs. In databases that contain residential addresses the distribution of numbers depends on the length of streets and number of levels in the residential blocks. Newspaper corpora contain several numbers related to pages, tables and figures, corpora of belletristic texts contain small and round numbers, which are freuent in t everyday situations. In the Frequency Dictionary of Modern Hungarian Fiction 56 is a frequent number referring to 1956, the Hungarian Revolution. Tabulated values of standard functions are defined by the properties of the function tabulated. The Japanese, Dutch and other corpora mentioned in Dehanene and Mehler (1992) contained „artefacts".

14. táblázat. *Frequency of number-words and Arab numerals in The Frequency Dictionary of Modern Hungarian Fiction (1965-1977). By Füredi and Kelemen. Compiled by the present author. 33 169 superlexemes, 91 471 word forms, 508 008 words.*

Number	Number of occurrences	Number	Number of occurrencec	Number	Number of occurrences
1	1304	20	62	39	<10
2	1075	21	<10	40	27
3	342	22	<10	50	46
4	116	23	<10	56	<10
5	92	24	<10	60	15
6	72	25	17	70	<10
7	51	26	<10	80	10
8	32	27	<10	90	<10
9	22	28	<10	100	48
10	158	29	<10	200	22
11	10	30	33	300	<10
12	20	31	<10	400	10
13	<10	32	<10	500	10
14	<10	33	<10	600	<10
15	29	34	<10	700	<10
16	14	35	23	800	<10
17	10	36	<10	900	<10
18	15	37	<10	1000	34
19	<10	38	<10	10000	14

In Hungarian „egy" is both Numeral and Indefinite Article. The data in the table refer to the Numeral only.

Generation of the gift size is a sequential psychological process

We have seen that digits of a gift size are dependent from their antecedent in the 1-2-5 system. The fact that frequency distribution of gift sizes is discontinuos and gift sizes can be reproduced sequentially by digits and by a Markovian mechanism, suggest that generation of gift sizes itself is a psychologically sequential process in a verbal system rather than choosing a most proper number from a set of numbers along a number line. A whole Markovian sequence can not easily be generated in one-step. The way how the first digit is choosen is the subject of later studies. The sequential hypothesis is supported by our experiments in which Ss generated „small", „great" and „very great" numbers in a tabular structure. Several Ss produced numbers, whose digits showed Markovity. More than half of the sequences of numbers produced were regular. (Dienes 2004c).

The distinguished role the system 1-2-5 plays in the reproduction of gift sizes may come from habitual manipulations with banknotes or from a „natural neural representation"

Why just the system 1-2-5 is the system, in which gift sizes and their frequency distribution are reproducible? Numbers 2, 3, 7, 99 or 324,1111 seem to be uniform at the level of algebra, mathematical analysis or of EXCEL users: Operations can be completed in an apparently uniform way. However, subitization, and several experiments with reaction times indicate that neural ensembles, „organs" of individual numbers, the psychological background of numbers are variegated.

Those are just 1, 2, 3 (= 2 + 1) that can be identified exactly in subitization. These numbers, as well as 5, are prominent in Baird and Noma's distribution. Hurford 2001 showed that morphology, syntax and inflection of numerals of small numbers is mostly irregular or have a special government. Corbett 2000 published data that various languages have various concepts of plurality like Singular, Dual, Minor Plural, Major Plural, Paucal. Feigenson, Dehaene és Spelke 2004 assume two core systems to represent and manipulate quantitative phenomena, one for representing large, approximate numerical magnitudes, and another for precise representation of a small number of objects. In the latter, the great numbers are represented by an ensemble of minor numbers. Feigensohn, Dehaene and Spelke's „approximate" system is used for estimation. However, estimations are also exact numbers. Euclidean algorithm might provide a way how estimations in certain situations are generated. It would not be surprising if the elementary „number-organs" that can be manifest precisely, and exact magnitudes would be the bases of a natural number system, which manifest as favourite gift sizes – and banknotes/coins. System 1-2-5 is a system with 0-1 coefficients, i.e. y/n representation, which may refer to the activity/inactivity of certain magnitude neuron-ensembles, emerging from handfingers and twohandfingers. In this system, Euclidean algorithm, the generation of an approximation or an exact representation of a number can be implemented by cascadic cells.

The Principle Koch-Crick and the opportunities it offers for psychology of numbers

The models published by Dehaene 1992, Dehaene és Cohen 1995, and Dehaene 2001a’s triple-code model, as well as „encoding-complex” model by Campbell és Clark 1986, Campbell 1994, and McCloskey 1992, McCloskey & Macaruso 1995’s modular hypothesis are not suitable for founding a detailed description of generation of gift sizes. These theories do not deal with the situations-frames, inside which numbers are generated. Numerals (number words) are used in the context of a sentence, without which situations are not conceptualized. Koch and Crick’s extended principle (Dienes 2004) offers the foundation to define and describe the representation which extends to the situation in which numbers are used, too. Principle declares that an ensemble of neurons exists to any permanent language element (speech sound, word, idiom) so that it is active whenever the element is produced. The principle implies that counting, enumeration, calculation, computation, comparison of numbers, estimation of sizes/magnitudes, timetelling are all different processes, which assume various number „representations”.

The principle claims that at least so many kinds of organs, organella are which can be manifest related to numbers. One also can say that neuron ensemble of each number includes a part, which is manifest with having seen, heard, written, thought etc. of this number. Whereas in Hungarian, for instance, we talk about „látt” (having seen), „írt” (having written), „fogalom” (concept), „eszme” (idea), „név” (name), ismeret (having known) of numbers, „látt”, „írt”, „hallat” (having heard) etc. of the name of a number, „látt”, „írt”, of a quantity of things and more, so numbers surely must have more than three representations. The principle also entails that the triple code may be a very simplified model. A complete „number line” may be a proper model to states of „number organs” rather than that of the organs themselves, because even a countable infinite manifold of integer number organs, each with a finite volume, would require an infinite volume to hold in the brain.
Manifestations of a language producer and his/her language organs can be denoted by an XML-like Koch-Crick (KL) notation. The kind of manifestation Z (speech sound, word etc.) of kind X (say, hear, think, etc. a psych verb) by a language producer Y at time t, recorded by observer Q is denoted by:

Q - Y_t: X „Z” /X

This formula denotes the fact that Q ascertained that language user Y at t produced Z in some way X. The representation assumed by Campbell and Clark 1988, and Campbell 1994 can all be denoted. The recording of digit „6” on a sheet of paper with pencil, typewriter, or on a magnetic medium by a keyboard can be denoted as **$ „6” /$** and **α „6” /α**, respectively. The organs that manifest by an utterance or hearing of number-word „six” can be denoted by **[~ „six” /~]**, and **[+ „six” /+]**, and **[~ „hat” /~]**, and **[+ „hat” /+]**, in Hungarian. The organ that manifest with accepting that numerosity of something is six is denoted by **[€ „six” /€].** Reading number „175” and the numeral „six” is denoted by: **[ρ „175” /ρ]** and **[ρ „six” /ρ]**. When the number „six” comes to our mind, this manifestation can be denoted by **@ „six” /@** and the neuron ensemble that is active meanwhile by **[@ „six” /@]**. These could be called – following McCloskey 1992 amodal mathematical representation. KC notation is illustrated by examples in Table 15. Quotation marks are omiited.

Table 15. *Examples for the use of KC notation..*

To be denoted	Having seen of number, numeral	Having heard of numeral	Reading out a number, numeral	Having thought at a numeral	Utterance of a numeral	Having written a number, numeral	Having keyboarded a number, numeral	Having perceived the numerosity of sg	Conscious repeating of an action several times
Activity	§hat/§, §6 /§	+hat/+	ρhat/ρ	@hat/@,	~hat/~	$hat/$, $6 / $, $vi/$	α6/α, αhat/α, αviα	€6/€.	□6/□
Active organ, organellum	[§hat/§], [§6/§]	[+hat/+]	[ρhat/ρ]	[@hat/@]	[<hat/>]	[$hat/$], [$6/$]	[α6/α], [αhat/α], [αvi/α]	[€6/€].	[□6/□]

In accordance with the principle Koch-Crick we assume that dimensioned quantities, and numbers like „six gallon” and „six” in „six gallon of water” as well as dimensionless numbers should have correlates. The correlates of dimenisoned numbers should differ from those of numbers, because dimensioned numbers are bound to units and kinds. A dimensionless number may stand for the classes of dimensioned numbers, which would inherit their operations: 1 + 1 = 2, since 1 kg coal + 1 kg coal = 2 kg coal as well, as 1 l milk + 1 l milk = 2 l milk.

If objects are manually counted, counting is done sequentially one ofter the other, and occasionally, later or in a cascade system by groups, pulling together the items into groups of five or ten. Euclidean algorithm is also suitable for counting objects by mentally grouping them, but it is efficient only, if groupers work parallel. Groups consisting of as many elements as base-numbers should be perceived in a parallel way, and then should be counted. In the binary and 1-2-3-5 systems, it is the presence of groups only, which should be perceived.

To explain the process of generation of a number we assume as many amodal „number-idea" neural ensembles (organella), as finite and infinite elementary numerals (@ one /@, @ two /@ ,.., @ many /@, @ few /@, etc.) exist. These ensembles perhaps may be related to those identified by Nieder and Müller (2003). Additionally other ensembles definable by KC notation are also assumed. Incative bases of the number manifest with zeroes when writing in the Arabic decimal system, but are not voiced when read. Operations that are manifest with –en and –ű when speaking or writing number words, are not manifest when writing Arabic numerals.
If subitization takes place in a system 1-2-(3?)-5, this can be KC formulated that organs [€ „four" /€], [€ „eight" /€], etc. are the structures of organs [€ „one" /€], [€ „two" /€], [€ „three" /€], [€ „five" /€]. The events like € „nine" /€, or € „twenty one" /€ take place in a process, in which both finite and infinite number-organs are activated. For instance, when 15 objects are subitized, the organs [€ „one" /€], [€ „two" /€], perhaps [€ „five" /€], the organs [@ „and" /@], and/or [@ „plus" /@] and/or [@ „times" /@] and [€ „few" /€] all could be activated, which might lead to an - erroneous - activity of the organ [~ „theirteen" /~]. Finally, this activity leads to saying thirteen.
Grossberg and Repin proposed multi-digit numerical representations that obey a place-value principle (actually serial cognition) to arise through learned interactions between categorical language representations in the What cortical processing stream and the Where spatial representation. Their sequential model, if assuming 1-2-5 and multiples by 10, as place-values at least in the donation situation, might explain our findings.

Many times, giving is a habitual activity

The experiments by Groen & Parkman (1972) made known that calculation is a learned habitual activity. Persistent and frequent donors usually keep giving the same amount and this may indicate that the cheque-writing phase of donation is a habitual, even „automatic" activity triggered by reading the inviting mail. This may contribute to scarcity of the distribution. The significance of habituality in donation hasve been proved by the results of our surveys (Dienes 2004b, 2004d).

A possible process of the manifestation of a number to be a gift-size in non-habitual donation

If gift-size is not „retrieved automatically", the sizes of denominations of banknotes seem to be the bases of donors, when they set forth the amount of their donation according to an Euclidean algorithm. The system 1-2-5 is a 0-1 i.e. y/n representation, which may refer to the activity/inactivity of their natural magnitude neuron-ensembles. The first neural magnitude-ensemble may be activated by financial position, motives and habits of the donor.   The first ensemble, which manifest by the first digit may be followed by others, but due to low motivation of donors to give more or to produce many digits, the generation of gift size soon decays. In the process of the generation of gift-sizes, the amodal neural ensembles of abundant favourite dimensionless digits like [@ „100 000" /@], [@ „50 000" /@], and [@ „20 000" /@] etc. or dimensional neural ensembles like [@ HUF 100 000 /@], [@ HUF 50 000 /@], and [@ HUF 20 000 /@] etc. manifest in the order of decreasing.
One could hypothesize that neural ensembles [@ „HUF 10" /@], [@ „HUF 20" /@] etc. are getting freqeuented just as a simple consequence of their abundant use in everyday situations of paying, without being involved in percepting or generating numerosities or operations above them. Then, however, experienced Markovity of digits would remain unexplained. Furthermore, if donors would conceptualize a complete „paying situation", they would consider cases when banknotes or coins are returned, which is surely not the case.

Summary

In non-habitual donation, gift size is generated by digits in an 1-2-(3)-5 system, in a sequential Markovian process, which is similar to the standard Euclidean algorithm. Grossberg and Repin's model could be adopted for explaining the results. The process can be conceptualized as sequential appearance of imaginary banknotes/coins. In the habitual donation donors give the usual amount, whose size was generated earlier in the sequential process.

Acknowledgements

Thanks to Cég-INFO Ltd. for supporting the work, Dénes Szűcs for his friendly help in finding and discussing references and invaluable consultancy.

References

Baird J.C - Noma E.: Psychophysical study of numbers. I. Generation of numerical responses.
Psychological Research, 1975. 37, 281-297

Banks W.P., Hill D.K.: The apparent magnitude of numbers scaled by random production.
Journal of Experimental Psychology Monographs. 1974. 102, 353-376.

Russ Reid Company & Barna Research Group
The Heart of Donor, A Lifestage Analysis, 1996
in The Direct Marketing Association: *Statistical Fact Book '99 21st edition.* New York, 2000, P. 236.

Benford F.: The law of anomalous numbers.
Proceedings of the American Philosophical Society, 1938. 78, 551-572

Bock Kathryn et al.: Minding the clock
Journal of Memory and Language, 2003. 48, pp. 653-685

Campbell J.I.D. - J.M. Clark: An encoding-complex view of cognitive number processing: Comment on McCloskey, Sokol, and Goodman (1986)
J. of Exp. Psych: General 1988. 117, 2, pp. 204-214

Campbell J.I.D.: Architectures for numerical cognition
Cognition, 1994. 44, pp. 1-44

Corbett, Greville C.: *Number*
2000, Cambridge University Press

Dehaene S., Mehler J.: Cross-linguistic regularities in the frequency of number words.
*Cognition.*1992. 43, 1-29

Dehaene S.: Varieties of numerical abilities
Cognition, 1992a 44, pp. 1-42.

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

Dehaene S.: Précis of the number sense
Mind and Languague . 2001a. *16, 1, pp. 26-36.*

Dehaene S., Marques J.F.: Cognitive Euroscience: Scalar variability in price estimation and the cognitive consequences of switching to the Euro.
Quarterly Journal of Experimental Psychology, 2002. 55, 703-731

Dehaene, S. et al.: Arithmetic and the brain
Current Opinion in Neurobiology, 2004, 14, pp. 218-224.

Dienes I.: On the one that exists – from the one, which is out of the one that exists (in Hungarian)
Az, amiből van az, ami van.
In László J., Kállai J., Bereczkei T. (eds) Levels of representation (in Hungarian)
A reprezentáció szintjei
Budapest, 2004a, Gondolat Kiadó, 382 p., pp. 106-134.

Dienes I. (ed.): Phone poll among active and inactive donors. 2004-i. (in Hungarian)
Az aktív és az inaktív donorok 2004. januári telefonos megkérdezése.
CID Internal Report, Budapest, 2004b.

Dienes I.: How much is „few", „many" and „ very many"? How great is a „little", „large" and „very large" number? An empirical study of the contents of concepts of indefinite numerosity by a spreadsheet test. (in Hungarian)
Presentation for *XII.th Hungarian Cognitive Science Conference, Debrecen, January 29, 2005*.
Mennyi a „néhány", a „sok" és a „nagyon sok"? Mekkora a „kis", „nagy" és „nagyon nagy" szám? Néhány határozatlan-számosság fogalom tartalmának vizsgálata számítógépi teszttáblázattal.
Manuscript, Budapest, 2004c, 16 p.
Előadva 2005.01.29, Debrecen, MAKOG'05

Dienes I. (ed.) Survey among active donors (in Hungarian)
Kérdőíves kutatás az aktív donorok között 2004. első negyedévében.
CID Internal Report, Budapest, 2004d.

Dienes I. (2006) Why are the favourite numbers favourite? (in Hungarian)
Miért kedvenc számok a kedvenc számok?
In: Ed. V. Altbacker: *Proceedings of the XI.th Hungarian Cognitive Science Conference, Tihany, January 29,* 2004.
Gondolat Kiadó, Budapest, in press

Feigensohn L., S. Dehaene, E. Spelke (2004) Core Systems of Number.
Trends in Cognitive Sciences 2004. Vol.8. No.7.

Füredi M. - Kelemen J.: Frequency Dictionary of Modern Hungarian Fiction (1965-1977). (in Hungarian)
A mai magyar nyelv szépprózai gyakorisági szótára.
Akadémiai Kiadó, Budapest, 1989.

Hurford J.R.: Languages treat 1-4 specially
Mind and Language 2001. 16, pp. 69-75.

McCloskey M.: Cognitive mechanisms in numerical processing: Evidence from acquired dsycalculia
Cognition, 1992. 44, pp. 107-157.

McCloskey M., P. Macaruso: Representing and using numerical information
American Psychologist May, 1995. pp. 351-363

Nieder, A. and E.K. Miller, Coding of cognitive magnitude: Compressed scaling of numerical information in the primate prefrontal cortex. *Neuron* 2003. 37 pp. 149–157.

Meeuwissen, M, A. Roelofs, W.J.M. Levelt: Incrementality in naming and reading complex numerals: Evidence from eyetracking.
Quart. J. of Exp. Psy. Ser. A: Human Experimental Psychology. 2005

Grossberg, S., D.V. Repin: A neural model how the brain represents and compares multi-digit numbers: spatial and categorial processes.
Neural Networks, 2003, 16, pp. 1107-1140.

Damian, M.F. Asymmetries in the processing of Arabic digits and number words.
Mem. and Cognition, 2004, 32(1), pp. 164-171.