ON ASSOCIATION EXPERIMENTS

Ekaterina Vylomova
Bauman MSTU, Moscow
e-mail: evylomova@gmail.com

ABSTRACT

The article presents general information about free association experiments, their results analysis and modeling of these results.

1 . INTRODUCTION

Associative experiment is one of the first projective methods which is based on assumption that uncontrolled associations are symbolic or direct projection of internal content of consciousness.
This property makes possible to use association experiment to discover and describe affect complexes.
There are 4 types of association experiments:

- pair associations;
- serial associations;
- verbal differentiation;
- free associations

Each experiment could have single/multiple and free/controlled associations. Depending on their combinantions one could differ such kinds of associations [1],[2]:

- single-response free association (the respondent is asked to write the fisrt word that comes to his/her mind);
- single-response controlled association (the difference from previous one is that the respondent is asked to write specific type of association, e.g. synonym or antonym);
- multiple-response free association (the respondent could use any number of words);
- multiple-response conrolled association (almost the same as previous one, but with some limitations on the type of response).

First association experiments were hold in the USA [3], [4], [5] and Belgium[6].

According to the results of experiments, most answers are typical and cultural primaries. Also there was noticed that respondents give more typical answers under time pressure [1].
Association tests show individual and socio-cultural differences in respondents, how mass media, environment, culture, ideology influence on forming subject's association system.

2 . ANALOGUES

2.1 Word Association, rhyme and fragment norms [7]

Type: single-response free association
The experiment was hold in the USA in 1973-1975, number of participants was ~6,000. Overall number of cues — 5,019, number of reactions - ~750,000.
Every participant was presented with 100-120 english words in a booklet containing 25-30 words per page. The pages and words were unsystemically randomized.
Respondents were asked to write only one word as reaction (single association).
Stimuli contained 76% nouns, 13% adjectives, 7% verbs and 4% other parts of speech.
In 1986 experiment data was translated from print to computer representation.

In current moment the data is opened to everyone in the Internet.

2.2 Russian-american experiment [2]

Type: multiple-response controlled associations

Experiment was hold in April of 1999, in the time of crisis in Kosovo, Yugoslavia when NATO defence forces were bombing Yugoslavian objects.

In experiment participated 314 russian and 182 american students. Individual and geografic differences between universities at the same country were not considered.

Oraganizators asked to write down 5-10 first nous or 2 words collocations (noun+noun or adjective+noun) that come to subject's mind to descibe current situation in Yugoslavia.

Overall number of responses from russian students — 2,746 (avg. 8.75), from american — 1,140 (avg. 6.26).

As the result there was discovered that russian students had more similar answers. Most frequent associations from russians: death(39%), war(36%), bomb(31%), fear, refugees, NATO, horror, blood, murder, hunger; from americans: war(26%), ethnic(22%), death(21%), refugees, NATO, bomb, sad, genocide, religion, suffering.

2.3 Word associations: Norms for 1,424 Dutch words in a continuous task [8]

Type: multiple-response free associations

The experiment was hold in 2003-2006 in the university of Leuven, Belgium.

Number of participants was 10,292. From this group 6,329 were male, 3,582 were female and 381 didn't indicate their sexes.

Average age of paritcipants was 24 (SD=10.55). Stimuli set contained 1,424 concepts: 1,266 nouns, 77 verbs, 80 adjectives and 1 numeral.

Every word layed in specific category (e.g. «fruits», «vegetables», «animals», «musical instruments», «vehicles»).

Respondents were asked to write down first 3 associations. If respondent didn't have any associations it was allowed to skip stimulus.

Overall number of concepts (stimuli and reactions) was 30,311, unique pairs «cue-reaction» - 133, 401, ovarall number of pairs — 381,909.

3 . RUSSIAN ASSOCIATION-COMPARATIVE DICTIONARY [9]

Tun: single-response free association

Russian associaion dictionary includes data from 4 free association experiments that were hold from 1967 till 2000 with 10-20 years between them. The participants were russian speakers.

3.1 Russian language association norms dictionary ed. by Leontiev A. A. (1967-1973)

Dictionary includes 196 stimulus. The participants were russian speakers of age 16-50 y.o. and high/uncompleted high education.

Number of respondents for every stimulus was 500-800. Every participant received 100 stimuli list and had 5-7 seconds to respond on each stimulus.

Dictionary final size was 12,178 pairs.

Also every participant could support the answer with informtion about year and place of birth, native language, education and speciality.

3.2 Russian association experiment (1986-1997) [10]

Association experiment had 3 stages.

Stimuli for the first stage were formed by experiment authors and included 1,277 concepts. For 2nd and 3rd stage stimuli were selected from most frequent reactions of previous stage and included 2,690 and 2,930 words.

The overall number of stimuli — 6,624.

According to the rules of experiment participants were asked to write the frist word-reaction that came to their mind on represented stimulus.

Stimuli were randomly placed in the lists.
There were 5,000 forms with 100 stimuli in each one. There were no same ones among them and all stimuli are uniformly distributed.
Every participant had 7-10 minutes to fill the form.
Number of participants was 11,000, most of them were 1-3 year students from different university (34 specialities).
On bases of this experiment was formed verbal association network (VAN).
In addition, as in previous experiment, respondents were asked to put information about their sexes, speciality and place (town, region) they live in.

3.3 Slavic association dictionary(1998-1999)[9]

The dictionary contains 112 stimuli, each of them is represented on russian, byelorussian, ukranian and bulgarian languages.
Respondents were 500 students at the age of 18..25. There were represented 11 different specialities (mathematics, physics, biology, philosophy, science of law, medicine and others).
Every participant was asked to fill the form containing 112 stimuli. The time for each form was 10 minutes.
As in previous experimens every participant was asked to put information about sex, age, native language, speciality and date.

3.4 Associations in information technologies: russian-french experiment(1999-2000)[11]

Stimuli were taken from «Computer week» magazine, 1995 corpus.
The dictionary was generated from most frequent «IT-specific» words from this corpora (names, trade marks and words of general usage were removed).
2 stimluli lists were formed, first one — for respondents of technical specialities, second one — for humanities. 2nd list was translated to french.
1st list contained 253 stimuli, 100 words per form, number of filled formes — 126, different association pairs — 7,331, different reactions — 4,057.
2nd list contained 128 stimuli, 100 words per form, number of filled formes — 111, different association pairs — 5,328, different reactions — 3,318.
3rd list contained 118 stimuli, 60 words per form, number of filled formes — 113, different association pairs — 3,906, different reactions — 2,459.

4 . RESEARCH BASED ON ASSOCIATION EXPERIMENTS DATA

4.1 Web-sites optimization[12]

Verbal-association experiment makes possible to find non-evident key phrases and collocations. The list of associations based on the results of experiment mostly matches list of «evident» key phrases. Filippovich A. And Kirnarski A. experimented with web-site optimization [12]. They measured attendance for 8 months.
3 lists of key phrases were compared: «evident», verbal associations, russian thesaurus concepts.
Best results were obtained for first 2 lists..

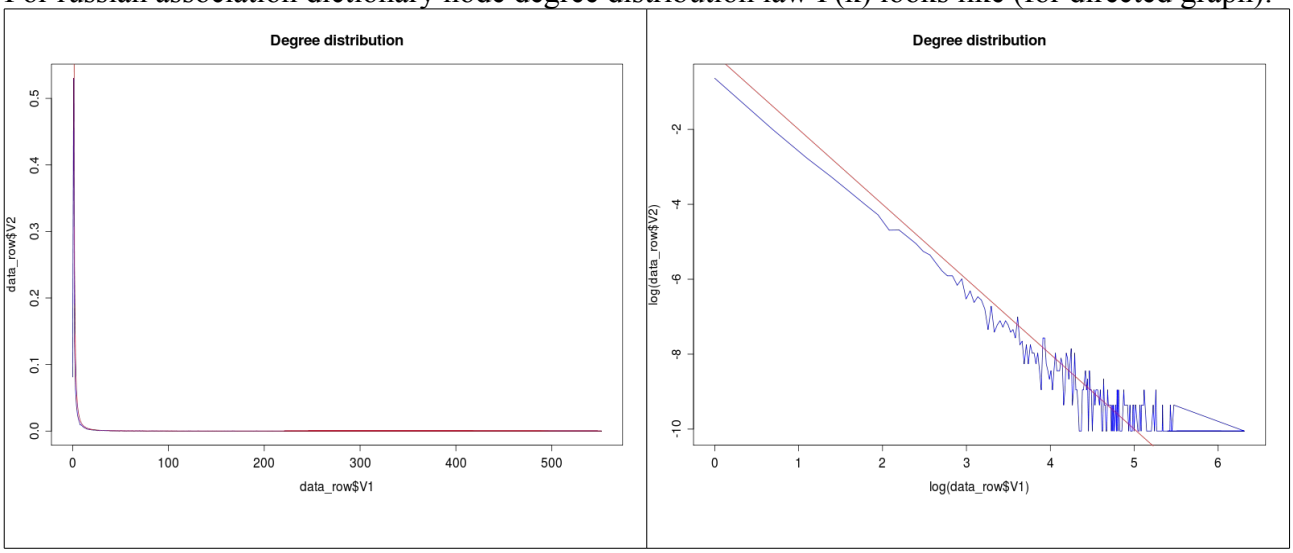
4.2 Dynamics of verbal associations model of russian consciousness[13]

Since russian association-comparaive dictionary was forming for a long period of time, the data from the dictionary could show changes in russian verbal consciousness.
With the help of this data one could not only observe the dynamics of these changes, but also the constant «core» of associations. For instance, the frequency of association like «bad->good», «dad->mom», «dot->comma», «real->man», «form->content», «pass->exam», «child's->babble», «nice->man» changed not more than 10% after 20 years. Such word combinations mostly contained synonyms, antonyms and stable combinations.

4.3 «Small-world» phenomenon

This phenomenon was first noticed by Migram in his experiments[14]. The essence is that any 2 US

citizens are connected on average through very small number of friends (usually it's called «6 degrees of separation»).
In 1998 Watts&Strogatz renewed interest to this phenomenon, they applied this theory to graphs and networks[15]. Such data structures were found in many kinds of networks, inculding WWW[16], networks of scientific collaborators [17], US power grid [15], metabolic networks in biology [18]. These networks usually have 2 types of distributions: for US power grid — exponentional, for WWW and metabolic networks — power law[19],
$P(k) \approx k^{-\gamma}$, где $\gamma \in [2..4]$.
Steyvers&Tenenbaum [20] analysed data from Nelson's free association experiment and discovered that node degree of graph created from the data has power distributions with $\gamma \in [3.01..3.19]$.
The most frequent concepts were «food», «money», «water», «car», «good».
For russian association dictionary node degree distribution law P(k) looks like (for directed graph):



In this case there is power distribution law with $\gamma = 2$.
The most frequent concepts are: «friend», «time», «child», «eat», «water» and others.
Analysis of this network showed that its main part is strongly connected and any node could be reached with not more than 6 steps (for undirected graph) [21].

4.4 Verbal consciousness modeling

Based on the results of association experiments there were attempts to create mathematical model for generating associations.
Mathematical models include:

- probabilistic graph with concepts as nodes and associations between them as edges. The task of discovering «close» and related associations transforms to shortest paths search problem[20];
- multidimensional scaling and singular value decomposition. These methods are usually used to decrease space dimensionality and save distances between objects at the same time[22].
  By distance between objects (words in current case) I mean euclidean distance or cosine of angle between 2 vectors.
  In case of multidimensional scaling we create semantic undirected graph (having edge weight as $S_{ij} = A_{ij} + A_{ji}$) and based on this graph we generate dissimilarity matrix (we store there distances between words, even those, which are not directly connected):
  $\hat{S}_{ij} = -\log(S_{ik} S_{kl} \dots S_{nj}) = -logS_{ik} - logS_{kl} - \dots - logS_{nl}$, where $S_{ik} S_{kl} \dots S_{nj}$ - weight of shortet path from node i to node j [23][24].
  After that we solve minimization problem: we try to minimize the error between real distance (difference) in vectors (representing concepts) and their «truth» value in

dissimilarity matrix.

Finally, this problem transforms to the problem of eigenvectors and eigenvalues discovering.

5 . CONCLUSION

In this article were represented main association experiments and some methods for their results analysis. Currently there is active research, development and comparison of neural, graph and tensor models for russian association dictionary. Also there is research of possiblity to apply these results for search quality improvement, homonimy discocvery and artificial modeling of complex logical chains.

BIBLIOGRAPHY

[1] Experimental Psychology. Vol. II.(1972) / Ed. by J.W. King and L.A. Riggs. Holt, Rinehart and Winston Inc.. 740 p.

[2] Gorodetskaya, L.(2002). Theory of communication and applied communication. / Ed. by I.N. Rozina, Rostov-on-Don: Institute of Management, Business and Law Publishing.

[3] Cramer, P. (1968).Word Association. NY: Academic Press.

[4] Deese, J. (1965). The Structure of Associations in Language and Thought. Baltimore, MD: The Johns Hopkins Press.

[5] Jenkins, J. J., & Palermo, D. S. (1964). Word Association Norms. Minn.:University of Minnesota Press

[6] De Groot, A. M. B. (1988). Woordassociatienormen met reactietijden (Word association norms with response times). Nederlands Tijdschrift voor de Psychologie (Dutch Journal of Psychology), 43, 280-296.

[7] Nelson, D. L., McEvoy, C. L., & Schreiber, T. A. (1999). The University of South Florida Word Association, Rhyme and Fragment Norms. http://luna.cas.usf.edu/~nelson/ .

[8] De Deyne, S. & Storms, G. (2008). Word associations: Network and semantic properties. Behavior Research Methods, 40, 213-231.

[9] Cherkasova, G.(2008). Russian comparative dictionary of associations, Moscow: IoL RAS.

[10] Cherkasova, G. (2004) Formal model of associative research.// Problems of applied linguistics. Ed. 2. / Moscow: «Azbukobnik», 400p.

[11] Philippovich, Y., Delft, D., Cherkasova, G. (2001). Associations in IT: exeriment on russian and french languages.: MGUM.

[12] Philippovich, A., Kirnarski, B(2008). Web-sites promotion based on the results of associative experiments.

[13] Cherkasova, G. (2007) Analysis of dynamics of russians' language cognition..//Questions of psycholinguistics, 6.. Moscow, p.105-122

[14] Milgram, S. (1967, May). The small-world problem. Psychology Today, 2, 60–67.

[15] Watts, D. J., & Strogatz, S. H. (1998, June 4). Collective dynamics of 'small-world' networks. Nature, 393, 440–442.

[16] Albert, R., Jeong, H., & Barabási, A. L. (1999). Diameter of the world wide web, Nature, 401,130-131

[17] Newman, M. E. J. (2001). The structure of scientific collaboration networks. Proceedings of the National Academy of Sciences, 98, 404–409.

[18] Jeong, H., Tombor, B., Albert, R., Oltval, Z. N., & Barabási, A. L. (2000). The large-scale organization of metabolic networks. Nature, 407, 651–654.

[19] Barabási, A. L., & Albert, R. (1999). Emergence of scaling in random networks. Science, 286, 509–512.

[20] Steyvers, M., Tenenbaum, J. (2005). The large-scale structure of semantic networks: statistical analyses and a model of semantic growth:Cognitive Science, 29(1).

[21] Sirenko A. (2007). Lingvocultural thesaurus of russian language. Intellectual technologies and systems.. Edition 9 / / Ed. by Philippovich Y., Moscow .: «Eliks+».

[22] Steyvers M., Shiffrin R., & Nelson D.(2004). Word Association Spaces for Predicting Semantic Similarity Effects in Episodic Memory. In A. Healy (Ed.), Experimental Cognitive Psychology and its Applications.

[23] Klauer, K., & Carroll, J. (1995). Network models for scaling proximity data.In R. D. Luce, M. D'Zmura, D. D. Hoffman, G. Iverson & A. K. Romney (Eds.), Geometric Representations of Perceptual Phenomena Mahwah, NJ: Erlbaum, 319-342.

[24] Cooke, N. , Durso, F. & Schvaneveldt, R.(1986). Recall and measures of memory organization. Journal of Experimental Psychology: Learning, Memory, and Cognition, 12, 538-549.