



# Using Latent Semantic Analysis for Simulating of Children's Cognitive Development

A.Solovyev

St.-Petersburg State University



# Target

In the 20th century Noam Chomsky formulated the so-called Plato's problem: why is the amount of our knowledge much more than we can extract from our everyday experience? For example, the pre-school children (6-7 years old) vocabulary increases in 3-8 words every day on the average. And not every word has a correlation with any reality or action (for example, abstract concepts, words carrying "phatic" or uninformative assignment, etc.).

The question arises: how does the child recognize each new meaning of the word and its relation to other, or why do the new "meanings" form?



## Why LSA?

This paper presents the results of three experiments of 4-7-year-old children cognitive development using a Latent Semantic Analysis (LSA) that permits to compare the semantic similarity between pieces of textual information.

A method to use in our research - LSA - makes possible to simulate the associative-semantic relations between words. On the one hand it eliminates the rigid binding of lexical unit to any of the clusters, and on the other it presents a complete system of relationships between words.



# Experiments

In our experiments we used a technique developed by G.Denhière and B.Lemaire. The principal difference of our research is: 1) experiments were performed on the Russian language for the first time; 2) experiments were performed on pre-school 4-7-year-old children for the first time.

Children were grouped into two categories 4-5 and 6-7 years old, which corresponded to the age variability of cognitive development.



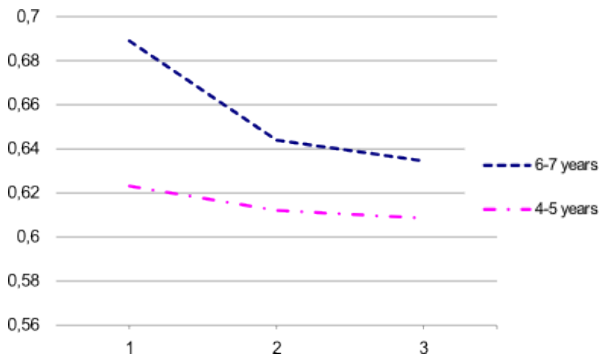
Two experiments describe semantic and associative similarity between LSA-models and children's cognitive development. The third experiment describes LSA using to measure children's semantic memory.

These three experiments results are compared to children's model data and adult's model data.

The computational models are built from the LSA of a multisource child corpus and of an internet mass media corpus.



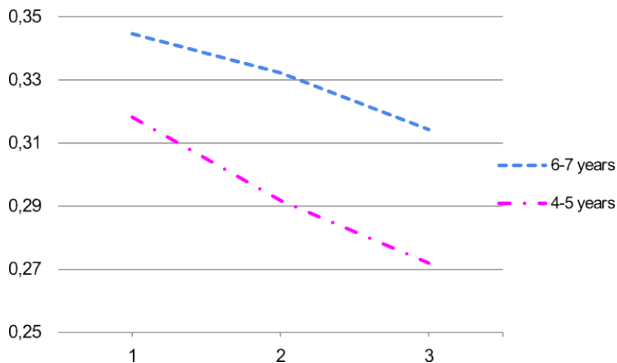
## Experiment 1



**Figure:** 1. Average values of the scalar products for the first three words associations with LSA model for the first group of children (4-5 years) and the second group of children (6-7 years). Abscissa - numbers of the first three words, the vertical axis - the value of the scalar product.



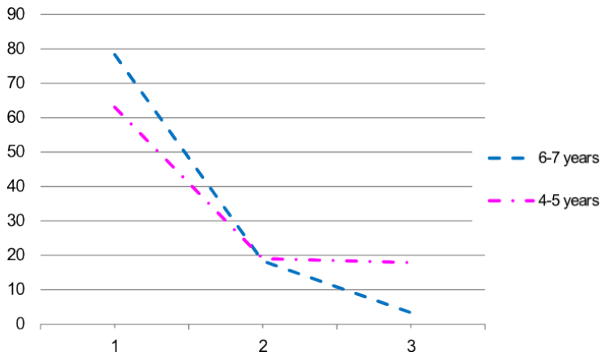
## Experiment 2



**Figure: 2.** Average values of the scalar products for the first three words associations of the LSA model received from the media corpus for the first group of children (4-5 years) and the second group of children (6-7 years). Abscissa - numbers of the first three words, the vertical axis - the value of the scalar product.



## Experiment 3



**Figure: 3.** Trendline responses for first and second children groups. Abscissa - response options, the vertical axis - the sum of the percentage of responses.





## Conclusion

Our findings confirm that:

- 1) LSA can be used to simulate a variety of children's cognitive processes;

## Conclusion

Our findings confirm that:

- 1) LSA can be used to simulate a variety of children's cognitive processes;
- 2) LSA models represent the development of different age children groups cognitive processes, in particular, associative semantic processes and short-term and long-term memory work;

## Conclusion

Our findings confirm that:

- 1) LSA can be used to simulate a variety of children's cognitive processes;
- 2) LSA models represent the development of different age children groups cognitive processes, in particular, associative semantic processes and short-term and long-term memory work;
- 3) This method may be recommended for the comparative study of children's cognitive development, in particular, the development of associative-logical thinking, verbal discourse, the development of memory.