

THE MENTAL LEXICON: RESULTS OF SOME WORD ASSOCIATION EXPERIMENTS

MÁRIA GÓSY – MAGDOLNA KOVÁCS

Abstract

There are numerous hypotheses concerning the structure, size, and strategies of adults' mental lexicon. This is the first time, however, that children's mental vocabularies are analysed using the technique of free word associations (with the participation of two hundred 12-year-old and two hundred 13-year-old pupils). The analysis focuses on both quantitative and qualitative characteristics of the data like types of associations, lexical representations, distribution of word categories or semantic analysis of words. Comparisons are also made with a very similar material found in the Hungarian literature that provides a unique opportunity to look at the differences of the mental lexicon after 60 years. The discussion concerns (i) the patterns of the tested children's mental lexicon (including the individual performances) and (ii) vocabulary changes seen as a multifactorial consequence of the progress of time. The hypothesis about the speed of lexical access being a definitive factor in the development of the mental lexicon has been confirmed and may be applied to other languages as well.

1. Introduction

The size and structure of the mental lexicon is a subject that has been intriguing people for centuries. Numerous hypotheses have been advanced (cf. Gósy 1999b); and modern science, especially psycholinguistics, tries to answer the arising questions by advanced, mainly experimental, methods. In the study of the process of first language acquisition, for instance, it is a central issue how the child's first words come into being and how they are later modified or changed during his/her maturation. How does it happen that the totality of his/her words come to constitute the child's mental lexicon, and how are the lexical representations of those words constructed and stored? Although we have learnt a lot from the accumulated empirical data and the results of relevant experiments, a number of related facts are still unknown and researchers are bound to formulate hypotheses.

Our idea of the time of appearance, articulatory features, and underlying phonological representations of children's first words is fairly accurate (cf. McNeill 1970; Waterson 1987; Steinberg 1993 for Hungarian: Kenyeres 1926;

Vértés O. 1955; Meggyes 1971; Gósy 1999b). However, there is no consensus among researchers concerning the origin of word meanings in the beginning phases of language acquisition. Theories are sometimes concurrent, but they often exhibit large differences, too (cognitive theory, association theory, probability theory, prototype theory, etc., cf. Berko Gleason 1985). The child's word stock comes about via perception/comprehension, by far antedating production. For instance, a 13-month-old child was reported to understand as many as 50 different words, whereas the same child did not become capable of producing 50 different words until she reached 19 months of age (Benedict 1979). This might suggest that people have two mental lexica, one for production and one for perception; but that situation is not likely to obtain even at the beginnings of language acquisition. It is just a matter of two different uses of the same mental lexicon (neurological arguments and experimental data both support the claim that humans have a single mental lexicon, cf. Coleman 1998). The literature on first words is extensive, but research on what happens to the word stock in later years is rather sparse. It centres around the acquisition of meanings, the phonetic and phonological form of words, and the length of words making up the child's utterances; these are relatively easy to analyse due to the limited number of words at hand. Special experiments are devoted to the investigation of the vocabulary of older children (for instance, the characteristics of the acquisition of parts of speech, or the development of the meanings of function words at various ages).

The size of the word stock does not necessarily exhibit linear correlation with the child's age, and shows rather extensive variation across individuals. On the other hand, the development of the comprehension and production of single concepts is a lot more uniform and regular. Learning a new word does not simply involve the acquisition of the relevant concept (or meaning) but also that of its connections with other items. The most widely accepted assumption is that the acquisition of meaning is primary, whereas the recognition and use of semantic relationships is secondary (McNeill 1970). Indeed, semantic fields are rather late to develop in child language. This observation does not contradict the assumption that, due to a general strategy of language acquisition, lexical items are organised into networks (that assumption being supported by physiological facts about how neurons work). Results of association experiments have confirmed the "cobweb-like" organisation of the word stock even at a relatively early age (Aitchison 1987). That organisation is almost unquestionably a semantic one. At any rate, it appears that children possess some kind of preprogramming with respect to the way newly acquired words are to be stored in their mental lexicon such that they are the most

appropriately accessible when needed. The children are able to incorporate the newly identified lexical items in their mental network; but the newcomers are connected to one another and to the items that were present before their arrival a lot more loosely than the latter are to each other.

Research in the past ten years seems to confirm the assumption that the development of children's vocabulary does not only go on in terms of articulatory and acoustic imprints but also by phonological mapping (Ingram 2001). That hypothesis is backed up by the idea that the child's "word stock" based solely on his/her perception of speech and general perceptual abilities is enough for phonological organisation (a prerequisite for the production of the first words) to begin. The large-scale initial development of the production vocabulary is based on the combinability of its items. In other words: the network of the child's mental lexicon begins to be built. This is also supported by the fact that the acquisition of as many as 100 new words a week until a particular child became 22 months old was recorded without noticeable changes in the phonological system of that child during the same period (Ingram 2001). Thus, later development takes the form of increasingly more detailed phonological representations and a more limited number of acquired new words. By the age of three or four, the following points probably characterise the child's mental lexicon: fundamental language specific phonological representations have taken shape, the semantically based network has been structurally completed, storage takes place as in the adult system, and the processes of lexical access approximate those observed with adult speakers.

Adapting Levelt's (1989) speech production model to our present purposes we can say that, in lexical representations, the levels of lemmata and lexemes are fully formed from the very beginning, except that their use may differ between children and adults. With the gradual elimination of children's double (phonetic) storage and a simultaneous increase of their word stock, the lemmatic and lexemic levels get clearly separated. This means that children's lexical access in speech production is characterised by strategies similar to the adult ones (at least from the age of the second biological barrier onwards). It has been experimentally demonstrated that with English-speaking five-year-olds neither the length of words nor their phonemic similarities affect the process of vocabulary access, whereas with seven-year-olds both factors play important roles (Gathercole–Baddeley 1993). The semantic associations of five-year-olds still contained concepts that did not fit the category at hand, but this never happened with seven-year-old subjects. Word activation based on the initial sound sequence was hardly ever possible with four-year-olds,

whereas at the age of seven or later children can even access several words satisfying the conditions of the given “rule”.

The question repeatedly arises of what size the child’s vocabulary may be at the various ages and how much it grows from one age group to another. The initial word stocks of children, despite often very large individual differences, show considerable similarity (Jarovinskij 1995). The differences are quantitative rather than qualitative (cf. Bakonyi 1918). Some data in the literature suggest that the acquisition of words is relatively slow until 1;6 and that at that point a sharp increase usually begins. A child who only knew 22 words at 1;6, had acquired 118 words by the age of 1;9, a mere three months later (Ingram 2001). Janota (1970) compared her own child’s word stock with the data of three other Czech-speaking children, tested by the Czech linguists Ohnesorg and Pačesová. At the same age, the four children had a vocabulary of 500, 921, 1613, and 1913 items, respectively. An analysis of their first 500 words showed that these sets of words included identical items in 48%, respectively 56%, for two pairs of the children (the latter pair was made up by siblings). Comparing all four children, 24% of their 500 words were completely identical. In terms of the estimated size of word stock at 3 years of age, individual variance is, again, large; a summary figure referring to English and German children’s data sets the lower limit at 500, and the upper limit at 3600 words. Hungarian children’s vocabularies are also given as varying between wide limits: 150–1500 words. For the production vocabulary of thirty nursery-school children (between 3;0 and 3;3) an average of 1256 words was found (Gósy 1984). Ildikó Meixner observed the first appearances of words with eight children until they became three and a half years old; she recorded 260 words from the child that had the fewest, and 1468 words from the one that had the most (Meixner 1976). The elder son of one of the authors of the present paper had 450 words at the age of 2; his word stock was estimated at 1600 words when he was three. If we take that estimate to be correct, then the average frequency of his acquiring new words comes to 3.15 words a day, 22.11 words a week, or 95.83 lexical items per month (of course, the real growth of vocabulary was not necessarily linear). And those figures do not include the acquired morphological units (suffixes), pieces of knowledge concerning the structure of the language (the numerical characterisation of that knowledge being problematic anyway) or pragmatic features. It is clear that that speed of acquisition cannot be consistently upheld for a long time; however, estimations concerning later years are even more vague.

The word stock of 6–7-year-old children is interesting in a number of respects; again, we have to rest content with estimates. According to a figure

from 1961, the vocabulary of American children at the age of six averages on approximately 23,700 words (Carroll 1961). Yet other data suggest that a six-year-old American child knows 7800 roots (and 13,000 words) and that his/her vocabulary grows by 14.5 roots (21 words) per day (Lindblom 1999). Accordingly, (s)he knows 10,585 roots and 15,330 words by the time (s)he is eight years old. Rather than commenting on those figures, let us note that the dictionary of the Hungarian writer Mór Jókai's complete works consists of 22,000 entries. However, individual researchers differ in their estimations. Clark (1995) estimates the vocabulary of the average five-year-old American child at 14,000 words, and she thinks that (s)he acquires some 300 new words a year. That means that the word stock of an eight-year-old child should consist of approximately 14,900 words. Tentatively accepting that growth rate, the vocabulary of the Hungarian boy referred to above should have consisted of 3100 words at the age of eight (?). In comparison, the word stock of an American adult is assumed to vary between 75,000 and 150,000 words (Berko Gleason – Ratner 1998, 158). Other estimations say that the 50 most frequent English words make up 60% of conversations, that is, 10–15 words are uttered before one occurs repeatedly. One of the most widely known dictionaries, the *Longman Dictionary of Contemporary English*, containing 55,000 headwords, explains them using as few as 2,000 words. All that is definitely enough to convince us that we know next to nothing about the size of children's (or adults') vocabularies and even the most serious estimations, sometimes based on experimental data, are extremely nebulous.

András O. Vértes set forth the methodological difficulties of establishing children's vocabulary sizes as early as in 1955. It is known that, at any stage of language acquisition, very large differences can be found in the linguistic performance of children. Concerning individual differences in speech perception, there are Hungarian data as well (Gósy 1999b); hence, individual variation in vocabulary size can be safely assumed, too. The primary aim of vocabulary tests known today is not to estimate the number of lexical items used at a given age, but rather to compare the subject's performance to the standard value of the given test and hence the early screening of children who do not perform as expected (e.g., the Peabody vocabulary test or for Hungarian: Lórik et al. 1995).

Word associations are often used as a method of exploring semantic representations and lexical access. The first known word association experiment was due to Francis Galton (1883) who tested the functioning of word associations, as well as their speed, on himself. The method has three familiar versions. (a) The procedure known as free word association requires that the subject ut-

ters the first word that occurs to him/her on hearing a (pre)determined clue. (b) In restricted word association tests, the clue is to be followed by a response of a given category (like its opposite, or a similar-sounding word). (c) In open-ended tests, the subject has to list as many words of a given semantic category (like *animals* or *pieces of furniture*) as (s)he possibly can.

Data obtained from word association tests often shed light on other linguistic phenomena as well. The syntagmatic/paradigmatic shift in language acquisition takes place at around the seventh year in English-speaking children (Berko Gleason–Bernstein Ratner 1998). This means that while younger children tend to activate words in free word association tests on the basis of syntactic relationships (e.g., they respond to the clue *dog* by the verb *bark*), older children and adults remain within the same semantic category (they activate *cat* on hearing *dog*). The fact of the syntagmatic/paradigmatic shift has been explained in a number of ways, and it can also be assumed that, due to certain structural features of the language concerned, the shift takes place at different ages in speakers of different languages.

The **aim of the research** reported in the present paper is a multifunctional analysis of the vocabulary of native Hungarian adolescents. Our experiment is a repetition of one performed in 1939 by János Cser, with a comparison of the data obtained, and their evaluation in themselves and in conjunction. Vocabulary size experiments found in the literature mostly focus on the process of lexical access as part of speech comprehension. On the other hand, an analysis of the results of lexical access activated during speech production may yield a more accurate view of the organisation of the mental lexicon. This claim can be substantiated by the following considerations: (a) In free word association tasks, both the active and the passive aspects of the mental lexicon are activated. (b) The faster the lexical access, the larger number of words can be activated. (c) The larger number of words are available, the easier it is to access them (since there are no linguistic or other rules to restrict the process). (d) The network character of the mental lexicon can be rendered probable in this way. (e) Conclusions can be drawn concerning the way individual linguistic categories are encoded. Finally (f) cautious estimations can be formulated concerning the size of the subjects' word stocks.

We have chosen subjects aged 12–13 because we wanted to make sure that we investigate the characteristics of the vocabulary of the period following what is called the second biological barrier. In addition, this was made possible by the fact that the experiment published in 1939 included children of that age group. Thus, it also became possible to get an idea of which words had become obsolete or arose as new lexical items during the intervening sixty-odd years.

Also, the fact that the two experiments can be compared makes it possible that we observe linguistic changes (and changes in language use) in the field of word stock.

2. Method and material

The method we have chosen was a repetition of the version of free word association tests that was employed and described by János Cser in his 1939 book (other references to this particular type of free word associations were not found in the literature). The point of that procedure is that the child is to activate words for 15 minutes without constraints of any kind (the subjects had to write down the words). Cser motivates his choice of that procedure by noting that the data can thus be made use of in three different ways. First, by analysing the number of words produced; second, by gaining insight into the child's way of thinking via exploring the characteristics of the lists of words obtained; and third, by getting a tool for estimating the active vocabulary sizes of the children tested, as well as the frequency of the words occurring in the lists.

In János Cser's experiment a total of 4483 boys and girls were involved; they all came from Budapest and were aged between ten and fourteen years. They were pupils of various kinds of schools: elementary, higher elementary, or secondary. However, the data coming from only 1000 of them (200 in each age group) were taken into consideration. Testing took place under classroom conditions. Although the aim was to elicit free word associations, two restrictions were nevertheless imposed on the procedure: the children had to avoid proper names and suffixed forms. In our own experiment 12 and 13-year-old boys and girls participated (from various parts of the country); sixth- and seventh-formers of primary schools, as well as secondary school pupils of the same ages. Both age groups included 200 children (100 girls and 100 boys), a total of 400 children. They also performed their task during a school class. In giving instructions to them, we did not make any restrictions (thus, they were allowed to write proper names and suffixed forms as well). We recruited our test groups both from Budapest and from other parts of the country. Participants of Cser's experiment had learned to write with oblique strokes, used dip-in pens, and had to observe "neatness" of their writing as was customary in those days. Our own subjects had learned upright writing, used ball-point pens or (rarely) pencils, and neatness was not an aspect of writing that was usually graded in those age groups.

In spite of the minimal amount of deliberate changes and the differences caused by the time elapsed between the two experiments, the series of tests conducted in the autumn of 1999 was comparable to the one that had been

conducted more than sixty years beforehand. The word lists that our 400 children wrote were entered in a computer file; the evaluation was performed by one of the present authors (M.K.) using an MS Access database management system and special task programs developed by herself. The statistical analysis was done with an SPSS for Windows software.

3. Results

3.1. Data from our own corpus

In order to determine the size of our corpus, the first thing we had to do was to define the notion of ‘word’. Indeed, free word associations unambiguously suggest that the units stored in the speaker’s mental lexicon are but partly characterizable in terms of the notion ‘word’ as defined in classical grammars. A ‘word’ can be a stem or a suffixed form, a compound or even a phrase. It is not infrequently difficult to tell the latter two from each other, if we think in terms of grammatical constructs, not in terms of spelling. In addition, a single activation sometimes resulted in accessing a multi-word sequence that had to be considered a unit just as much as a monomorphemic word did. Therefore the type/token differentiation could not be used unambiguously with our corpus. In order to solve the above dilemma, we will refer to the units stored in the mental lexicon as ‘mental words’ which may correspond to a ‘traditional’ word in the grammatical sense but may just as well differ from it structurally as well as in their ability to occur as syntactic phrases or indeed as sequences of phrases.

Thus, a mental word is a stored lexical unit as opposed to a lexeme that stands for a single concept as in Levelt’s (1989) definition. The lemmatic level of Levelt’s speech production process need not detain us here. A lexeme—in Hungarian—is a semantically, syntactically, and phonologically defined form of a concept. (For simplicity, the term ‘word’ will also be used below as a synonym of ‘lexeme’.) The type/token distinction was taken into consideration in two different ways. With respect to the repetitions across (as well as within) children, whereas in ‘lexeme’ counts each lexeme is counted just once. With respect to individual children’s word lists, repeated items count as a single lexeme each but are included in the number of mental words as many times as they occur. Our 12 and 13-year-old subjects wrote a total of 52,764 mental words, of which 289, respectively 250, that is, a total of 539 words were illegible and will be disregarded in what follows (the ratio of illegible items was a

mere 1.02%). The corpus then contains **52,225 mental words**, this is what we considered to be 100%. In the analyses, we simplify by taking the average age of sixth-formers to be 12 years, and that of seventh-formers to be 13 years.

The number of mental words activated by sixth-formers was **25,761**, the same number for seventh-formers was **26,464**, the difference being 703 words. The difference can be explained in two ways. First, the more hoped-for explanation is that the size of the mother-tongue mental lexicon differs in this manner between the two age groups or else its use becomes faster and more accurate as time goes by. This would entail that the average difference between 12 and 13-year-olds is the knowledge of 703 words, a value that does not contradict the assumptions in the literature. It cannot be excluded either that the difference between the two age groups is not in the size of their mental lexica but rather in their processes of lexical access. This would mean that access by the older children is faster and more successful, a factor that in turn leads to their association to more numerous words.

The other possible explanation runs as follows. Seventh-formers obviously write faster than sixth-formers do, and this makes it possible for them to keep up with their own activation of their mental lexicon to a larger extent. In this case there need not be an inequality in the number of activated mental words, the difference may be simply due to the speed of handwriting. The analysis of the number of syllables showed that older pupils wrote longer words. The 12-year-old subjects activated mental words of 2.14 (girls) and 2.15 (boys) syllables on average (the scatter was 0.18 for girls and 0.25 for boys). For 13-year-olds the average was 2.22 (girls) and 2.20 (boys) syllables per mental word (the scatter was 0.22 for girls and 0.25 for boys). The age-bound difference is statistically significant, thus we can conclude that the older group activated somewhat longer words. (For comparison: the average length—in number of syllables—of Hungarian monomorphemic dictionary items is 1.94, cf. Papp 1973.)

General experience and the analysis of our data both make us conclude that the surplus is due to the better functioning and more loaded character of the mental lexicon, rather than to the superior writing skills of the older group. Seventh-formers do not write considerably faster than sixth-formers do, and the neatness of their handwriting does not differ much either. Faster lexical access, as well as activations more rapidly following one another, obviously make it possible to access a larger number of words in a given amount of time. The question still is whether this in itself explains the difference of 703 words or the growth of the mental lexicon in a year has to be taken into consideration as well.

Our 12-year-old subjects activated **4642** different **lexemes**, whereas the 13-year-olds wrote **5280** (discounting proper nouns). The younger ones wrote 1930 lexemes that did not occur in the seventh-formers' material; the older group wrote 2568 that the others did not. The number of lexemes that both groups activated was 2712, that is, 58.42% in the case of 12-year-olds and 51.36% in the case of 13-year-olds. Examples of items that only the younger group mentioned and whose frequency of occurrence was 3 or more (a total of 116 such items were found): 10: *gömbölyű* 'round', 7: *boszorkány* 'witch', *komoly* 'serious', *muskáтли* 'geranium', *zsarú* 'cop', 6: *jószívű* 'warm-hearted', *csóró* 'poor', *jeles* 'very good', *sündisznó* 'hedgehog' (etc.), 5: *anakonda* 'anaconda', *csukló* 'wrist', *krampusz* 'bogey-man', *utcalány* 'street-walker' (etc.), 4: *asztalos* 'joiner', *csacsi* 'donkey', *emésztő* 'cesspool', *harmonika* 'accordion', *hőmérséklet* 'temperature', *adás* 'broadcast', *páva* 'peacock' (etc.), and 3: *bojt* 'tassel', *csipke* 'lace', *csorda* 'herd', *előhang* 'prologue', *féreg* 'worm', *nyomorék* 'disabled', *pelyhes* 'downy', *szalonna* 'bacon', *zsalu* 'shutters', *tejszínhab* 'whipped cream' (etc.). The material contained a few obscene words, too. It is interesting that a few 'neologisms' were also found that illustrate the changes of the words stock, e.g., *csúcsragadozó* 'top predator' that as many as three pupils included in their lists.

The number of lexemes occurring only with the 13-year-olds (and occurring three or more times) was 189, significantly more than those mentioned by the younger ones only. Examples: 9: *deltoid* 'id.', *egészség* 'health', *internet* 'id.', 8: *fájdalom* 'pain', *takarít* 'clean', *völgy* 'valley', *kémcső* 'test-tube', 7: *higany* 'mercury', *sor* 'row', *szex* 'sex', 6: *hélium* 'helium', *pia* 'booze', *nyávog* 'mew', *óvszer* 'condom', *alvázzám* 'chassis number', 5: *készít* 'make', *tanító* 'teacher', *szenvedés* 'suffering', *őrült* 'mad', *puszi* 'kiss', *pszichopata* 'psychopath', *himnusz* 'anthem' (etc.), 4: *nemi szerv* 'genital organ', *óda* 'ode', *paróka* 'wig', *hangkártya* 'sound card', *koporsó* 'coffin', *jövő* 'future', *veszekszik* 'quarrel', *utálat* 'disgust', *vallás* 'religion' (etc.), and 3: *fogamzásgátló* 'contraceptive', *féltékenység* 'jealousy', *bimbó* 'bud', *költészet* 'poetry', *horgászik* 'angle', *kocog* 'jog', *rendszám* 'registration number', *intelligens* 'intelligent' (etc.). The numbers and the examples conceal the possibility of a semantic analysis that may involve a number of psychological conclusions as well. Since the aims of the present paper do not include such an analysis, let us rest content with making a few points here. The most frequent associations of 13-year-olds contained names of animals in only 1.05% of the cases, whereas the same ratio with 12-year-olds is still 14.65%. On the other hand, words having to do with sexuality and negative emotions exhibit a sharp rise between the two groups. With 12-year-olds, these two sets of words constitute 2.58% each, whereas a year later

they go up to 4.23% and 18.9%, respectively. Frequent words in the field of hobbies show a shift, too. With 12-year-olds, fairy tales dominate (*boszorkány* 'witch', animals, *krampusz* 'bogey-man'). The effect of television also shows up in their associations (*szappanopera* 'soap opera', *bemondó* 'announcer'). 13-year-olds mention leisure activities characteristic of older age groups, although *tévé* 'television' occurs frequently, too (*táncol* 'dance', *internet*, as well as other computer terms). Although using a different lexeme, both groups mention *pecázás* 'dangling a line' (12-year-olds) or *horgászás* 'line-fishing' (13-year-olds). The effect of school subjects can be seen in both groups, with a characteristic difference of activated words; items like *barnakőszén* 'brown coal', *hőmérséklet* 'temperature', *nagyító* 'magnifier', *paraszt* 'peasant', *világegyetem* 'universe', or *viasz* 'wax' (sixth-formers) contrast, as it were, with the associations of the seventh-formers like *elektron* 'electron', *deltoid* 'id.', *kémcső* 'test-tube', *higany* 'mercury', *tápanyag* 'nutriment', *autotróf* 'autotrophic', *magnézium* 'magnesium', *stroboszkóp* 'stroboscope', or *vegyjel* 'chemical symbol'.

Average values, as well as minimal and maximal numbers of activated words, are summarised in Figure 1 (overleaf). The scatter values are 41.94 (girls) and 36.6 (boys) for 12-year-olds and 39.5 (girls) and 39.06 (boys) for 13-year-olds. Individual differences, as was to be expected, are rather large. Girls activated more words than boys did, although in minimal and maximal amounts the differences are slight. This means that in both age groups more girls activated significantly more words. The difference between younger and older girls is larger than that between younger and older boys; boys exhibit stagnation across age groups or even declining performance (in terms of the minimum values). With girls, the minimum value grows twice as large in a year.

The total number of activated mental words, for the 12-year-old girls, was 13,877, whereas for the 13-year-old girls it was 14,685. That is, figuratively speaking, girls activated some eight hundred words more 'a year later', hence their 'word acquisition rate' was 2.21 words a day. The number of words activated by the 12-year-old boys was 11,874, that of 13-year-old boys was 11,751. The difference is a mere 123 words (approximately 1%), but with the younger group having the larger number, that is, having activated that many more words in the allotted time. We can conclude that the boys of the two age groups did not show any difference in the process of lexical access.

The above questions become especially important if we ask them for boys and girls separately. The apparent stagnation of boys' mental lexica may correspond to their writing skills (boys usually write more slowly and less neatly than girls do), but it may also be independent of them. In either case, there is an obvious disadvantage compared to girls, one that affects boys'

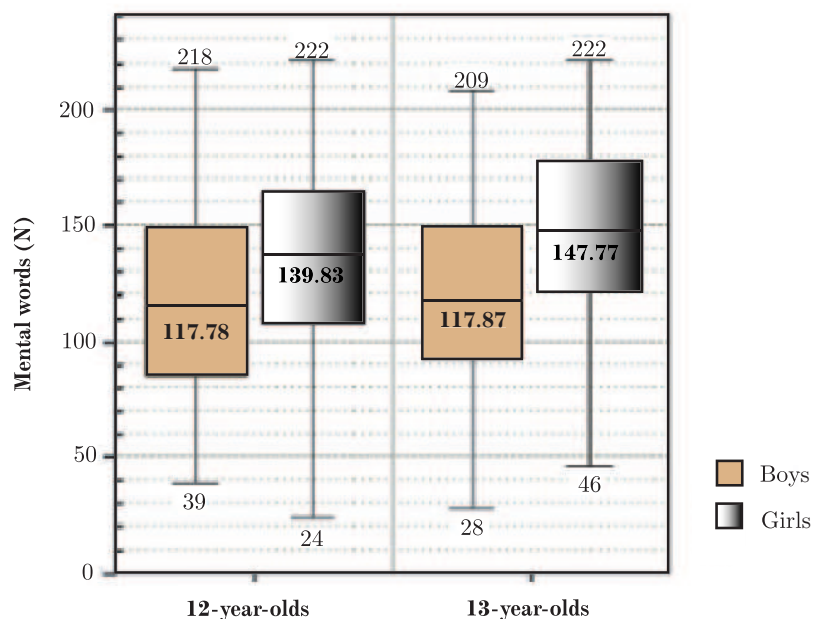


Fig. 1

Results of the word association test by age and gender

learning processes, their results in testing, and altogether adversely affects their appreciation at school.

With respect to the limiting values, an important question arises. Did children who activated extremely few words write frequent or infrequent words (compared to the full corpus)? Our hypothesis with respect to this question was that children accessing few words would necessarily activate words that are the most frequent for the given age group. The data confirmed our assumption. We picked the ten sixth-formers and the ten seventh-formers who wrote the fewest words. Table 1 shows how the words they wrote—a total of 858 lexemes—were distributed in terms of frequency ranges. We made five word groups on the basis of how many pupils activated the same lexeme. The figures demonstrate that the occurrence of identical items is very high with those who activated few words, that is, individual differences are the smallest in their case.

A statistical analysis of the whole corpus also revealed a negative correlation between performance per pupil and the frequency index of the activated lexemes. It can be stated that the more words a pupil activated the higher number of his/her words fell in the $3 \leq n \leq 39$ range.

Table 1

Frequency of words with children who wrote few words

FREQUENCY	RATIO OF LEXEMES WITH POOR PERFORMERS (%) n = 858 → 100%	RATIO OF LEXEMES IN THE FULL CORPUS (%) n = 7397 → 100%
n < 3	8.28	8.93
3 ≤ n ≤ 19	23.19	28.06
20 ≤ n ≤ 39	13.52	16.76
40 ≤ n < 79	17.02	18.16
n ≤ 80	38.00	28.10

3.1.1. A “typology” of word associations

Word associations are a peculiar kind of lexical access (both in terms of the process and of its results) that come into being as a result of phonological, semantic, or structural relationships or their combinations. The system we established on the basis of the relevant literature and our own material is the following.

1. Phonological relationships
 - (a) homonymy
 - (b) partial phonological identity
2. Structural relationships
 - (a) suffixation
 - (b) compounding
 - (c) phrases
3. Semantic relationships
 - (a) coordination
 - (b) subordination/superordination
 - (c) part-of-speech identity

3.1.1.1. Phonological identity may concern the full word form (this is the case of homonymy), or part of it (its initial or final segment, sequence, or syllable). In our material, the following subtypes of phonological similarity can be distinguished.

(a) Adjacent words begin with the same CV sequence (plus may or may not exhibit a coordinative semantic relationship), e.g. *baba* ‘doll’ – *baka* ‘infantryman’, *család* ‘family’ – *csalán* ‘nettle’, *fotel* ‘armchair’ – *fogas* ‘coat-stand’, *fuvola* ‘flute’ – *furulya* ‘recorder’. The effect of a semantic relationship can be modified by phonological similarity (this is what Jakobson exemplifies by the preference of *horrible Harry* over *terrible Harry*, cf. Jakobson 1969, 222). Examples where a monosyllabic word triggers a similar-sounding longer item in which even the original syllable structure may be eliminated also belong here: *csat* ‘buckle’ – *csattog* ‘clatter’, *hód* ‘beaver’ – *hódol* ‘pay homage’, *öt* ‘five’ – *ötlet* ‘idea’, *pad* ‘bench’ – *padló* ‘floor’, or the same relationship working ‘in reverse’: *sörét* ‘small shot’ – *sör* ‘beer’.

(b) Adjacent words end in the same VC/CV sequence or share more material at the end (again, with or without a semantic relationship), e.g., *fotel* ‘armchair’ – *hotel* ‘id.’, *blöki* ‘pooch’ – *löki* ‘he pushes it’, *gatyá* ‘pants’ – *kutyá* ‘dog’, *vipera* ‘adder’ – *opera* ‘id.’, *terület* ‘area’ – *kerület* ‘district’. In most cases, the pairs consist of an identical number of syllables, exceptions: *szemöldök* ‘eyebrow’ – *köldök* ‘navel’, *ajtókilincs* ‘door-handle’ – *bilincs* ‘handcuffs’, *leopárd* ‘leopard’ – *gepárd* ‘cheetah’, *billentyű* ‘key’ – *kesztyű* ‘glove’, *moly* ‘moth’ – *zsámoly* ‘hassock’.

3.1.1.2. In **structural relationships**, again, additional semantic strategies may also be at work. The data that we classified as belonging here are the ones in which we think structure was the dominant basis of association. Collocation is a relatively frequent strategy whereby words that usually occur in sequence in normal speech trigger each other’s appearance in the list (in Hungarian, these are usually attributive phrases or compounds). For instance: *ravasz róka* ‘sly fox’, *örök szerelem* ‘eternal love’, respectively *pók* ‘spider’ – *pókháló* ‘cobweb’.

(a) The association strategy based on suffixation may take four different forms. (i) Stem followed by derived word: *fagy* ‘frost’ – *fagyi* ‘ice cream’, *harc* ‘fight’ – *harcos* ‘fighter’. (ii) Derived word followed by its stem: *hatványozás* ‘raising to a higher power’ – *hatvány* ‘power (of a number)’, *ruházat* ‘clothing’ – *ruha* ‘clothes’. (iii) Association based on shared stem: *borító* ‘cover’ – *boríték* ‘envelope’, *ellenőrző* ‘report book’ – *ellenőrzés* ‘supervision’. (iv) Association based on shared suffix: *szervetlen* ‘inorganic’ – *ehetetlen* ‘inedible’ – *verhetetlen* ‘unbeatable’, *lábatlan* ‘footless’ – *álmatlan* ‘sleepless’ – *vágyatlan* ‘undesired’.

(b) With respect to compounds, several subtypes are found: (i) Simple → compound, the basis of association becomes the anterior member: *barát* ‘friend’ – *barátnő* ‘girlfriend’, *cipő* ‘shoe’ – *cipőfűző* ‘shoelace’. (ii) Simple → compound, the basis of association becomes the posterior member: *csuka* ‘pike’

– *focicsuka* ‘soccer shoes’, *szivacs* ‘sponge’ – *tengeri szivacs* ‘sea sponge’, *szülők* ‘parents’ – *nagyszülők* ‘grandparents’. (iii) Compound → simple, the anterior member is retained: *asztalláb* ‘table leg’ – *asztal* ‘table’, *pénztárca* ‘wallet’ – *pénz* ‘money’. (iv) Compound → simple, the posterior member is retained: *állatkert* ‘zoo’ – *kert* ‘garden’, *rövidnadrág* ‘shorts’ – *nadrág* ‘trousers’. (v) Both words are compound, sharing their anterior member: *Csipkerózsika* ‘Sleeping Beauty’ – *csipkebogyó* ‘rose-hips’, *fogkefe* ‘tooth-brush’ – *fogkrém* ‘toothpaste’; verbs sharing the same preverb are listed also in this category: *megáll* ‘stop’ – *meglátogat* ‘visit’ – *megcsal* ‘cheat on’ – *meghív* ‘invite’. (vi) Both words are compound, sharing their posterior member: *dióbél* ‘walnut kernel’ – *vakbél* ‘caecum’, *sípcsont* ‘shin-bone’ – *lapos csont* ‘flat bone’ – *kulcscsont* ‘collar-bone’ – *lábszárcsont* ‘shank-bone’.

(c) Phrases are closely related to compounds, they are usually attributive or adverbial phrases (occasionally the children divide them by commas in their lists): *szőke nő* ‘blond woman’, *nemi szerv* ‘genital organ’, *illatos kölni* ‘scented perfume’. Predicative phrases and consequence relationships also occur: *tanár* ‘teacher’ – *szid* ‘scold’, *csoki* ‘chocolate’ – *olvad* ‘melt’, *harap* ‘bite’ – *vicsorit* ‘snarl’ – *kutya* ‘dog’ – *macska* ‘cat’, *gólya* ‘stork’ – *csecsemő* ‘baby’. Their connection with compounds is further shown by associative pairs that are members of an existing compound: *harisnya* ‘stockings’ – *cipő* ‘shoe’ – *kanál* ‘spoon’ (cf. *cipőkanál* ‘shoe-horn’) – *jég* ‘ice’ – *korong* ‘disk’ – *foci* ‘soccer’ (cf. *jégkorong* ‘ice hockey’), *fizika* ‘physics’ – *óra* ‘hour’ – *rend* ‘order’ – *őr* ‘guard’ – *irodalom* ‘literature’ (cf. *fizikaóra* ‘physics class’, *órarend* ‘timetable’, *rendőr* ‘policeman’).

Taking the number of phonological and structural associations together (1461, or 2.8% of the whole corpus) to be 100%, their relative proportions are 38% vs. 62%, that is, structural relationships are more frequent. In the case of phonological associations, shared word beginnings are a lot more frequent, 31% out of the total 38%. Among structural associations, compounds are more frequent than derivations, the latter making up 5% of the cases, whereas the former make up 57%. Among compounds, the *cipő* ‘shoe’ – *cipőfűző* ‘shoelace’ type is the most frequent, and the *állatkert* ‘zoo’ – *kert* ‘garden’ type is the least frequent. Phonological associations occurred with 68.75% of the children (two items on average, with one child who wrote 17). Morphological associations occurred with 75.5% of the children (2.5 items on average).

3.1.1.3. Among **semantic relationships**, the most general type in free word association is usually taken to be coordination (Aitchison 1987, 74); the point here is that the associatum represents the same level of generality as the trigger,

e.g., *só* ‘salt’ and *bors* ‘pepper’ or *nap* ‘sun’ and *hold* ‘moon’. Opposites like *fekete* ‘black’ and *fehér* ‘white’ or *nappal* ‘day’ and *éjszaka* ‘night’ also belong here. Subordination is also attested; this is where the associatum is in a genus/species relationship with what precedes it, e.g., *madár* ‘bird’ – *gólya* ‘stork’ or *szín* ‘colour’ – *zöld* ‘green’. A similar relationship “in the opposite direction” can be called superordination, e.g., *puska* ‘rifle’ – *fegyver* ‘weapon’. The occurrence of synonyms is rather infrequent, e.g., *kövér* ‘fat’ – *hájás* ‘obese’, *kerékpár* ‘bicycle’ – *bicó* ‘bike’, *mikulás* – *télapó* (both: ‘Santa Claus’), *rohan* ‘rush’ – *szalad* ‘run’, *együttes* ‘band’ – *zenekar* ‘orchestra’.

(a) Coordination will be defined as a relationship between items in a semantic field that are of the same level of generality. Metonymical and situation-based associations are also classified as belonging here. Examples: *blúz* ‘blouse’ – *nadrág* ‘trousers’ – *cipő* ‘shoes’ – *zokni* ‘socks’, *tojás* ‘egg’ – *embrió* ‘embryo’; spatial contiguity: *kémény* ‘chimney’ – *füst* ‘smoke’, *szánkó* ‘sledge’ – *hó* ‘snow’, *pattanás* ‘pimple’ – *arc* ‘face’, *cseresznye* ‘cherry’ – *kukac* ‘maggot’. A specific subcase of spatial contiguity based on the test situation is the one in which the children list objects they can see in the classroom: *pad* ‘school-bench’ – *szék* ‘chair’ – *asztal* ‘desk’ – *konnektor* ‘wall socket’ – *könyv* ‘book’ – *füzet* ‘exercise-book’. Part-whole relationship: *háló* ‘net’ – *kapu* ‘goal’, *monitor* ‘id.’ – *számítógép* ‘computer’, *autó* ‘car’ – *rendszám* ‘registration number’; ‘made of’ or causal relationship: *bor* ‘wine’ – *szőlő* ‘grapes’, *álmosság* ‘sleepiness’ – *elalvás* ‘falling asleep’, *csúszik* ‘be slippery’ – *jeges* ‘icy’. Opposite or complementary pairs: *hülye* ‘stupid’ – *okos* ‘clever’, *tűz* ‘fire’ – *víz* ‘water’, *haver* ‘chum’ – *ellenség* ‘enemy’, *gyerek* ‘child’ – *felnőtt* ‘adult’, *anya* ‘mother’ – *apa* ‘father’.

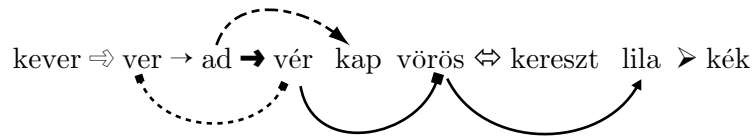
(b) Examples of sub/superordinate terms: *rovar* ‘insect’ – *légy* ‘fly’, *zöldség* ‘vegetable’ – *répa* ‘carrot’, *szervek* ‘organs’ – *szív* ‘heart’, *hajó* ‘ship’ – *Titanic*, *bolygó* ‘planet’ – *Merkúr* ‘Mercury’; respectively *puska* ‘rifle’ – *fegyver* ‘weapon’, *galóca* ‘amanita’ – *gomba* ‘mushroom’, *zokni* ‘socks’ – *fehérnemű* ‘underwear’.

A separate category could be called “other associative relationships”. This includes cases that do not, strictly speaking, fit into any of the above categories but are motivated. For instance, a list of plants may include kinds of fruit: *fa* ‘tree’ – *nárcisz* ‘narcissus’ – *narancs* ‘orange’; or a list of colours may shift into a list of fruit when ‘orange’ comes up: *barna* ‘brown’ – *lila* ‘violet’ – *narancs* ‘orange’ – *citrom* ‘lemon’.

(c) The category of lists based on part-of-speech identity includes series of adjectives or verbs. No obvious conceptual relationship can be found within each pair of items: (*puska* ‘rifle’ –) *lő* ‘shoot’ – *beszél* ‘speak’ – *gondolkozik*

'ponder', *ül* 'sit' – *áll* 'stand' – *mosogat* 'wash up' – *tévézik* 'watch television', (*busz* 'bus' –) *utazni* 'to travel' – *sétálni* 'to walk' – *fagyizni* 'to have an ice cream' – *nyal* 'lick' – *nézni* 'to watch' – *bámulni* 'to stare' – *olvasni* 'to read'; respectively *szép* 'nice' – *jó* 'good' – *okos* 'clever' – *ügyes* 'skilful' – *olasz* 'Italian', *hibbant* 'slightly cracked' – *mulya* 'doltish' – *ifjú* 'young' – *leleményes* 'inventive'.

A specific feature of associative access is that associative pairs/series of diverse types may be embedded into one another, resulting in a decrease of immediate relationships. For instance (glosses: mix, beat, give, blood, get, red, cross, violet, blue):



We have selected the twenty top-scoring children (i.e., five boys and five girls in both age groups who wrote the largest number of mental words) and looked at their numerical data of the various types of association. The twenty children wrote a total of 4168 mental words. Taking 4168 to be 100%, we could establish that in 41.5% of the cases there was no relationship belonging to any of the above categories between adjacent words. The phonological and structural relationships (that were analysed with respect to the full corpus above) made up 1.3% and 2.3% of this sample. In the case of phonological relationships, we only counted cases of pure similar sounding; if that was coupled with some kind of conceptual relationship, we counted the case at hand as an instance of the latter. The category that turned out to be the most populous was coordination (32.5%), followed by subordination (0.7%) and superordination (0.4%). Phrases were a full 1%. Within that sample of 4168 mental words, 58.5% of adjacent words showed some observable pairwise relationship. Taking the latter set to be 100%, the following proportions were found for the individual types. Coordination: 59.3%, opposites: 5.2%, sub- or superordination: 1.9%, synonyms: 3.5%, other associative links: 7%, situational relationships 10.8% (of these, some 2.8% were made up by objects seen in the classroom), all these add up to 88.5%. Part-of-speech identity: 6.6%, compounds: 1.9%, attributive and adverbial relationships: 2.1%, predicative relationship: 0.7%, other: 0.3%.

The associative lists also included **repetitions**, whose occurrence was subject-dependent. The total number of these was 695 (cases where the same child wrote the same word twice, occasionally three times). Compared to the total number of activated words, this is negligible, a mere 1.3%. What is more

interesting is the number of children who repeated words at all. 271 children, 67.25% of the subjects repeated at least one word, with no statistical difference between girls and boys or sixth-formers and seventh-formers. The question arises of whether the number of repetitions correlates with total performance. According to our calculations, it does: the more words one child wrote the more chance there was for a few repetitions to crop in. Hence, the repeated occurrence of the same lexical item was not a sign of accessing failures but rather a side effect of good performance. Repetitions mainly concerned objects in the classroom or school concepts. The lexemes repeated the most often (with the number of repetitions in parentheses) were *óra* ‘hour/watch/clock/school class’ (21), *tábla* ‘blackboard’ (12), *kép* ‘picture’, *könyv* ‘book’, *virág* ‘flower’ (11), *cipő* ‘shoes’ (10), *fa* ‘tree’, *toll* ‘pen’ (9), *ceruza* ‘pencil’, *kabát* ‘coat’, *radír* ‘eraser’, *víz* ‘water’ (8), *szék* ‘chair’, *üveg* ‘glass’ (7), and *haj* ‘hair’, *pad* ‘bench’ (6). It appears that whenever the children got into a dead end in the course of free associations, it was the objects of their immediate physical environment that served as clues to help the series of associations move on.

Several characteristics of repetitions can be observed that may refer to the explanation (or cause) of their occurrence. The same word form may appear in different meanings within different association environments (this amounts to a 5–6% of cases). For instance: *hold* ‘moon’, *nap* ‘sun’, *fény* ‘light’; *hónap* ‘month’, *nap* ‘day’, *év* ‘year’; or *kenyér* ‘bread’, *kosár* ‘basket’, *túrórudi* ‘curd dessert’; *foci* ‘soccer’, *kosár* ‘basket(ball)’, *jégkorong* ‘ice hockey’; or *festmény* ‘painting’, *rajz* ‘drawing’, *vetítő* ‘projector’; *éneklés* ‘singing (lesson)’, *rajz* ‘drawing (lesson)’, *biológia* ‘biology’; or *tanterem* ‘classroom’, *óra* ‘lesson’, *osztály* ‘class’; *perc* ‘minute’, *óra* ‘hour’, *kevés* ‘too few’; *gyönyörű* ‘beautiful’, *óra* ‘(wrist)watch’, *karkötő* ‘bracelet’; or *szív* ‘heart’, *lép* ‘lien’, *máj* ‘liver’; *lépés* ‘step’ (noun), *lép* ‘step’ (verb), *megy* ‘go’; or *vasárnap* ‘Sunday’, *egy* ‘one’, *kettő* ‘two’; *tető* ‘roof’, *egy* ‘an’, *öregember* ‘old man’; or *irodalom* ‘literature’, *föld* ‘ground’, *rajzfilm* ‘cartoon’ (cf. *földrajz* ‘geography’); *Mars* ‘id.’, *Föld* ‘Earth’, *Jupiter* ‘id.’. Homonyms sometimes occur right in a row: *út* ‘road’, *daru* ‘crane’, *daru* ‘derrick’, *lyuk* ‘hole’; *egér* ‘mouse’, *nyúl* ‘rabbit’, *nyúl* ‘reach for’, *róka* ‘fox’; *menekül* ‘flee’, *ég* ‘burn’, *ég* ‘sky’, *bevásárlóközpont* ‘shopping mall’. First names are also often repeated as if each classmate with the same name appeared separately in the lists. Another factor that might contribute to repetitions is the effect of associative relationships of diverse types. For instance, synonymy and coordination: *eb* ‘dog’, *kutya* ‘dog’, *szőnyeg* ‘rug’ + *ló* ‘horse’, *kutya* ‘dog’, *tábla* ‘blackboard’; coordination and antonymy: *egér* ‘mouse’, *fekete* ‘black’, *párduc* ‘panther’ + *fehér* ‘white’, *fekete* ‘black’, *gyors* ‘fast’; phrase and coordination: *szép* ‘nice’, *kék* ‘blue’, *virág* ‘flower’ + *barna*

'brown', *kék* 'blue', *piros* 'red'. Uncategorizable, probably individually explainable repetitions include *élet* 'life', *halál* 'death', *fent* 'high up' + *karambol* 'collision', *halál* 'death', *szenvedés* 'suffering'; *műt* 'operate on', *halál* 'death', *temet* 'bury' + *pap* 'priest', *halál* 'death', *fáj* 'it hurts'; *irodalom* 'literature', *felmérés* 'written exam', *lesés* 'peeping' + *felelés* 'oral exam', *felmérés* 'written exam', *folyók* 'rivers'; *szeretet* 'love', *szív* 'heart', *halál* 'death' + *boldogság* 'happiness', *szív* 'heart', *vers* 'poem'; *haj* 'hair', *mell* 'chest', *műszer* 'instrument' + *pénisz* 'penis', *mell* 'bust', *hegy* 'tip'. Often no explanation can be found for repetition; in such cases it is difficult to tell if the child was aware of the fact that (s)he was repeating a word. Examples: *füzet* 'exercise-book', *ceruza* 'pencil', *tolltartó* 'pencilbox' + *radír* 'eraser', *ceruza* 'pencil', *ellenőrző* 'report book'; *kulcs* 'key', *táska* 'bag', *papír* 'paper' + *filc* 'felt(-tip pen)', *táska* 'bag', *könyv* 'book'; *pulóver* 'sweater', *nadrág* 'trousers', *póló* 'jersey' + *dolgozat* 'test', *nadrág* 'trousers', *cipő* 'shoes', etc.

3.1.2. Part-of-speech categories

The corpus contains data from practically all part-of-speech categories, although the various word classes are of course represented very unequally. It is interesting to note, first, that it includes 41 English words, e.g., *episode*, *thriller*, *enter*, *transport*, *skate*, *city*, *alien*, *market*, *yellow*. These will be excluded in what follows, even though some of them may be in the process of domestication (*thriller* has perhaps become a loanword in Standard Hungarian). Part-of-speech distribution will be analysed in terms of **mental words** (total: 52,184), then in terms of **lexemes** (total: 7397, including proper names).

In investigating the association strategies of the children, we took non-finite verb forms to be suffixed forms (even though they appear to be regarded by children as simple, morphologically unanalysable forms). The various types of non-finite verb forms (abbreviated as *nf* in figures and tables below) occurred in the following numbers: 167 infinitives, 13 "adjectival" participles ('one that X-es/is X-ing', e.g., *ugráló* 'jumping', *szelő* 'cutting', *vágyakozó* 'wishing', *focizó* 'playing football', *lehulló* 'falling', *húzó* 'pulling', *viszահúzódzkodó* 'withdrawing', *faxoló* 'sending a fax'), 7 "adverbial" participles ('in an X-ing manner, while X-ing, having X-ed', e.g., *körülírva* 'paraphrasing', *rajzolva* 'drawing', *megbántva* 'hurting', *futva* 'running'). Most words that could be adjectival participles as well are included in the categories 'adjective' or 'adjective/noun' (e.g., *lökött* 'loony' (lit. 'pushed'), *üdítő* 'soft drink' (lit. 'refreshing'), *fényképező* 'camera', *fűző* 'corset'/'shoelace' (lit. 'lacing'), *erősítő* 'amplifier' (lit. 'strengthening'), *lopakodó* 'stealth bomber' (lit. 'one who walks stealthily'). The item *ugráló* 'jumping' could mean 'skipping-rope', and *bukó*

‘falling/diving’ could mean ‘a pupil about to fail a school subject’, it is difficult to tell out of context. Part-of-speech labels that are asterisked stand for ambiguous items (noun* = noun/adverb, e.g., *éjszaka* ‘night/at night’, *föl* ‘skimmings/up’, *hétköznap* ‘weekday/on weekdays’, *korán* ‘Koran/early’, *reggel* ‘morning/in the morning’, *vasárnap* ‘Sunday/on a Sunday’, *otthon* ‘home/at home’, *haza* ‘motherland/(go etc.) home’; adv* = adverb/conjunction, e.g., *így* ‘in this way/hence’ or adverb/postposition, e.g., *alá* ‘(to) below’, *mögé* ‘(to) behind’, *mellé* ‘(to) beside’, *kívül* ‘outside/in addition to’, or adverb/pronoun e.g., *arra* ‘in that direction/onto that’, *ki* ‘out/who’; num* = numeral, including *hat* ‘six/make an effect’ and *hét* ‘seven/week’).

The category ‘other’ includes postpositions, articles, conjunctions, interjections (e.g., *hoppá* ‘oops’, *nyekk* ‘sound of beeing floored’, *kuss* ‘shut up!’), and modifiers, as well as a few items whose categorisation is debatable (e.g., *szia* ‘hi’, *mizújs* ‘what’s up’, *kösz* ‘thanks’, *anno* ‘a long time ago’). The part-of-speech distribution of mental words and lexemes in the full corpus is shown in Figures 2 and 3, respectively.

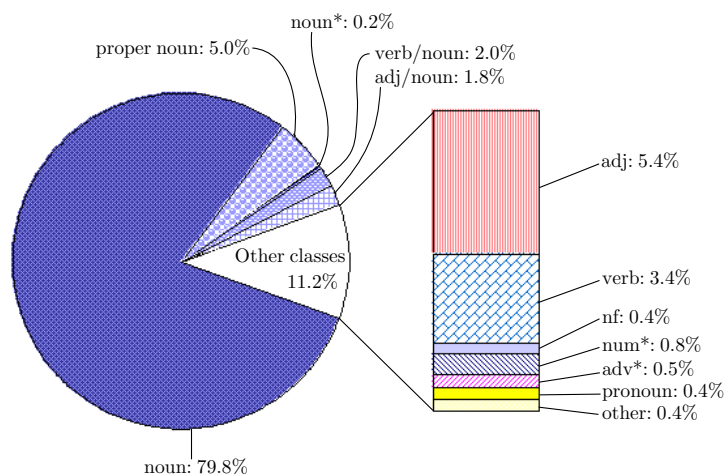


Fig. 2

The part-of-speech distribution of mental words in percentage

We have also analysed age and gender differences in mental words and in lexemes (Table 2).

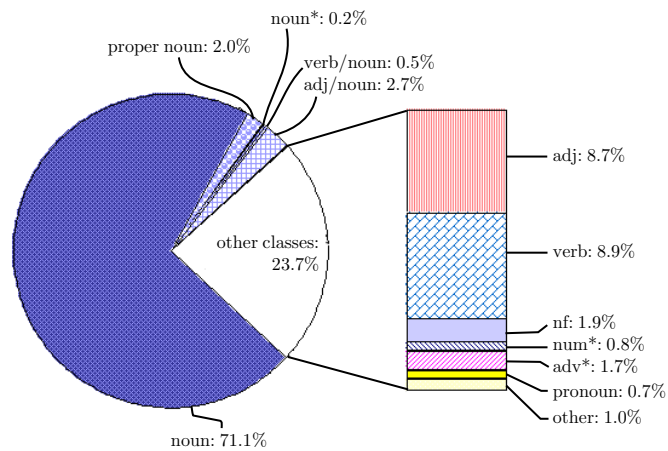


Fig. 3

The part-of-speech distribution of lexemes in percentage

Table 2

Part-of-speech distribution of mental words and lexemes in terms of age groups and genders

	12-YEAR-OLDS		13-YEAR-OLDS		BOYS		GIRLS	
	Mental words	Lexemes	Mental words	Lexemes	Mental words	Lexemes	Mental words	Lexemes
noun	80.3%	74.6%	79.2%	72.4%	78.8%	73.4%	80.6%	72.5%
proper noun	5.3%	1.5%	4.8%	1.4%	6.2%	1.8%	4.0%	1.1%
noun*	0.2%	0.2%	0.3%	0.2%	0.2%	0.2%	0.3%	0.3%
verb/noun	2.0%	0.6%	2.0%	0.6%	2.1%	0.6%	2.0%	0.6%
adj/noun	1.6%	2.8%	1.9%	2.6%	1.9%	3.0%	1.7%	2.5%
adj	5.4%	8.1%	5.3%	7.7%	5.0%	8.3%	5.7%	7.6%
verb	2.9%	7.4%	3.8%	8.9%	3.0%	6.9%	3.7%	9.6%
nf	0.1%	0.5%	0.6%	2.0%	0.4%	1.6%	0.4%	1.6%
num*	1.0%	1.1%	0.5%	0.6%	1.0%	0.8%	0.6%	0.9%
adv*	0.4%	1.5%	0.6%	1.7%	0.5%	0.8%	0.4%	0.8%
pronoun	0.3%	0.7%	0.6%	0.8%	0.6%	1.5%	0.4%	1.7%
other	0.4%	0.9%	0.4%	0.9%	0.5%	1.0%	0.3%	0.8%

The proportions of lexemes are similar to those of mental words. The largest differences are with common nouns and proper nouns (relatively fewer lexemes) and with verbs (relatively more lexemes). This is in accordance with the fact

that nouns are the most frequent words in the lists. Proper nouns constitute 5% of all mental words (6.18% of all nouns, common nouns and proper nouns added up). Most proper nouns are first names, probably those of concrete persons (e.g., classmates), but full personal names also occur. The occurrence of geographical names and titles was as expected, the latter mainly include titles of television programmes, films, less frequently those of literary pieces. Brand names occurred in surprisingly large numbers, a total of 121 with 12-year-olds and a total of 176 with 13-year-olds. Boys activated more than twice as many brand names as girls did, the difference is significant in both age groups. Boys wrote 3.55, whereas girls merely wrote 1.7 such words on average. The most populous groups of brand names were car models and/or Formula 1 stables, e.g., *Fiat*, *Ferrari*, *Porsche* (37.8% of all occurring brand names), soft drinks and other consumables, e.g., *Pepsi*, *Fanta*, *Milka* (15.5%), sports equipment producers, e.g., *Adidas*, *Nike*, *Fila* (15.2%). The word *rotring* was the single most frequent brand name, suggesting that it is becoming a common noun. The relatively frequent occurrence of brand names witnesses their common use, but it is also a sign of a change of the vocabulary, indicating the emergence of new lexemes.

260 of the four hundred subjects wrote one or more verbs, each of them wrote 7.28 verbs on average. There were 95 children who wrote a single verb each. Seventh-formers wrote significantly more verbs than sixth-formers, and girls wrote significantly more than boys did. It can be assumed that for the younger age group *word* ('onoma') equals *name*, meaning primarily nouns, to a larger extent. Wherever a verb occurs, another one usually follows suit. Disregarding those who wrote a single verb in their whole lists, the average number of verbs that immediately follow the first verb is 1.45. The longest sequence of verbs (consisting of 16 items) was written by a 12-year-old girl: (*tallér* 'thaler'), *dolgozik* 'work', *tévézik* 'watch television', *alszik* 'sleep', *eszik* 'eat', *iszik* 'drink', *dohányzik* 'smoke', *rágózik* 'chew gum', *ír* 'write', *rajzol* 'draw', *fest* 'paint', *másol* 'copy', *olvas* 'read', *néz* 'watch', *kér* 'ask for', *kap* 'get', *örül* 'be glad', (*krém* 'cream') – in this case, we arbitrarily took *ír* 'write/Irish(man)' to be a verb.

The number of lexemes classifiable both as nouns and as verbs was 35 (=1133 mental words). Listed in an order of decreasing frequency, they were the following: *fej* 'head/milk (e.g., a cow)', *fal* 'wall/devour', *ing* 'shirt/wobble', *nyúl* 'rabbit/reach for', *fog* 'tooth/hold', *csap* 'tap/strike', *ír* 'Irish(man)/write', *szív* 'heart/inhale', *tűz* 'fire/(fasten with a) pin', *csavar* 'screw/twist', *ég* 'sky/burn', *nő* 'woman/grow', *dob* 'drum/throw', *terem* 'hall/yard', *áll* 'chin/stand', *zár* 'lock/shut', *vár* 'castle/wait', *sír* 'grave/cry', *ér* 'blood vessel/be worth',

szeg 'nail/hem', *volt* 'volt/was', *fél* 'half/be afraid', *él* 'edge/live', *lép* 'lien/step', *fagy* 'frost/freeze', *nyom* 'trace/push', *sejt* 'cell/guess', *nyű* 'maggot/wear down', *les* 'offside/peep', *szán* 'sleigh/pity', *követ* 'envoy/follow', *zavar* 'confusion/disturb', *szignál* 'signature tune/sign', *tett* 'deed/did', *varrat* 'seam/have a dress made'. The fundamental question in this respect is whether it is the nominal or the verbal meaning that is activated; also, whether something suggests that both meanings may be activated for a single form (about multiple activation of ambiguous lexical items in perception cf. Pléh 2000). An associative link may be present with the previous word, the following word, or both, or indeed neither. There were cases in which a mental word of the type verb/noun (or verb/noun/adjective for *volt* 'was/volt/former' and *ír* 'write/Irishman/Irish') was found in a sequence where there was a link with the previous word but a sudden semantic shift occurred between the word at hand and the following one: *sétál* 'walk', *ül* 'sit', ***áll*** 'stand/chin', *csend* 'silence'; *nap* 'sun', *felhő* 'cloud', ***ég*** 'sky/burn', *emelet* 'storey'; *kabát* 'coat', *dzseki* 'jacket', ***ing*** 'shirt/wobble', *ágy* 'bed'-this was the case in 27.7% of such sequences. With the following word, there was a clear connection in 25.9% of the cases: *ügyes* 'skilful', *ügyetlen* 'clumsy', ***áll*** 'chin/stand', *ül* 'sit'; *alma* 'apple', *kivi* 'kiwi', ***szív*** 'inhale/heart', *bél* 'bowels'. In 26.7% of the instances, connections were observable both ways: *könyök* 'elbow', *csukló* 'wrist', ***áll*** 'chin/stand', *száj* 'mouth'; *gatyá* 'pants', *ruha* 'dress', ***ing*** 'shirt/wobble', *cipő* 'shoes'. A preceding activation can influence the subject's choice at a later point, too. For instance, the sequence *levegő* 'air', *víz* 'water', ***csap*** 'tap/strike', *föld* 'earth' may be based on the collocation *víz, csap* 'water, tap' interrupting the coordinative series 'air, water, earth', i.e., the 'elements' of ancient and medieval natural history.

In some cases, words of several more or less contiguous conceptual fields are accessed almost alternately, e.g., *mosdószivacs* 'toilet-sponge', *szappan* 'soap', ***fürdőkádb*** 'bathtub', *zokni* 'socks', *sampon* 'shampoo', *papucs* 'slippers', ***csap*** 'tap', *ruházat* 'clothing', *törülköző* 'towel'; *zokni* 'socks', *zsiráf* 'giraffe', ***ing*** 'shirt', *nyuszi* 'bunny', ***papucs*** 'slippers'; *hangya* 'ant', *sün* 'hedgehog', *beszél* 'talk', *ír* 'write', *nyúl* 'rabbit', *ól* 'sty'. Coordinative relationships may be intertwined with collocation-based associations: *karóra* 'wrist-watch', *csavar* 'screw', ***anya*** 'mother', *apa* 'father', ***anyacsavar*** 'screw nut', ***csavarhúzó*** 'screwdriver', *szerelés* 'installation'; *tábla* 'blackboard', *kocka* 'cube', ***fal*** 'wall', *dobókocka* 'dice'. It is impossible to tell what might cause a mental detour like *gólya* 'stork', *macska* 'cat', *kutya* 'dog', ***táska*** 'bag', *nyúl* 'rabbit', *zebra* 'id.', *kígyó* 'snake'.

Structural relationships with preceding/following items are less clear indicators of which meaning is probably present. In particular, they are incapable of excluding the irrelevant reading: *kalapács* ‘hammer’, **fogó** ‘pliers/holder’, **fog** ‘hold?/tooth?’, *iroda* ‘office’; *lepedő* ‘bed-sheet’, **háló** ‘bed-room’, **ing** ‘shirt?/wobble?’, *magas* ‘tall’ (cf. *hálóing* ‘night-dress’); **szőke** ‘blond’, **nő** ‘woman?/grow?’, *nemi szerv* ‘genital organ’. Similarly, with primarily phonological relationships, it is uncertain which meaning is activated (both?) e.g., *körömlakk* ‘nail varnish’, **fok** ‘degree’, **fog** ‘hold?/tooth?’, *vilanykörte* ‘light-bulb’; *nyál* ‘saliva’, *nyelv* ‘tongue’, **nyúl** ‘rabbit?/reach out?’, *nyel* ‘swallow’; *fekete* ‘black’, *szín* ‘colour’, *szív* ‘heart?/inhale?’, *kirándulás* ‘excursion’. In the sequence *fű* ‘grass’, *virág* ‘flower’, *les* ‘peep?/offside?’, *foci* ‘soccer’, it is likely that *foci* is triggered by *les* as a noun, but this does not necessarily prove that it was only the nominal reading of *les* that was activated.

The frequency values of the above 35 ambiguous lexemes are diverse; *fej* ‘head/milk’, *fal* ‘wall/devour’, *ing* ‘shirt/wobble’, and *nyúl* ‘rabbit/reach for’ are among the most frequently occurring items, belonging to the ‘80 or more’ range, whereas *szignál* ‘signature tune/sign’, *tett* ‘deed/did’, *zavar* ‘confusion/disturb’, *követ* ‘envoy/follow’, and *varrat* ‘seam/have a dress made’ only occur once. We have made a calculation of the proportions of the various readings of ambiguous lexemes occurring at least twenty times. The category “impossible to tell” collects items whose environment does not disambiguate them. Examples: *táska* ‘bag’, *térkép* ‘map’, **ing** ‘shirt/wobble’, *ajtó* ‘door’; *gólya* ‘stork’, *kökörcsin* ‘meadow anemone’, **vár** ‘castle/wait’, *Lengyelország* ‘Poland’. This is where phonologically-based occurrences are also included. Table 3 furthermore contains frequency values taken from Füredi–Kelemen (1989). Vacant rubrics mean that the given part-of-speech reading did not occur in the material of that frequency dictionary at least ten times.

Comparing our data with those of the frequency dictionary, the items *csap*, *csavar*, *fog*, *nyúl*, *zár* appear to be interesting in that part-of-speech frequency shows the opposite tendency in the two sets of data. With the children, “importance” matters, and also whether objects can be directly seen around them. For some lexemes—*fej*, *fal*, *ír*—several factors seem to conspire to achieve the attested asymmetrical results. On the other hand, some items clearly support multiple activation, even though the given experimental setup does not make it possible for us to decide if these are cases of simultaneous or retrospective activation (cf. Pléh 2000, 968). This is because the given item is available for an unlimited amount of time in the case of a written task, therefore “multiple recognition” is possible.

Table 3

Frequency of occurrence of ambiguous words (*main verb, **auxiliary)

LEXEME	NOMINAL MEANING (%)		VERBAL MEANING (%)		IMPOSSIBLE TO TELL (%)	FREQUENCY VALUES OF FÜREDI-KELEMEN (1989)	
						NOUN	VERB
áll	chin	13.79	stand	68.97	17.24	30.48	820.17
csap	tap	73.61	strike	2.78	23.61	–	45.68
csavar	screw	53.06	twist	2.04	44.90	–	11.83
dob	drum	32.50	throw	10.00	57.50	–	39.79
ég	sky	42.86	burn	30.61	26.53	101.3	64.88
fal	wall	70.64	devour	0.92	28.44	214.03	—
fej	head	80.18	milk	0.90	18.92	664.07	—
fog	tooth	61.64	hold	8.22	30.14	53.18	163.45*
			will				341.41**
ing	shirt	61.18	wobble	1.18	37.65	69.30	–
ír	Irish	1.41	write	95.77	2.82	–	164.82
nő	woman	90.70	grow	0.00	9.30	347.6	46.68
nyúl	rabbit	75.29	reach	2.35	22.35	–	65.34
sír	grave	19.05	cry	57.14	23.81	44.82	112.02
szív	heart	62.69	inhale	2.99	34.33	175.63	32.26
terem	hall	81.25	yield	0.00	18.75	63.97	21.78
tűz	fire	74.07	pin	0.00	25.93	70.28	20.45
vár	castle	14.81	wait	22.22	62.96	14.46	464.02
zár	lock	72.41	shut	0.00	27.59	13.9	31.55

One of the cardinal issues pertaining to the structure of the mental lexicon is **the way suffixes are encoded**. According to some hypotheses, the grammatical structure of a language determines the structure of its mental lexicon (Aitchison 1987). It is claimed that in agglutinating languages the relationship between stems and suffixes within the mental lexicon is different from that relationship in non-agglutinating languages. One possible consequence is that suffixed forms are easier to access in languages with rich morphologies. This can be accompanied by the occurrence of bare suffixes, in addition to root morphemes, during word activation. We had anticipated that we would find relatively numerous suffixed lexical items in our corpus. Although such forms did occur, their ratio of occurrence was small. The 12-year-olds listed 419, and the 13-year-olds listed 784 such forms. It is true that in absolute numbers these figures seem to be considerably large and the trend is increasing; however, suffixed forms constitute a mere 2.3% of the whole corpus of mental words. Hence, word activation gives preference to word stems (root morphemes), i.e., children access meanings or concepts, not forms that can be

directly inserted into syntactic structures. The most frequent suffixed forms are plural nouns or adjectives, making up 51% of all suffixed forms; infinitives constitute 15.1%; and all the others occur in the remaining 32.46% of cases.

Inflected word forms occurring ten times or more in the corpus are the following: *állatok* ‘animals’ (21), *színek* ‘colours’ (17), *számok* ‘numbers’ (17), *emberek* ‘people’ (14), *szavak* ‘words’ (13), *betűk* ‘letters’ (11), *gyerekek* ‘children’ (11), *barátok* ‘friends’ (10), *írni* ‘to write’ (10), *csillagok* ‘stars’ (10), *fiúk* ‘boys’ (10). Words that are unambiguously nouns, are not plurals but some other inflected forms, and were written by at least two children are *megszentségteleníthetelenségeskedéseitekért* ‘for your (pl.) repeated pretending to be undeconsecrateable’ (4), *egészségedre* ‘cheers!’ (3), *anyád* ‘your mother’ (an insult) (3), *megszentségteleníthetelenségeskedéseiteket* ‘your (pl.) repeated pretending (acc.) to be undeconsecrateable’ (2), *tesóm* ‘my brother/sister’ (2). Other attempts at being “funny” include *megegészségesedésetekért* ‘for your (pl.) becoming healthy’, *megvezekelésétekért* ‘for your (pl.) repeatedly being punished’ and, among verbs, *elkelkáposztásítottalanították* ‘they deprived it from having been turned into savoy cabbage’. The adj. and adj./noun groups also included plural forms like *rosszak* ‘the bad ones’ (3), *magyarok* ‘Hungarians’ (2), *élők* ‘the living ones’ (1); also, there sporadically occurred other inflected forms like *vörösben* ‘in red’, *jókkal* ‘with the good ones’. Verbs occurred in objective conjugation, in the imperative, etc. (in addition to infinitives mentioned previously): *viszi* ‘carries it’ (8), *veszi* ‘takes it’ (7), *lesz* ‘will be’ (7), *vigyázz* ‘take care!’ (4), *csináld* ‘do it!’ (2), *nézd* ‘look at it!’ (2), *megyek* ‘I go’ (2), *utálok* ‘I hate it’ (2), *utállak* ‘I hate you’ (2), *hívna* ‘they call me’ (2), *bejött* ‘came in’ (2), *gondoskodott* ‘made sure’ (2), *lógni fogsz* ‘you will be hanged’ (1). The only occurring conditional forms were *szeretne* ‘would like to’ (1), *lenne* ‘would be’ (1). Examples of other parts of speech occurring in an inflected form: *többet* ‘more (acc.)’ (1), *nyolcadika* ‘the eighth (day of the month)’ (1), *mit* ‘what (acc.)’ (5), *ezt* ‘this (acc.)’ (3), *ezzel* ‘with this’ (2). The frequency of activation of a suffixed form is not to be confused with its accessibility. Experiments involving reaction time confirmed that there is no difference, in terms of ease of lexical access, between suffixed vs. unsuffixed words (Gósy 1998b).

3.2. Comparing word activations 60 years apart

The question arises of how our own data relate to János Cser’s material recorded more than six decades ago. The identical number of subjects participating in the two experiments makes direct comparisons possible. The total

number of words activated by the “grandparents” of our present-day subjects was 37,912—i.e., they wrote 14,313 words fewer than those tested in 1999. This difference is extremely large, and the reason for it can again be twofold: first, the size of the vocabulary and/or the ease of access may differ, and second, the changing speed of handwriting may be held responsible. While six decades ago an average of 6.3 words were written in a minute, today 8.7 words were. The difference is 2.4 words per minute. This is relatively such a large number that it cannot be simply due to the different speed of handwriting. We had to conclude that—whether or not in an absolute sense—the vocabulary of today’s 12/13-year-olds is probably larger than that of their (great)grandparents, and/or their accessing processes are faster. Another potential factor would be the possibly different strategies employed by the two groups of subjects. János Cser notes that his Hungarian subjects were slower than American and Swiss (probably French-speaking) children tested in similar word association experiments had been. He mentions the following factors as responsible for that difference: “the speed of association, the size of concept stock [vocabulary], the ability to express what they mean, the length of words, and the speed of writing” (Cser 1939, 15). Of these, the length of words—that is undoubtedly a decisive factor—is irrelevant here as the language of the two experiments was the same, thus what remains is the array of causes that we also assumed (albeit in a different wording), i.e., the size of vocabulary and ease of lexical access, as well as (to some extent) the speed of writing.

The factors having to do with the historical change of the language can be illustrated by an analysis of the conceptual areas involved. That analysis is partly one of content and partly one of a statistical nature. In the following comparisons with Cser’s material, relative frequency is invariably computed with respect to the actual number of subjects mentioning that particular category. The analysis encompassed the following aspects: **1.** How dominant is the given conceptual area for the given age group? This may be revealed by (a) what percentage of the 400 subjects wrote words belonging to that area—cf. the second column of Table 4; (b) what percentage of all occurrences is made up by words belonging to that area—cf. the fifth column of Table 4. **2.** Are there “girlish” and “boyish” conceptual areas? Cser does make such a distinction. For instance, he says that boys are interested in fight, technical progress, etc., whereas topics like colours, foodstuffs, etc. are more characteristic of girls. He draws this conclusion from a comparison of the number of frequent words (occurring more than in 5%) of a given conceptual area mentioned by boys, respectively by girls. In our own material, we made the following calculations in this respect. (a) Among subjects who wrote words in the given area, what

is the proportion of girls vs. boys—cf. the third column of Table 4. (b) For all words of all conceptual areas, we performed a two-way ANOVA test for the factors gender and age group. This shows whether there are significant differences by gender or by age group in the number of words written per person. In Table 4, the fourth column shows the average number of words written by those who activated words in the given area. Figures in bold stand for cases where the gender distinction was significant. (c) For frequent words—meaning words that occurred 20 times or more—what are the areas where the difference between girls and boys is 20% or larger. (The calculations in (a) and (b) were based on all occurrences, those in (c) were based only on words that occurred in at least 5%, i.e., 20 times or more.)

Table 4
Conceptual areas, occurrence, distribution, and proportions

CONCEPTUAL AREA	CHILDREN MENTIONING WORDS BELONGING THERE (%)	GIRLS VS. BOYS (%)		AVERAGE NUMBER OF WORDS GIRLS VS. BOYS		PROPORTION OF OCCURRENCE (%)*
Animals	93	50.8	49.2	10.96	10.47	8.02
Plants	90	53.3	46.7	4.05	3.55	2.77
Colours	51	64.2	35.8	4.75	3.92	1.83
Occupations	63.25	53	47	2.31	2.54	1.24
Nationalities	26.25	56.3	43.7	2.13	2.31	0.47
Food	88.75	50.7	49.3	6.22	5.05	5.41
Body parts	84	53.3	46.7	6.10	5.39	3.93
Clothing	95	51.3	48.7	11.02	7.53	7.25
House & household	97	50.5	49.5	16.15	11.38	10.84
Traffic	83.25	47.7	52.3	2.81	3.64	2.18
Engineering	91	49.7	50.3	4.09	4.79	3.26
Nature	87	51.7	48.3	5.59	4.89	3.68
Materials	70	50.6	49.4	2.45	2.74	1.51
Leisure, sports	96.5	51.3	48.7	6.84	6.15	5.06
Tools	78.25	50.2	49.8	3.01	3.15	1.94
School	99.25	50.6	49.4	13.89	11.13	10.03
Curriculum	82.5	50.3	49.7	5.07	4.31	3.1
Man	77.25	53	47	5.18	3.76	2.8
Negative	47.25	50.3	49.7	2	2.49	0.85
Abstract noun	77.75	50.5	49.5	7.29	4.94	3.84

*Taking the whole corpus into consideration.

Our analysis of conceptual areas covered common nouns and “colours” which are adjectives. We started from the categories set up by Cser, and defined 19

conceptual areas in his terms, adding “abstract nouns”. (Of these areas, 18 covered 93.6% of all occurring nouns.)

Colours. The corpus contains a total of 908 colour adjectives identifying 24 different colours: *barna* ‘brown’, *bordó* ‘claret’, *bordópiros* ‘wine red’, *citromsárga* ‘lemon-yellow’, *fehér* ‘white’, *fehéres* ‘whitish’, *fekete* ‘black’, *fekete-fehér* ‘black and white’, *feketés* ‘blackish’, *hupikék* ‘gaudy blue’, *hupilila* ‘gaudy violet’, *kék* ‘blue’, *királykék* ‘royal blue’, *lila* ‘violet’, *narancssárga* ‘orange’, *okkersárga* ‘ochre’, *piros* ‘red’, *rózsaszín* ‘pink’, *sárga* ‘yellow’, *sötétbarna* ‘dark brown’, *sötétzöld* ‘dark green’, *szürke* ‘grey’, *vörös* ‘red’, *zöld* ‘green’ (colours set in bold occurred three or more times; those underlined were only written by girls). 51% of all subjects mentioned some colour or colours. Among those writing colours, there were almost twice as many girls (64.2%) than boys (35.85). The girls wrote an average of 4.75 colours, significantly more than the boys’ average 3.92. All colours occurring 20 times or more were written by girls a lot more times than by boys; the difference is largest for *rózsaszín* ‘pink’, and smallest for *vörös* ‘red’. The frequency order is blue, green, black, red (*piros*), white, yellow, violet, brown, pink, grey, red (*vörös*). No boy mentioned *bordó* ‘claret’.

Girls made more exact distinctions among shades of colour (with respect to sociolinguistic differences between genders in using colour terms, cf. Lakoff 1975, 8). Experimental investigation of Hungarian children and adults also confirmed this conclusion (Gósy 1998a). In Cser’s material, the frequency toplist is different (red, green, blue, white, yellow, black), and shades (non-primary colour terms) are either missing or their frequency of occurrence has grown since (e.g., orange – relative frequency 0.2 → 3, pink – relative frequency 3.8 → 9).

Occupations. The most frequently occurring item is *tanár* ‘teacher’, together with its synonyms (e.g., *tanárnő* ‘lady teacher’, *tanító* ‘primary school teacher’, *pedagógus* ‘pedagogue’, *tanár néni* ‘schoolmistress’). The tendency is similar in Cser’s material. There is no significant gender difference, although girls wrote *orvos/doktor* ‘physician’ more often than boys did, whereas *rendőr* ‘policeman’ occurred more often with boys than with girls.

Nationalities. The full corpus contained 236 occurrences of 33 different names of nationalities. The sweep of *angol* ‘English’ (relative frequency: 2.9 → 13) may reflect a difference between the two historical/social periods. Another interesting difference is that of *olasz* ‘Italian’ (relative frequency: 3 → 1.25). The word *roma* ‘Gypsy’ was not attested in the old material.

Animals. The activation of names of animals was very frequent. In our material the number of all occurrences was 3981, that is, 8.02% of the whole corpus, containing 270 different names of animals. 93% of the 400 children wrote one or more names of animals. The most frequent ones, obviously the prototypical ones, were *kutya* ‘dog’ and *macska* ‘cat’. It is a commonplace that names of animals constitute an important part of children’s vocabularies. According to the two-way ANOVA test, age group is the determining factor, sixth-formers wrote 11.92 names of animals on average, whereas seventh-formers wrote significantly fewer, 9.52. Gender, however, is not a significant factor. Among names of animals occurring 20 times or more, girls more often wrote *delfin* ‘dolphin’, *mókus* ‘squirrel’, *cica* ‘kitten’, *malac* ‘piglet’, *hörcsög* ‘hamster’, *csiga* ‘snail’, whereas boys more often wrote *bika* ‘bull’, *galamb* ‘pigeon’, *őz* ‘roe’, *sas* ‘eagle’. The largest increase of frequency as compared to Cser’s material was found for wild species (in the frequency order of our own corpus: *zsiráf* ‘giraffe’, *delfin* ‘dolphin’, *bálna* ‘whale’, *cápa* ‘shark’, *hörcsög* ‘hamster’, *giliszta* ‘worm’, *pingvin* ‘penguin’, *jaguár* ‘jaguar’, *polip* ‘octopus’). The largest decrease of frequency with respect to Cser’s top list was found for *őszvér* ‘mule’, *szamár* ‘donkey’, *veréb* ‘sparrow’. On the other hand, *teknős* ‘turtle’ (40) and *teknősbéka* ‘tortoise’ (8) do not figure in his material at all.

Plants. Names of parts of plants were also classified as belonging here. 90% of our subjects mentioned plants, but relatively few ones. General terms were frequent: *fa* ‘tree’, *virág* ‘flower’, *bokor* ‘shrub’, *növény* ‘plant’. The most popular flower was *rózsa* ‘rose’, whereas among cereals *kukorica* ‘maize’ and *búza* ‘wheat’ had priority, again confirming the validity of prototype theory. Compared to Cser’s list, an overall decrease of frequency characterises this category, with cereals losing ground to the largest extent. In his material plant names are among the “girlish” words; this is also found in our corpus (gender is a significant factor), although the difference is less spectacular than for colours.

Food. The most frequently mentioned foodstuff was *kenyér* ‘bread’ (85) followed by chocolate (79); with the occurrences of *csoki* ‘chocolate (dimin.)’ (60) and *csokoládé* ‘chocolate’ (19). In the rest of the frequency list, the relatively good position of *répa* ‘carrot’ is somewhat surprising. The most frequently mentioned drinks were *tej* ‘milk’ and *kóla* ‘coke’; among alcoholic beverages, *bor* ‘wine’ scored highest. According to ANOVA, girls consistently wrote more items of food than boys did. The “girlish” food words were *pizza* ‘id.’, *torta* ‘cake’, *étel* ‘food’, *tea* ‘id.’, *cukor* ‘sugar’, *hamburger* ‘id.’, *só* ‘salt’, *vacsora* ‘dinner’, *leves* ‘soup’, *csoki* ‘chocolate’, *ital* ‘drink’, *hús* ‘meat’, *vaj* ‘butter’, *kóla* ‘coke’, *tejföl* ‘sour cream’, *kifli* ‘crescent’, *krumpli* ‘potato’, *bab* ‘bean’,

whereas the single (frequent) food word that boys wrote a lot more often than girls did was *szalámi* 'salami'.

Cser's list obviously lacks a number of types of food as they were unknown at that time; in other cases, it is just the choice of name that is different. His subjects wrote *burgonya* 'potato', *fagylalt* 'ice cream'; ours preferred *krumpli* and *fagyi*, respectively. Among his top-scoring items, *főzelék* 'vegetable dish' lost the most ground. Alcoholic drinks not mentioned by his subjects were *pezsgő* 'champagne' (16), *whisky* 'id.' (11), *vodka* 'id.' (4), *alkohol* 'alcohol' (6), whereas the frequency of *sőr* 'beer', *bor* 'wine', *rum* 'id.', *pálinka* 'brandy' diminished. The most frequently occurring fruit was *alma* 'apple' (137), followed by *körte* 'pear' (73) and *naranacs* 'orange' (71). The frequency of *ananász* 'pineapple' and *banán* 'banana' had increased significantly since Cser's time. With fruits, too, the gender difference was significant (in favour of girls).

Body parts. In our corpus, the order is *haj* 'hair', *szem* 'eye', *láb* 'leg', *fül* 'ear', *kéz* 'hand', *fej* 'head', *orr* 'nose', *száj* 'mouth', *köröm* 'nail', *fog* 'tooth', *ujj* 'finger', *szív* 'heart', *nyak* 'neck' (with respect to *nyelv* 'tongue/language', written by 10 boys and 19 girls, it is difficult to tell which reading they meant). The order is similar in Cser's material. Gender differences are not significant, although some frequent body parts were written more times by girls: *köröm* 'nail', *kar* 'arm', *száj* 'mouth', *fog* 'tooth', *nyelv* 'tongue', *has* 'belly', *kéz* 'hand'. Private parts were of low frequency, but showed a characteristic gender distinction; boys mentioned them a lot more often (52) than girls did (6). In 32 cases (only with boys) the words chosen were obscene.

Clothing. This category includes pieces of clothing, their accessories (*cipzár* 'zip fastener', *zseb* 'pocket'), as well as jewels. The word *óra* 'watch/clock/hour/lesson' (mentioned by 136 boys and 131 girls)—that could in principle belong here just as well as to "school" (etc.)—was classified as belonging here because Cser did the same. 95% of all subjects mentioned words belonging here, representing 7.25% of all common nouns. The most popular ones were *cipő* 'shoes', *kabát* 'coat', *nadrág* 'trousers'; among accessories, *óra* 'watch', *szemüveg* 'glasses', and among jewels, *nyaklánc* 'necklace', *gyűrű* 'ring', *fülbevaló* 'earrings' was the top order. ANOVA indicates a clear gender difference: girls wrote 10.46, whereas boys wrote 6.95 items of clothing on average. The words *hajgumi* 'rubber hairband', *szoknya* 'shirt', *zsebkendő* 'handkerchief', *gyűrű* 'ring', *bugyi* 'panties', *nyaklánc* 'necklace', *esernyő* 'umbrella', *nadrág* 'trousers', *cipőfűző* 'shoelace', *farmer* 'jeans', *szemüveg* 'glasses', *sapka* 'cap', etc. were more frequently mentioned by girls than by boys; interestingly, *melltartó* 'bra' occurred more often in boys' than in girls' lists.

Underwear was amply represented, although the frequency of the individual items was relatively low. Among items of underwear that were only mentioned by boys, or that were mentioned by more boys than girls, the various names for ‘men’s drawers’ were on top: *alsógatya* (4), *gatyó* (2), *alsónaci* (1), *alsónadrág* (1), *férfialsó* (1), respectively *gatya* (boys: 14, girls: 4), *bokszeralsó* (boys: 3, girls: 2). Among words for girls’ underwear, *bugyi* ‘panties’ was the most frequent (boys: 8, girls: 19); others occurred more or less sporadically (*tangabugyi* (boys: 10, girls: 2), *tanga* (boys: 3, girls: 2), *bugyogó* (boys: 1, girls: 1)). The neutral terms *alsónemű* ‘underwear’ (boys: 2, girls: 3), *fehérnemű* ‘lingerie’ (boys: 2, girls: 9), occurred more with girls than with boys. The most frequent word that was only written by girls was *top* ‘id.’ (7).

The largest increase compared to the frequency values found by Cser can be seen for *bugyi* ‘panties’, *pulóver* ‘sweater’, *bakancs* ‘boots’, showing the various sources of changes clearly. These days, various types of underwear are less taboo—in Cser’s list only *fehérnemű* ‘lingerie’ appeared with 19 occurrences. Usage and fashion have both changed. Words that show a large drop in frequency are also as expected. Sixty years ago *kötény* ‘apron’ occurred 171 times, *gallér* ‘collar’ 127 times; today, both were mentioned 4 times only. The “disappearance” of *kalap* ‘hat’ is also conspicuous: in Cser’s material it occurred 348 times, in ours it occurred 38 times (relative frequency: 34.8 → 9.5). In all, words for clothing had a much larger share in the old material than they had today.

House and household. This category contains words connected with residential buildings, their furnishings (following Cser, also objects that are typically found around a house like *kút* ‘well’, *kutyaól* ‘dog-house’), and pieces of furniture. It was one of the conceptual areas represented in large numbers. The most popular words were *szék* ‘chair’ (264), *asztal* ‘table’ (247), *ablak* ‘window’ (230), *ajtó* ‘door’ (205), *ház* ‘house’ (198), that is, the prototypical items. Gender differences were again significant. *WC* ‘toilet’ did not occur at all in Cser’s material; of its synonyms, only *illemhely* (3) and *toalett* (3) were to be found. With our subjects, *WC* occurred 55 times, and *budi* twice. Words that showed the largest increase of frequency were *csempe* ‘wall tile’ (relative frequency: 0.2 → 8.5), *fürdőszoba* ‘bathroom’ (relative frequency: 0.7 → 6.75). The largest decrease of frequency was shown by *roló* ‘roller-blind’, *mennyezet* ‘ceiling’, *kályha* ‘stove’, *sáml* ‘stool’, *fogas* ‘coat-rack’. In Cser’s list, *kályha* ‘stove’ occurred 458 times, today it occurred only 33 times (relative frequency: 45.8 → 8.25); on the other hand, *radiátor* ‘heater’ that obviously did not occur in his list at all, was mentioned 88 times by our subjects (relative frequency: 22).

The category “house” includes words for “household”, too: this is where words referring to objects found in one’s home, including kitchen and bathroom articles, household utensils, perfumery, as well as bed-linen were collected. The most frequent items were *villa* ‘fork’ (76), *kanál* ‘spoon’ (72), *pohár* ‘glass’ (70). Of words having to do with sanitation and cosmetics, *szappan* ‘soap’ (46), *körömlakk* ‘nail varnish’ (40), *fogkefe* ‘toothbrush’ were the most frequent. As expected, girls wrote a significantly larger number of words in this area.

Traffic. Cser’s category “technical progress” was divided into “traffic” and “engineering” (see below). The most frequent word within the area of “traffic” was *autó* ‘car’ (210), to which we can add *kocsi* ‘car’ (41) and *gépkocsi* ‘motorcar’ (2); the next one was *repülő* ‘airplane’ (76), *repülőgép* ‘aircraft’ (11). The largest drop in frequency concerned *autóbusz* ‘motor bus’ (our subjects wrote *busz* ‘bus’ instead), *mozdony* ‘locomotive’, *villamos* ‘tram’ (this perhaps because trams are less used in the country than in Budapest where all of Cser’s subjects were from). The relative frequency of *autó* ‘car’ grew from 33 to 52.5. Gender differences were significant, this time in favour of boys.

Engineering. This is where the words of telecommunication, computer technology, household appliances, machines and engines, and other technological terminology were collected. The most frequent items were *tévé* ‘television’ (163, plus *televízió* 46) and *számítógép* ‘computer’ (131, plus *komputer* 2 and *computer* 2). Boys wrote a larger number of “engineering” words than girls did, and seventh-formers more than sixth-formers did. In addition to *tévé* and *számítógép*, there were 33 other lexemes, a total of 103 occurrences. Boys wrote e.g., *traktor* ‘tractor’, *hangfal* ‘(hi-fi) speaker’, *gép* ‘machine’, *űrhajó* ‘spaceship’, *égő* ‘bulb’, *CD* ‘id.’ a lot more often, whereas with girls *mosógép* ‘washing-machine’, *számológép* ‘calculator’, *magnó* ‘cassette deck’, *videó* ‘video recorder’, *televízió* ‘television’ occurred similarly more frequently. It is not surprising that the overlap with Cser’s material is rather slight in this category; a mere 15% of all lexemes belonging to this area occurred there, too.

Nature. The most frequent words were *víz* ‘water’ (144), *hó* ‘snow’ (110), *föld* ‘earth’ (76), *folyó* ‘river’ (74), *hegy* ‘mountain’ (65). In ANOVA terms, gender was not significant here, but its correlation with age group was; especially 13-year-old girls wrote relatively many, and 13-year-old boys wrote relatively few, words in this category. The largest gender difference was shown by the names of seasons; for some reason, it was primarily 13-year-old girls who mentioned them.

Materials. Frequent names of materials were *üveg* ‘glass’ (120), *arany* ‘gold’ (75), *vas* ‘iron’ (63). Neither gender nor age group was significant. The differ-

ences with Cser's list were typical again: *beton* 'concrete' was ten times more frequent in our case, whereas *szén* 'coal', *ólom* 'lead', *selyem* 'silk', *márvány* 'marble', *vászon* 'linen' were included a lot less often.

Leisure, sports. The most popular toy (or sports requisite) was *labda* 'ball' (117); the next most frequent word occurring here was *játék* 'toy/game' (103); adding up synonyms, 'bicycle' occurred in the largest number (*bicikli* 101, *kerékpár* 18, *keró* 1, *bicó* 1). There was no significant gender difference, but significant "gender → age group" differences were found again (13-year-old girls wrote a lot more words belonging here than 13-year-old boys did). Cser's material lacked *foci* 'soccer' but included *futball* '(association) football' instead.

Tools. The most frequent tools were *kés* 'knife' (93), *olló* 'scissors' (76), *cső* 'pipe' (57), *csavar* 'screw' (49), *doboz* 'box' (44), *kalapács* 'hammer' (42), *tű* 'pin' (42). Gender and age group were not significant, although there were characteristically "boyish" words like *csavar* 'screw', *kötél* 'rope', *cső* 'pipe'. The lexemes *kalapács* 'hammer' and *szeg/szög* 'nail', however, were written by equal numbers of boys and girls.

School. This category included names of school subjects, and typical objects, concepts, and persons of school life. It is a very important conceptual area, all except three children wrote words belonging here, and 10% of all occurrences come under this heading. Statistically, it is an unanimously "girlish" category. The most frequent words were *toll* 'pen' (303), *ceruza* 'pencil' (295), *könyv* 'book' (283), *tábla* 'blackboard' (278), *tolltartó* 'pencilbox' (249). Compared to Cser's material, the largest drop in frequency was shown by *számítan* 'arithmetics', *töltőtoll* 'fountain-pen', *tinta* 'ink'; the first one of these was replaced by *matek* 'maths' or *matematika* 'mathematics'.

Curriculum. This is where lexemes belonging to the terminology of some school subject or field of science were classified. The most frequent ones were those occurring very early in one's studies (e.g., *szám* 'number', *betű* 'letter', *vers* 'poem'). Gender and age group were not significant. 73.7% of words belonging here occurred less than three times (e.g., *allegória* 'allegory', *deltoid* 'deltoid').

Man. This is where we classified terms of kinship, words for people in their various ages, as well as words like *barát* 'friend', *osztálytárs* 'classmate'. Gender differences were significant; this was a "girlish" category again; but the words for various kinds of 'human being' were very frequent irrespective of the gender of the subject. E.g., *fiú* 'boy', *kölyök* 'kid', *kisfiú* 'little boy', *csávó* 'guy', *haver* 'pal', *nő* 'woman', *asszony* 'woman', *hölgy* 'lady'. Six different words were used for 'mother', with a characteristic cline of frequency: *anya* (94), *mama* (41), *anyu* (27), *mami* (3), *anyuka* (2), *édesanya* (2). In comparison with Cser, the

most interesting thing was that *barát(nő)* '(girl)friend' gained frequency to the largest extent. The 6 occurrences of *barátnő* 'girlfriend' increased to 24 in sixty years' time (relative frequency: 0.6 → 6), and the 16 occurrences of *barát* 'friend' increased to 54 (relative frequency: 1.6 → 13.5). The occurrence of *család* 'family' was roughly three times as numerous as it had been sixty years before.

Negative. This is where words having to do with illness, war, death, crime, and drugs were counted. Also, words referring to e.g., homosexuality and other socially negative terms like *csöves* 'bum', *hajléktalan* 'homeless' were included here. Items of this category, a total of 653 occurrences (179 different words), were found with 59.75% of the subjects. Gender differences were not significant. Names of weapons, as expected, were written mainly by boys; only boys mentioned *puska* 'rifle' (28), *gépfegyver* 'heavy machine gun', *tőr* 'dagger' (3), *géppágyú* 'automatic cannon', *gránát* 'grenade', *harckocsi* 'tank' (2), and *géppisztoly* 'machine gun', *golyószóró* 'light machine gun', *mustárgáz* 'mustard gas', *páncélcocsi* 'armoured car', *plasztik* 'explosive', *rakétavető* 'rocket launcher', *stukker* 'shooter' (1). In the 'illness' subgroup, 12 common nouns reached the frequency value of 10; the subgroup contained 147 occurrences, 40 different words. Examples: *alkoholista* 'alcohol addict' (5), *influenza* 'id.' (5), *amnézia* 'amnesia' (3), *AIDS* 'id.' (2), *holdkóros* 'sleepwalker' (2), and *nyúlszáj* 'cleft palate', *reuma* 'rheumatism', *vakbélgyulladás* 'appendicitis' (1), *mentő* 'ambulance' (21), *beteg* 'patient/ill' (19), *kórház* 'hospital' (18), *seb* 'wound' (10). Illness and mortality were mainly on girls' minds. The word *halál* 'death' and *kórház* 'hospital' (12/6) appeared with twice as many girls (18) as boys (9), and the difference was even larger for *sír* 'grave/cry' (15/6), *beteg* 'patient/ill' (13/6).

Other objects. All objects that did not fit into any of the previous categories were collected here. This group consisted of 1544 occurrences, 181 different lexemes. Examples of items occurring 30 times or more: *kép* 'picture' (198), *levél* 'letter/leaf' (83), *kulcs* 'key' (72), *gyertya* 'candle' (66), *kosár* 'basket' (49), *zászló* 'flag', *fénykép* 'photograph', *koszorú* 'wreath'; examples of less frequent items in this category: *cigi* 'cigarette', *akvárium* 'fish-tank', *kupak* 'hood', *zacskó* '(paper) bag', *zsák* 'sack'.

Abstract nouns. In principle, nouns containing the derivational suffixes *-ás/-és* '-ing' or *-ság/-ség* '-ness', would have been clear candidates for this category. However, a number of them represented one of the above conceptual areas just as clearly. As a compromise solution, we classified the latter set (218 occurrences, 77 different lexemes) into the appropriate other categories; therefore, in this category we have all and only abstract nouns that did not fit into any

of the previous ones (altogether 1905 occurrences). Of these, a mere 11 lexemes reach the frequency value of 20 (7.8% occur only once or twice). Abstract nouns mentioned 20 times or more were: *szerelem* ‘love’ (80), *szín* ‘colour’ (54), *írás* ‘writing’ (49), *idő* ‘time’ (47), *szeretet* ‘affection’ (38), *barátság* ‘friendship’ (33), *élet* ‘life’ (24), *áram* ‘electricity’ (22), *beszéd* ‘speech’ (21), *munka* ‘work’ (20), *öröm* ‘joy’ (20). The lexeme *élet* ‘life’ occurred 7 times with boys and 17 times with girls, similarly *öröm* ‘joy’ was mentioned 7 times by boys and 13 times by girls. Also, girls listed *szerelem* ‘love’, *szeretet* ‘affection’ and *barátság* ‘friendship’ several times as often as boys did. *Boldogság* ‘happiness’ was written by a single boy and 14 girls. A large number of different lexemes occurred here (649); this was the most populous group in terms of number of different lexemes. Gender differences were significant, in favour of girls. The correlation of genders and age groups was also significant: 13-year-old girls wrote the largest number of abstract nouns.

The question arises which conceptual areas are the most important for the children, i.e., which areas are represented the most heavily. This can be judged on the basis of the total number of occurrences within the given conceptual area, as well as according to what percentage of the subjects mentioned words belonging to the given area. The order in these terms is: **house and household, school, animals, clothing, food, leisure**. The order based on the percentage of the subjects mentioned words of a given area is **school, house and household, leisure, clothing, animals, plants, food**. As we have already mentioned, Cser says that there are “girlish” and “boyish” conceptual areas; that is, gender-bound differences in what is interesting for the subjects are reflected in free word associations. Our analyses have shown the following. Many more girls wrote words belonging to the areas of **perfumery** within the category of **house and household, colours, nationalities**, and **fruit** within the category of **food** (the difference was over 10% in each case) and **house and household** (where the difference was 9.8%). Girls wrote at least one word more (on average) than boys did in the categories food, clothing, school, man, and abstract noun. On the other hand, for **traffic**, the difference was 4.6% in favour of boys. Therefore, gender differences are indeed characteristic. On the basis of earlier word association experiments it was assumed that words are activated from both the active and passive parts of one’s vocabulary. The data of the present investigation confirm the claim that words accessed in the active vocabulary constitute a larger part of the total number of words activated, given that this is what explains the unambiguous distinction between “girlish” and “boyish” conceptual areas. According to our statistical analyses, gender was a determining

factor, whereas age group was significant only for **animals**. It appears that this conceptual area loses some of its importance as the child grows older.

4. Free word associations at the age of six

From a certain age onwards and to some extent, even nursery children are capable of word associations (given a CV syllable as trigger, they are able to access an existing word or several existing words already at the age of 4–5, cf. Gósy 1999a). However, we have been unable to find data in the literature concerning an investigation in which children were to say words one after the other, with no limitation whatsoever. Word associations—as we saw above—provide us with valuable information concerning characteristics of the processes of access to the mental lexicon and the strategies employed, hence they are a good tool of analysing children’s development. Therefore, we conducted another experiment involving 6-year-old nursery children in order to learn if they are able to activate their mental lexicon by the strategy of free word associations, as well as what the results of such activation may be. We did not even have a vague notion of how many and what kinds of lexical items a small child of this age would be able to activate within a given period of time. Although a comparison with the material of schoolchildren (see above) is not unambiguously possible because of the differences in the number of subjects and methods of experiment, the acquired data make us able to draw some conclusions and recognise some tendencies.

In this experiment, 19 six-year-old (between 6;1 and 6;8) nursery school children participated (8 girls and 11 boys) in the framework of the afternoon leisure activities in the nursery (this being exactly as familiar surroundings as the classroom is for schoolchildren). The experiment was conducted by the nursery teacher of the children, on the basis of instructions given to her. Testing was done individually. The children got the following task: “Say things like *ball*, *sleeps*, *joy*, *sweet*, as many as you can”, that is, the model list contained words belonging to diverse parts of speech. Subsequently five minutes were at each child’s disposal to activate words that were immediately put down in writing by the nursery teacher. One third of the fifteen minutes allotted to schoolchildren was allowed because of the age difference and the oral nature of the task. Most children were able to go on activating words for some four minutes.

The children were able to solve the task; they uttered a total of 709 mental words. The average values obtained are given in Table 5. The six-year-olds

tested were able to activate an average of 7.46 words per minute in this setup of free word association.

Table 5
Numerical results of free word associations
by nursery children

GENDER	MEAN	STANDARD DEVIATION
Boys	37.81	14.35
Girls	36.62	13.41
Total	37.31	13.59

Girls and boys practically did not differ in the number of words activated. The fewest words activated by a child were 17, the largest score was 66. Girls activated a total of 191 different lexemes, 23.87 on average, whereas boys listed a total of 255 lexemes, 23.18 on average. The total number of mental words given above (709) does not include repetitions (by the same child). The incidence of repetitions was very large, 12 out of the 19 children did produce repetitions (this is 63.2%, the same figure for schoolchildren was 67.25%). The repetitions constituted 8.38% of all activated items (on average). Compared to the data of the schoolchildren, this is many times as large as the value we got there (1.85%). Table 6 shows the numerical comparison of repetitions.

Table 6
A comparison of repetitions

SCHOOLCHILDREN			NURSERY CHILDREN		
GENDER	MEAN %	STANDARD DEVIATION	GENDER	MEAN %	STANDARD DEVIATION
Boys	2.01	1.25	Boys	5.87	3.38
Girls	1.69	1.13	Girls	13.40	5.35
Total	1.85	1.20	Total	8.38	5.37

We have analysed the part-of-speech distribution of the data. Of the 709 mental words obtained, 704 were nouns, 4 were adjectives (*elektromos* ‘electric’, *világos* ‘clear’, *plüss* ‘plush’, *barna* ‘brown’), and 1 was an adverb (*korán* ‘early’). There were a few lexemes that are in principle ambiguous in terms of part of speech (*szív* ‘heart/inhale’, *hal* ‘fish/die’, *fal* ‘wall/devour’, *nyúl* ‘rabbit/reach for’, *ég* ‘sky/burn’, as well as *repülő* ‘aeroplane/flying’, *dongó* ‘blue-bottle fly/buzzing’, *kuka* ‘dustbin/tongue-tied’, *üdítő* ‘soft drink/refreshing’). These were counted as nouns. We had anticipated (on the basis of the data of the schoolchildren) that an overwhelming majority of the words would be nouns, especially that this is the dominant part of speech at the early stages

of language acquisition. That assumption was borne out by the data. Of the 704 nouns, 671 were common nouns and 33 were proper nouns, of which 15 were first names. It is conspicuous that not a single verb occurred (the few noun/verb items listed above are almost certain to have been meant in their nominal function).

The six-year-olds mentioned very few suffixed words, a total of 11, that is, 1.5%. All of these suffixed words were plural nouns, e.g., *állatkák* 'little animals', *videokazetták* 'videotapes', *majmok* 'monkeys', *békák* 'frogs', *ceruzák* 'pencils', *könyvek* 'books', *halak* 'fish (pl.)', *virágok* 'flowers'. (Note that almost half of these plural words were activated by the same little boy.) The three most frequently occurring items were *virág* 'flower', *szék* 'chair', and *asztal* 'table', whereas with schoolchildren they were *toll* 'pen', *ceruza* 'pencil', and *könyv* 'book'. That is, in both groups, words for the objects of the children's immediate surroundings were top scoring. The average number of syllables was 2.32, roughly the same as the corresponding figure for schoolchildren.

The word associations of six-year-olds contained sequences of words in which activations based on semantic, morphological, and/or phonological similarities could be found. However, due to the relatively small number of items, the frequency of accessing strategies is hard to determine reliably. It was the semantic strategy that occurred the most often, e.g., *virágtaró* 'flower stand', *virág* 'flower'; *baba* 'doll', *Barbi baba* 'Barbie doll', *napocska* 'sun', *felhő* 'cloud', *elefánt* 'elephant', *zsiráf* 'giraffe' or *szék* 'chair', *asztal* 'table', *ágy* 'bed', *szekrény* 'cupboard'. Sequences based on phonological similarity include *szekrény* 'wardrobe', *szem* 'eye', *virág* 'flower', *villanykörte* 'lightbulb'; *kabát* 'coat', *kacsa* 'duck', *harang* 'bell', *halak* 'fish (pl.)', *gomba* 'mushroom', *gomb* 'button', *szék* 'chair', *szél* 'wind', *társasjáték* 'indoor game', *táska* 'bag'; *kamion* 'truck', *kacsa* 'duck'. Examples of associations that can be claimed to be morphological: *csibe* 'chicken', *csibe lába* 'chicken's leg', *zsebóra* 'pocket watch', *zseblámpa* 'pocket torch', *szekrény* 'cupboard', *szekrény lábai* 'legs of cupboard'. The following can be seen as mixtures of semantic and morphological effects: *villany* '(electric) light', *villanyoszlop* 'pole', *virág* 'flower', *virágszál* '(a single) flower'; *tyúk* 'hen', *tyúkól* 'hen-house'; *baba* 'doll', *babakonyha* 'doll kitchen', *babakocsi* 'pram', *fésű* 'comb', *fésűtartó* 'comb case', *gáztűzhely* 'gas-range', *gáz* 'gas'. An interaction of semantic and phonological factors may have produced *maci* '(teddy) bear', *majom* 'monkey'. The following sequence probably shows all three effects: *víziló* 'hippo', *víz* 'water', *vízisikló* 'grass snake'. All this suggests that the claim that children are prone to employing predicative constructions in their word associations was not borne out in our case (Sigurd 1996).

Comparing the lexemes listed by nursery children and schoolchildren, we can observe certain age-bound characteristics. The 354 lexemes of six-year-olds contained 44 (12.4%) that did not occur in the schoolchildren's material; these words are probably members of the passive vocabulary of the older, and of the active vocabulary of the younger, children. These words (i) have to do with "diminution": *halacska* 'fishlet', *emberke* 'little fellow', *kisegér* 'little mouse', *kisautó* 'toy car'; (ii) refer to toys for small children: *babakonyha* 'doll kitchen', *autóskártya* 'car card', *építőköcka* 'building bricks'; (iii) may go back to the differences in the immediate surroundings: *irattartó papucs* 'filing jacket', *fésűtartó* 'comb case'; (iv) occur in a different shape, e.g., *teknőc* 'turtle', rather than *teknős* 'tortoise'; or (v) are relevant for some personal reason like *egészségház* 'health care center', *orrszívó* 'nose pump', *dekóder* 'decoder'.

The results of our analysis of conceptual areas are also revealing. Of the areas we established for the schoolchildren, the ones that were (practically) missing were the **colours** (except *barna* 'brown'), **negative** items (except *fegyver* 'weapon'), obviously **curriculum** (except *betű* 'letter', *szám* 'number'), **occupations** (except *doktor* 'doctor', *orvos* 'physician', *bohóc* 'clown'), **abstract nouns** (except *memória* 'memory', *áram* '(electric) current'), **tools** (except *hőmérő* 'thermometer', *kalapács* 'hammer', *olló* 'scissors', *doboz* 'box', *zsinór* 'string'). Similarly, just a few items (1–3%) occurred in the categories **man** (e.g., *apa* 'father', *anya* 'mother', *gyerek* 'child', and *pólyás baba* 'baby'), and **materials** (*gyurma* 'plasticin', *plüss* 'plush'-these two may refer to the relevant toys-and *arany* 'gold', *kő* 'stone', *üveg* 'glass'). The category **body parts** included 18 occurrences of 12 lexemes, e.g., *haj* 'hair', *orr* 'nose', *szem* 'eye', *szív* 'heart'. The conceptual area **engineering** was represented by 19 occurrences of 12 different words, e.g., *rádió* 'radio', *számítógép* 'computer', *videó* 'video', *televízió* 'television'. The next frequency range was constituted by areas whose representation was 3–5%. These were **school**: 12 different words, 26 occurrences (the most frequent items were *könyv* 'book', *lap* 'page', *ceruza* 'pencil'), **food**: 19 different words, 27 occurrences (primarily various kinds of fruit like *alma* 'apple', *dinnye* 'melon', *cseresznye* 'cherry'), and **plants**: 10 different words, 31 occurrences. The item *virág* 'flower' occurred the most often here, both for girls and for boys. Other words in this area were *fű* 'grass', *fa* 'tree', and *tulipán* 'tulip'. The category **traffic** was represented by 15 different words, 33 occurrences. The most frequent ones were *autó* 'car', *repülő* 'plane', *vonat* 'train', *hajó* 'ship'. **Nature** (including weather) was represented by 12 different words, occurring 30 times, respectively. The most frequent items were *akvárium* 'fishbowl', and *felhő* 'cloud', *föld* 'earth', *nap* 'sun'.

At the age of six, in the group tested, the most populous conceptual areas were **clothing**: 28 different words, 64 occurrences (the most frequent ones were *cipő* 'shoes' (8), *óra* 'watch' (6), *zsebkendő* 'handkerchief' (6)); **leisure** and sports: 35 different words, 71 occurrences (e.g., *lego* 'lego' (7), *mászóka* 'climbing toy', *kocka* 'cube' (6), *baba* 'doll' (5)); **house** and household: 60 different words, 126 occurrences, 18.6% of all mental words activated (e.g., *szék* 'chair' (11), *asztal* 'table' (9), *ágy* 'bed' (6), *ajtó* 'door' (6), *ház* 'house' (5), *szekrény* 'cupboard' (5)); and finally—as expected—**animals**: 71 different words, 138 occurrences, 20.4% of all mental words in this test (e.g., *elefánt* 'elephant', *kutya* 'dog', *majom* 'monkey' (6), *kígyó* 'snake', *macska* 'cat' (5), *oroszlán* 'lion', *víziló* 'hippo', *pillangó* 'butterfly', *zsiráf* 'giraffe', *cica* 'kitten' (4).

5. Conclusions

Our results have proved that the method of free word associations can be put to good use in a complex investigation of the mental lexicon. The children activated both their active and passive vocabularies. Although in an implicit way, the hypothesis "the faster lexical access is, the more words can be activated" was also borne out. At least, this appears to be the most straightforward explanation for the fact that there was a significant difference between the two age groups of schoolchildren concerning the number of lexemes activated. The older group also wrote longer words on average. On the other hand, the average length of words did not change between the six-year-old nursery children and the schoolchildren.

Seventh-formers wrote a lot more lexemes that were not mentioned by the other group. In terms of the whole corpus, the overlap between the two age groups, i.e., the proportion of lexical items occurring in both, was 37.62%. The individual differences were considerable, both in the younger and the older group. Girls were able to activate a larger number of words than boys (in both age groups). It was possible to discern "girlish" and "boyish" conceptual areas among associations. 12.4% of the words activated by the nursery children (these were the lexemes not mentioned by the two school groups) can be unambiguously assumed to belong to the passive vocabulary of the schoolchildren; the semantic analysis above gave a clear answer to why this was the case.

The facts that were revealed by our data concerning the characteristics of associations and the manner of lexical access all support the claim that the mental lexicon has a network-like structure. The detailed typology above was the first analysis (that we are aware of) of the structure of the mental lexicon

and the strategies of access (phonological, structural, and semantic) for Hungarian speakers. It appears that with younger speakers semantic relationships are the dominant factor, but the use of structural and phonological criteria during the activation of the mental lexicon can be observed already at the age of six. Hence, the strategies of orientation within one's own vocabulary take shape relatively early on.

At present, there are no appropriate methods for establishing the size of a speaker's individual mental lexicon; therefore all hypotheses or estimates in this respect are equally important. The 400 children tested (200 girls and 200 boys) activated a total of over 50,000 mental words. If we assume that all children are familiar with all the lexemes in this corpus (7397 items including proper nouns and English words) and we furthermore assume that this is all they are familiar with, this would mean that between the ages of three and thirteen they have to acquire approx. 1.6 words a day (this is based on the approximately 1500 lexical items attested for 3-year-old children, cf. Gósy 1984 and Meixner 1976). It is likely that, in ten years' time, children are able to extend their vocabularies to a larger extent than this, even if such extension is not necessarily gradual over the whole ten years.

Our comparative analysis of the two word association experiments sixty years apart have revealed certain changes in linguistic knowledge and language use. Today's children wrote over 14,000 words more than their "grandparents" did; the difference being 2.4 words per minute. This makes it probable that the vocabulary of children is larger today (an important contributive factor may be television); but it is also possible that their lexical access is (also) superior, due to the requirements of faster speech rate (among other things). 68.8% of Cser's corpus occurred in our own material, too. Since the latter is a much larger set of words, this common stock is only 35.6% of the words activated by our subjects. 55.8% of the common core falls within the same frequency range (cf. Figure 4).

The differences in the extent to which individual conceptual areas are represented in the two materials are characteristic and reflect general changes in way of life, society, technical devices, environment, and spheres of interest. The two corpora illustrate changes in usage and differences in what counts as "taboo". From a pedagogical point of view it is interesting—and may be due to certain changes in the principles of upbringing (family? school?) or in social norms in general—that today's children hardly constrain their performance in word access tasks. They do not hesitate to put down in writing a number of words that children would not even dare to think of sixty years ago (including even *bugyi* 'panties'). School expectations may underlie the lack of diminu-

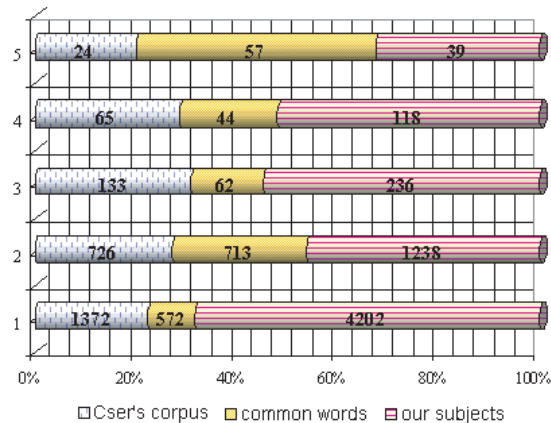


Fig. 4

Common and different words in the data collected 60 years ago and today
 (the units of the vertical axis are: 1 = $n < 3$, 2 = $3 \leq n \leq 19$,
 3 = $20 \leq n \leq 39$, 4 = $40 \leq n \leq 79$, 5 = $n \geq 80$)

tive forms sixty years ago (for instance, in words for school subjects). There are predictable differences that go back to the “obsolescence” of certain words due to a variety of reasons, words like *szatócs* ‘chandler’, *pengő* ‘old Hungarian currency’, *mángorol* ‘mangle’ (verb), *penna* ‘quill’, *plajbász* ‘(lead) pencil’, respectively *leadó* ‘radio set/transmitter’, *távirda* ‘telegraph office’, *sparherd* ‘kitchen range’, *stelázsi* ‘tier-stand’, *vegytan* ‘chemistry’, *ródlí* ‘sledge’, *prakker* ‘carpet-beater’, *trén* ‘Army Service Corps’, *plint* ‘plinth’.

The incidence of certain lexemes having to do with trouble or unpleasant phenomena has grown, e.g., *halál* ‘death’ (relative frequency: 1.6 → 6.75), *baleset* ‘accident’ (relative frequency: 0.3 → 3). The words *drog* ‘drugs’, *gyilkosság* ‘murder’ did not at all occur sixty years ago, although *gyilkos* ‘murderer’ did. It is interesting that *bomba* ‘bomb’ hardly ever occurs in Cser’s material; in our own list even *atombomba* ‘atomic bomb’ crops up. (Sixty years ago, children tended to write *ágyú* ‘cannon’ and *kard* ‘sword’ instead.) Positive emotions are also more heavily represented in today’s corpus, primarily *szerelem* ‘love’ (relative frequency: 0.5 → 20) and *barátság* ‘friendship’ (relative frequency: 0.2 → 8.25) but similar differences for *öröm* ‘joy’ (relative frequency: 1.6 → 5), *boldogság* ‘happiness’ (relative frequency: 0.2 → 3.75), and *nevetés* ‘laughing’ (relative frequency: 0 → 3.5) are also worth mentioning.

Both pedagogically and psychologically speaking, it is very important to draw the appropriate conclusions from individual differences among children.

The distance between children activating 24 (the minimum amount) and 222 mental words (the maximum amount) raises the issue of individual size of the mental lexicon and/or individual amount of successful access. (At the age of six the difference is less, the fewest is 17 and the largest is 66 words). All that interacts with communication processes and learning processes, hence it is considerably affecting the cognitive abilities, learning successes, and further development of the children. The individual differences in the process of first language acquisition, often extremely large ones, are familiar and are accepted as facts (especially in early stages of learning to speak). However, the word activation differences observed with six-year-olds forecast the possibility of those very large dissimilarities that we have attested with sixth- and seventh-formers. One year prior to starting school, there are children who are able to activate twice, three times, even four times as many words as some of their age-group peers. This is important because a successful and fast strategy of activating one's mental lexicon is indispensable for the appropriate working of the processes of both speech production and speech comprehension.

References

- Aitchison, Jean 1987. *Words in the mind. An introduction to the mental lexicon*. Blackwell, Cambridge MA & Oxford.
- Bakonyi, Hugó 1918. A gyermeknyelvi szókincs fejlődése [The development of children's vocabulary]. In: *A gyermek* 7: 337–8.
- Benedict, Helen 1979. Early lexical development: comprehension and production. In: *Journal of Child Language* 6: 183–200.
- Berko Gleason, Jean 1985. *The development of language*. Merrill, Columbus.
- Berko Gleason, Jean – Nan Bernstein Ratner (eds) 1998. *Psycholinguistics*. Harcourt Brace College Publishers, Orlando.
- Carroll, John B. 1961. Language development in children. In: Sol Saporta (ed.) *Psycholinguistics*, 331–46. Holt, Rinehart & Winston, New York.
- Clark, Eve 1995. Later lexical development and word formation. In: Paul Fletcher – Brian McWhinney (eds) *The handbook of child language*, 127–68. Blackwell, Cambridge MA & Oxford.
- Coleman, John 1998. Cognitive reality and the phonological lexicon: a review. In: *Journal of Neurolinguistics* 1: 1–26.
- Cser, János 1939. *A magyar gyermek szókincse* [The vocabulary of the Hungarian child]. Magyar Pedagógiai Társaság, Budapest.
- Füredi, Mihály – József Kelemen 1989. *A mai magyar nyelv szépprózai gyakorisági szótára: 1965–1977* [Dictionary of Hungarian literature language: 1965–1977]. Akadémiai Kiadó, Budapest.

- Galton, Francis 1883. *Inquiries into human faculty and its developments*. Dent, London.
- Gathercole, Susan E. – Alan D. Baddeley 1993. *Working memory and language*. Hillsdale, Lawrence Erlbaum.
- Gósy, Mária 1984. *Hangtani és szótani vizsgálatok hároméves gyermekek nyelvében* [Phonetic and lexical investigations in the language of 3-year-old children]. Akadémiai Kiadó, Budapest.
- Gósy, Mária 1998a. *Színmegnevezések gyermekkorban és felnőttkorban* [The use of colour terminology by children and adults]. In: Zsolt Lengyel – Judit Navracsics (eds) *Alkalmazott nyelvészeti tanulmányok II* [Studies in applied linguistics II], 55–69. Veszprémi Egyetem, Veszprém.
- Gósy, Mária 1998b. *Tőszók és toldalékolt szavak aktivizálása a mentális lexikonban* [The activation of roots and affixed words in the mental lexicon]. In: Mihály Hajdú – Borbála Keszler (eds) *Emlékkönyv Abaffy Erzsébet 70. születésnapjára* [A Festschrift for Erzsébet Abaffy on her 70th birthday], 59–64. ELTE Magyar Nyelvtörténeti és Nyelvjárási Tanszéke, Budapest.
- Gósy, Mária 1999a. *Evaluation of speech perception and comprehension processes of children: description and practical results*. In: Ben Maassen – Peter Groenen (eds) *Pathologies of speech and language*, 50–60. Whurr Publishers, London.
- Gósy, Mária 1999b. *Pszicholingvisztika* [Psycholinguistics]. Corvina, Budapest.
- Ingram, David 2001. *Toward a theory of phonological development*. In: William Kreidler (ed.) *Phonology*, 60–79. Routledge, London.
- Jakobson, Roman 1969. *Hang – jel – vers* [Sound – sign – poem]. Gondolat, Budapest.
- Janota, Přemysl 1970. *Development of children's vocabulary*. In: Karel Ohnesorg (ed.) *Colloquium Paedolinguisticum. Proceedings of the First International Symposium of Paedolinguistics*, 101–9. Mouton, The Hague & Paris.
- Jarovinskij, Alekszandr 1995. *Korai szókincs a gyermeknyelvben* [Children's early vocabulary]. In: *Általános Nyelvészeti Tanulmányok* 18: 91–101.
- Kenyeres, Elemér 1926. *A gyermek első szavai és a szófajok fellépése* [The first words of a child and the emergence of parts of speech]. Athenaeum, Budapest.
- Lakoff, Robin 1975. *Language and woman's place*. Harper & Row, New York.
- Levelt, Willem J.M. 1989. *Speaking. From intention to articulation*. MIT Press, Cambridge MA.
- Lindblom, Björn 1999. *How speech works: Questions and some preliminary answers*. Plenary talk at Eurospeech '99. Budapest.
- Lőrík, József – Péter Ajtony – Gábor Palotás – Csaba Pléh 1995. *Szókincsteszt gyermekeknek* [Vocabulary test for children]. Gyógypedagógiai Főiskola, Budapest.
- McNeill, David 1970. *The development of language*. In: Peter Mussen (ed.) *Carmichael's manual of child psychology I*, 1061–1162. Wiley, New York.
- Meggyes, Klára S. 1971. *Egy kétéves gyermek nyelvi rendszere* [The linguistic system of a two-year-old child]. *Nyelvtudományi Értekezések* 71. Akadémiai Kiadó, Budapest.
- Meixner, Ildikó 1976. *Hároméves gyermekek szókincse* [The vocabulary of three-year-old children]. In: József Molnár – Imre Wacha (eds) *A beszéd-szimpozium magyar előadásai*. Szeged, 1971 [Hungarian talks at the 1971 Szeged symposium on speech], 55–9. Nyelvtudományi Társaság, Budapest.

- Papp, Ferenc 1973. Tőszókincsünk etimológiai rétegenkénti hangstatisztikája [A sound statistics of Hungarian roots by etymological strata]. In: Nyelvtudományi Közlemények 75: 3–40.
- Pléh, Csaba 2000. A magyar morfológia pszicholingvisztikai aspektusai [Psycholinguistic aspects of Hungarian morphology]. In: Ferenc Kiefer (ed.) Strukturális magyar nyelvtan 3: Morfológia [A structural grammar of Hungarian 3: Morphology], 951–1020. Akadémiai Kiadó, Budapest.
- Sigurd, Bengt 1996. Computer simulation of word associations and crossword solving. In: Working Papers, Lund University 45: 157–67.
- Steinberg, Danny 1993. An introduction to psycholinguistics. Longman, London & New York.
- Vértés O., András 1955. A gyermeknyelv [Child language]. Gyógypedagógiai Főiskola, Budapest.
- Waterson, Natalie 1987. Prosodic phonology. The theory and its application to language acquisition and speech processing. Grevatt & Grevatt, Newcastle upon Tyne.

Address of the authors: Mária Gósy
Research Institute for Linguistics
Hungarian Academy of Sciences
Benczúr utca 33.
H-1068 Budapest
gosy@nytud.hu

Magdolna Kovács
Department of General and Applied Linguistics
University of Debrecen
Egyetem tér 1.
H-4032 Debrecen
mkovacs@fox.klte.hu