Psychology in Russia: State of the Art
Volume 12, Issue 1, 2019

Russian Psychological Society
Lomonosov Moscow State University

Comparing the Generation of Words from Different Semantic Categories in Native and Foreign Languages

Olga P. Marchenko*

Moscow State University of Psychology and Education, Moscow, Russia

* Corresponding author. E-mail: olga.p.marchenko@gmail.com

Objective. The aim of this study was to evaluate the degree of conformity of generation frequency scores for semantic categories between native (L1) and foreign languages (L2).

Design. Sixty-eight native Russian-speaking students were asked to list words belonging to different semantic categories in the Russian and English languages. Their generation frequency was calculated for each word in both languages. Hellinger Affinity scores were used to measure the conformity of the generation frequency norms between the Russian and English language.

Results. For culture-dependent categories, the level of conformity between the category frequencies in the Russian and English languages was greater than the conformity with North American norms. For culture-independent categories there were no significant differences in the level of conformity between the native and foreign languages as compared with North American frequencies. Furthermore, the number of concepts with significantly different levels of frequency between the Russian and English language was greater for culture-dependent categories than for culture independent-categories.

Conclusion. The low level of similarity for some categories can be explained by the subjects' different levels of experience using native and foreign languages in the context of these categories. A low level of category frequency accordance can be also explained by the fact that when subjects switch to a foreign language, they tend to name the concepts which are representative of the culture of this language. The strong level of similarity for other categories suggests that the vocabulary of these categories is less affected by cross-cultural and cross-linguistic diversity.

Keywords:
category norms, exemplar generation frequency, cross-language difference, Russian language, English language, L1, L2, Hellinger Affinity coefficient

ISSN 2074-6857 (Print) / ISSN 2307-2202 (Online)

http://psychologyinrussia.com

Introduction

The study of semantic categorization is one of the central problems being studied in cognitive psychology. Given that culture and conceptual behavior are tightly related, such questions need to be studied in the context of culture (Ojalehto & Medin, 2014). Cultural differences can be manifested in how semantic information is organized and used in tasks requiring categorization. One of the simplest and most widely available methods in the study of the features of categorization and organization of knowledge in long-term memory, is the method of generating items belonging to a particular category within a certain time interval (Voorspoels et al., 2013). This technique allows one to determine the content and structure of categories. For diagnostic purposes, one can determine the so-called category fluency (number of words in a category named during a certain period of time). Not only the category fluency but also the frequency of the naming of certain words in the category (generation frequency) is of particular interest.

Today, language is typically considered a part of culture. Language and other elements of culture do not act in isolation from each other, but are related by mutual influence and mutual reinforcement. Fixation on important cultural patterns in a language ensures its longevity and universality for all members of a given society (universal distribution within the culture). This is because the process of language acquisition automatically becomes the process by which person turns into a full member of that culture (Fausey et al., 2010). Therefore, the results of studies of bi-cultural individuals, in which it has been shown that a person changes his/her behavior after switching from one language to another, are quite understandable. Language, being a conductor of culture, can influence memory, thinking, self-esteem, and/or personality traits (Marian & Kaushanskaya, 2004; Oyserman et al., 2008; Ramírez-Esparza et al., 2006).

Category frequency and the richness of vocabulary in the designation of certain concepts (the "dictionary development" of the language) can be key indicators of the specific traits of different cultures (Wierzbicka, 2001). The most typical items within categories often coincide in different cultures (Yoon et al., 2004; Pekkala et al., 2009). However, along with the similarity of categories' content in different languages, there are also significant qualitative and quantitative differences caused by cultural factors (Pekkala et al., 2009; Eng et al., 2018). Cultural differences in categories' content and category fluency can be caused by many factors, such as differences in social, historical, economic, and educational systems; habitat; the degree of the population's familiarity with the category; as well as by purely linguistic factors such as word length (Pekkala et al., 2009).

The problem of inter-cultural stability and the diversity of categories naturally leads to the question of how bilinguals use categories. Bilinguals are worse than monolinguals in performing the categorical fluency task and demonstrate a smaller vocabulary even in their first language (Bialystok et al., 2008). Moreover, when bilinguals are tested, significant differences in the performance of a category fluency task in their native and foreign language may be found for some categories, while for other categories, their results are roughly the same (Ardila & Bernal, 2006; Grogan et al., 2009).

These results can be explained by two factors. First, the experience of interaction with objects from different categories may differ in the context of L1 and L2

usage. For example, it has been shown that emotional concepts in L2 do not evoke an emotional response (Pavlenko, 2012). Second, as explicated above, there may be some categories that are cross-culturally very stable, and other categories that are culturally specific. What's important is that the differences in the conceptual structures associated with L1 and L2 should be reflected in the differences in the structure of the categories when the subjects are tested in L1 and L2.

Empirically, the difference in the structure of a category in L1 and L2 can be operationalized by studying the similarity between the distributions of generation frequencies obtained for this category in L1 and L2. The similarity between the two frequency distributions can be computed by a number of means, one of which is the widely-used Hellinger Affinity (HA) coefficient (Yoon et al., 2004). Thus, the purpose of this study was to compare generation frequencies for different semantic categories generated for L1 and L2. Specifically, we computed HA scores for a number of categories reflecting the similarity (proximity) of generation frequency distributions obtained for these categories in L1 and L2. We expected that for some categories, the similarity measures would be greater than those for other categories, since a foreign language is used differently in different areas of human activity. In addition, we analyzed to what extent the categories' structures obtained for L1 and L2 were similar to the normative structures of these categories, as exhibited by native speakers of L2. This would allow us to show whether an individual starts to use a foreign language's conceptual structure when switching to L2.

Method

Sample

Sixty-eight students from the Moscow State Linguistic University, aged 18-27 years (58 women and 10 men, M=20.5, SD=2.2), participated in the study. All were native Russian speakers who had experience in learning English.

Procedure

The procedure used to obtain the categorical frequency scores was identical to the procedure developed by F.W. Battig and W.E. Montague (Battig & Montague, 1969). Participants received small notebooks in which they had to list items belonging to different categories. The experimenter read the instructions aloud and named the categories. Participants were asked to list as many items belonging to a category as they could in 30 seconds. In order to obtain a reliable result in computing the proximities between categories' structures, it was necessary to use a large number of different categories for testing (in many contemporary studies only one or just a few categories are used).

All participants performed the category fluency task in Russian (L1) and in English (L2). The order of languages was counterbalanced across participants. The category names were presented randomly to different groups of participants. The random order for both the native and foreign language was the same. Forty-five semantic categories of various types were selected (*Fruits, Fish, Insects, Vehicles, Furniture, etc.*). The frequency of each word in a category for Russian and English was calculated.

The similarity of frequency distributions

In order to measure the similarity (proximity) of the frequency distributions for each category between Russian and English, the Hellinger Affinity (HA) coefficients were used (Yoon et al., 2004). HA is calculated by summing the square root of the product of the two items' frequencies (p_i, q_i):

$$HA(p, q) = \sum\sqrt{p_i q_i}$$

The value of Hellinger Affinity coefficients ranges from 0 to 1, with 1 meaning that two frequency distributions are identical.

The median number of words named in each category in 30 seconds was measured. The median number of named words in Russian and English and their HA scores were then compared using the Wilcoxon T-Test. For each word, overall generation frequencies were compared between L1 and L2, with the help of Pearson's χ^2-test with Yates' correction for continuity. The number of significant differences ($p<0.01$) was compared between regions, using the one-tailed Pearson χ^2-test as well. The HA scores between L1, L2, and North American norms were compared using the Friedman and T-Wilcoxon tests. North American generation frequency norms for 36 categories in English language were taken from Van Overschelde et al. (2004). Intra-cultural HA scores for Russian generation frequencies were taken from Marchenko et al. (2018).

Results and Discussion

The average number of words named in L2 in 30 seconds for a category was less than the number of words named in L1 (6.56±2.8 vs. 4.27±2.8, $Z=-5.744$, $p<0.001$). Such differences can be explained by the fact that the size of a person's vocabulary is significantly smaller in a foreign language than in his/her native language. In addition, it was suggested that the naming speed in the L2 might be reduced because a person's ability to express knowledge in a foreign language is not as automatic as it is in his/her native language. The ease and speed of access to a lexicon depend upon the age of acquisition, frequency, and recency of access (Snodgrass & Tsivkin, 1995).

It has also been observed that even balanced bilinguals (these who have approximately similar experience using both languages in different areas of human activity), then using their dominant language, perform worse on the category fluency task than monolinguals (Bialystok et al., 2008). There are several explanations for this deficit. Perhaps bilinguals cannot suppress interference which is caused by the activation of another language. Perhaps in bilinguals, each language is supported by weaker ties connecting concepts with words, as compared to monolinguals who use only one language (Bialystok et al., 2008). In addition, if the word is activated in Russian at the first moment of the task performance, and it has to be translated into a foreign language, then it takes more time to name the word in English than when the task is performed in the native language.

Hellinger's Affinity coefficients between the frequency distributions of words generated for all categories in English and in Russian are presented in *Table 1*.

Table 1

Hellinger's Affinity scores for category frequency distributions in L1, L2, North American category norms, and Russian category norms.

Category	Hellinger's Affinity					
	L1&L2	**L2&NA**	**L1&NA**	**RN1**	**RN2**	**RN3**
Alcoholic Beverages	0.91	0.83	0.76	0.97	0.95	0.96
Amphibians	0.90	–	–	0.93	0.91	0.92
Birds	0.75	0.61	0.63	0.96	0.95	0.97
Body Organs	0.78	–	–	0.97	0.96	0.95
Body Parts	0.88	0.92	0.87	0.96	0.96	0.96
Furniture	0.84	0.68	0.74	0.96	0.96	0.96
Carpenter`s Tools	0.73	0.66	0.70	0.92	0.92	0.92
Clothing	0.83	0.76	0.75	0.96	0.95	0.96
Colors	0.92	0.94	0.91	0.98	0.98	0.98
Countries	0.90	0.79	0.73	0.95	0.95	0.95
Crimes	0.58	0.68	0.75	0.91	0.90	0.91
Diseases	0.53	0.46	0.57	0.91	0.91	0.91
Distance Units	0.89	0.93	0.85	0.97	0.97	0.96
Domestic Animals	0.91	–	–	0.96	0.95	0.97
Domestic Appliances	0.82	–	–	0.96	0.94	0.95
Fabrics	0.86	0.76	0.79	0.94	0.94	0.93
Family Members	0.90	0.92	0.86	0.98	0.98	0.98
Farm Animals	0.91	–	–	0.98	0.98	0.98
Female Names	0.50	0.33	0.04	0.91	0.93	0.91
Fish	0.51	0.60	0.64	0.91	0.84	0.87
Flowers	0.72	0.63	0.67	0.95	0.91	0.92
Foods	0.68	–	–	0.89	0.89	0.89
Four-footed Animals	0.89	0.90	0.81	0.96	0.96	0.96
Fruits	0.91	0.92	0.89	0.98	0.97	0.97
Insects	0.82	0.86	0.87	0.96	0.96	0.96
Kitchen Utensils	0.78	0.51	0.80	0.94	0.91	0.94
Male Names	0.37	0.57	0.03	0.93	0.93	0.92
Mammals	0.89	–	–	0.95	0.95	0.94
Metals	0.87	0.81	0.88	0.94	0.98	0.94
Musical Instruments	0.82	0.76	0.76	0.97	0.96	0.96
Nonalcoholic Beverages	0.84	0.72	0.61	0.95	0.95	0.96
Plants	0.64	–	–	0.87	0.94	0.87
Precious Stones	0.77	0.81	0.85	0.96	0.95	0.96
Professions	0.70	0.67	0.52	0.87	0.87	0.88
Reptiles	0.83	–	–	0.97	0.96	0.97
Sciences	0.83	0.66	0.65	0.9	0.92	0.89
Sports	0.72	0.72	0.70	0.93	0.93	0.94
Time Units	0.91	0.95	0.92	0.98	0.97	0.97
Toys	0.70	0.60	0.61	0.89	0.88	0.91
Trees	0.64	0.69	0.71	0.96	0.95	0.97
Types of Music	0.89	0.26	0.79	0.92	0.92	0.94
Vehicles	0.86	0.76	0.68	0.96	0.93	0.95
Vegetables	0.85	0.65	0.62	0.98	0.97	0.98
Weapons	0.71	0.75	0.66	0.92	0.92	0.92
Wild Animals	0.80	0.83	0.76	0.96	0.94	0.94

Notes. NA=North American norms; RN1=HA scores for Moscow-Yekaterinburg data; RN2 =HA scores for Moscow-Irkutsk data; RN3=HA scores for Yekaterinburg-Irkutsk data.

We performed a series of analyses of the HA scores. We analyzed the distribution of the HA scores between categories obtained for L1 and L2. As can be seen from the *Table 1*, the HA for L1 and L2 differ markedly between categories. There are categories with high HA scores (like *Alcoholic Beverages, Colors, and Domestic Animals*). There are also categories with much lower HA scores (like *Female Names, Mammals, and Precious Stones*). These results suggest that there are categories that have very similar structures when produced in a native and in a foreign language, but that there are also categories that have dissimilar structures when produced in a foreign language.

For categories with high L1&L2 HA scores, it can be suggested that either the subject's English language vocabulary is better developed, or that there are smaller cross-cultural differences for these categories. The lower L1&L2 HA scores can be explained by strong cross-cultural differences in these categories. Exploring these ideas in further research will require sophisticated statistical analyses of the shape of the L1&L2 HA scores distribution, which should be bimodal if there are indeed two distinct classes of categories.

Some analyses of these ideas can be also done with the current data. For example, some categories in our list may be highly dependent on culture (like categories of non-living things), while other categories (like categories of living things) are not. The first type should be culture-specific and, as such, items in that category should have lower L1&L2 HA scores. The second should be more cross-culturally stable and should have higher L1&L2 HA scores. We divided all categories into two groups (culture-dependent and culture-independent), but found no differences between mean L1&L2 HA scores between the groups ($p>0.05$).

However, we compared the frequencies of all concepts inside the categories between L1 and L2 separately. This statistically more powerful comparison showed that the number of concepts with significantly different concept frequencies between L1 and L2 was greater for culture-dependent categories than for culture-independent categories ($\chi^2=5.6$, $df=1$, $p<0.01$), thus supporting the distinction between the two types of categories.

Second, we compared the HA scores for categories generated in L1 and L2 to HA scores for categories generated in L1 (Russian) in several regions of Russia (Moscow, Yekaterinburg, and Irkutsk). These last scores reflect the level of intra-cultural stability of categories in Russian. It should be noted that comparison of category frequency norms within a single culture usually shows relatively high HA scores, indicating the high intra-cultural stability of categories (Marchenko et al., 2018; Marchenko et al. 2016; see the last three columns in *Table 1*). This is true even if the participants belong to different generations (Yoon et al., 2004), or if the results of two tests are separated in time (Marful et al., 2015). In this study, we replicated this result. The HA scores of L1 and L2 frequency distributions were significantly lower than the intra-cultural HA scores between frequency distributions of the three different regions of Russia ($Z=-5.845$, $p<0.001$; $Z=-5.844$, $p<0.001$; $Z=-5.845$, $p<0.001$).

Third, we compared the category frequency distributions obtained in our study for L1 (Russian) and L2 (English) with the normative frequency distributions collected in North America for English (North American norms=NA). For culture-dependent categories, L1&L2 HA scores were greater than the proximity scores

between L1 categories and NA, and between L2 categories and NA (χ^2=13.195, p=0.001; L1vsNA: Z=–3.391, p=0.001; L2vsNA: Z=–2.523, p=0.012). For culture-independent categories, the L1&L2 HA scores did not differ from both the L1&NA and L2&NA HA scores (χ^2=0.400, p=0.819; L1vsNA: Z=–0.455, p=0.649; L2vsNA: Z=–0.142, p=0.887). This result clearly supports the notion that there are different subtypes of categories which are more or less universal across cultures. Interestingly, this result also suggests that (for culture-dependent categories) when an individual switches to a foreign language, there is still a strong influence of the native language's conceptual organization on category generation in the foreign language.

On a general note, our results support the idea that when switching to a foreign language, an individual activates conceptual knowledge associated with the use of the foreign language. We frequently observed that when a subject switched to English, he/she began to name concepts which are inherent to the culture of that language (for example, English names in *Male* and *Female Names,* or *mile* and *foot* instead of *kilometer* and *meter*).

However, this idea should be qualified, as suggested, for instance, by our last empirical result. When a person generates items for categories in L2, activation of his conceptual knowledge takes place. In cases where there has been little direct interaction with the objects in the context of L2, the person will probably translate concepts that are familiar to him or her in the context of L1. This may explain how individuals, while to some extent switching to the culture representing a foreign language when they are using it, at the same time may be influenced by the conceptual knowledge of the native language. So, the degree of cross-cultural universality of generated categories, and the experience of the interaction with exemplars from these categories within different linguistic contexts, both affect performance in cross-language category fluency tasks.

Conclusion

This study demonstrated that categorical fluency rates are greater for an individual's native language than for a foreign language. When a person switches to a foreign language, he/she begins to use the words inherent in the culture of that language. The similarity between the frequency distributions of words generated in L1 and L2 is lower than the similarity of such data within one culture. Such differences, connected with the specificity of a person's experience using the foreign language, must be taken into account when assessing the features of semantic categorization. There is evidence that an individual using another language begins to behave as a representative of the culture whose language he uses. Such transformation is echoed in the differences of categorical norms obtained in a cross-linguistic category fluency task in this study.

Although appearing merely theoretical at first glance, these results can also have practical consequences. First, in the process of studying a foreign language, an individual usually acquires culture-specific words, which leads to the establishment of a foreign culture-specific vocabulary. Such data can be used to assess the vocabulary of individuals learning a foreign language, as well as to assess the level of proficiency in that new language. The practice of introducing new vocabulary in

semantic categories is a popular approach in second language teaching (Finkbeiner & Nicol, 2003).

Second, our results corroborate the previous findings that the category fluency of bilinguals, even in their native language, can be significantly different from that of monolinguals (Baus et al., 2013). This makes it necessary to develop separate category norms for such individuals, which can be used to diagnose different cognitive impairments, since multicultural contacts and the use of more than one language is a hallmark of the modern world, at least for the large population of Russian-speaking individuals who study and use a foreign language.

Acknowledgements

This work was supported by the Russian Foundation for Basic Research, grant 17-36-01131.

References

Ardila, A. & Bernal, B. (2006). Cognitive testing toward the future: The example of Semantic Verbal Fluency (ANIMALS). *International Journal of Psychology,* 41, 324–332. https://doi.org/10.1080/00207590500345542

Battig, W.F. & Montague, W.E. (1969). Category norms of verbal items in 56 categories, A replication and extension of the Connecticut category norms. *Journal of Experimental Psychology,* 80, 1–45. https://doi.org/10.1037/h0027577

Baus, C., Costa, A., & Carreiras, M. (2013). On the effects of second language immersion on first language production. *Acta psychological,* 142, 402–409. https://doi.org/10.1016/j.actpsy.2013.01.010

Bialystok, E., Craik, F., & Luk, G. (2008). Lexical access in bilinguals: Effects of vocabulary size and executive control. *Journal of Neurolinguistics,* 21, 522–538. https://doi.org/10.1016/j.jneuroling.2007.07.001

Eng, N., Vonk, J., Salzberger, M., & Yoo, N. (2018). A Cross-Linguistic Comparison of Category and Letter Fluency: Mandarin and English. *Quarterly Journal of Experimental Psychology.* https://doi.org/10.1177/1747021818765997

Fausey, C., Long, B., Inamori, A., & Boroditsky, L. (2010). Constructing Agency: The Role of Language. *Frontiers in psychology,* 1, 162. https://doi.org/10.3389/fpsyg.2010.00162

Finkbeiner, M. & Nicol, J. (2003). Semantic category effects in second language word learning. *Applied Psycholinguistics,* 24, 369–383. https://doi.org/10.1017/S0142716403000195

Grogan, A., Green, D.W., Ali, N., Crinion, J., & Price, C.J. (2009). Structural Correlates of Semantic and Phonemic Fluency Ability in First and Second Languages. *Cerebral cortex*, 19, 2690–8. https://doi.org/10.1093/cercor/bhp023

Marchenko, O.P., Bandurka, T.N., & Pavlov, Y.G. (2016). Geographical Stability of Category Norms for Russian Language. *Russian Linguistic Bulletin, 4*(8), 12–16. doi: 10.18454/RULB.8.25.

Marchenko, O.P., Bandurka, T.N., & Pavlov, Yu.G. (2018). Intra-cultural continuity and diversity of semantic category norm. *Russian Linguistic Bulletin, 2*(14), 3–7. doi: 10.18454/RULB.2018.14.2.9

Marful, A., Díez, E., & Fernandez, A. (2015). Normative data for the 56 categories of Battig and Montague (1969) in Spanish. *Behavior Research Methods, 47*(3), 902–910. https://doi.org/10.3758/s13428-014-0513-8

Marian, V. & Kaushanskaya, M. (2004). Self-construal and emotion in bicultural bilinguals. *Journal of Memory and Language,* 5, 190–201. https://doi.org/10.1016/S0749-596X(04)00045-2

Ojalehto, B. & Medin, D. (2014). Perspectives on Culture and Concepts. *Annual review of psychology.* 66. https://doi.org/10.1146/annurev-psych-010814-015120

Pavlenko, A. (2012). Affective processing in bilingual speakers: Disembodied cognition? *International journal of psychology: Journal international de psychologie,* 47, 405–28. https://doi.org/10.1080/00207594.2012.743665

Pekkala, S., Goral, M., Hyun, J., Obler, L., Erkinjuntti, T., & Albert, M. (2009). Semantic verbal fluency in two contrasting languages. *Clinical linguistics & phonetics,* 23, 431–45. https://doi.org/10.1080/02699200902839800

Ramírez-Esparza, N., Gosling, S., Benet, V., Potter, J.P., & Pennebaker, J. (2006). Do bilinguals have two personalities? A special case of frame switching. *Journal of Research in Personality,* 40. 99-120. https://doi.org/10.1016/j.jrp.2004.09.001

Snodgrass, J.G. & Tsivkin, S. (1995). Organization of the bilingual lexicon: Categorical versus alphabetic cuing in Russian-English bilinguals. *Journal of Psycholinguistic Research,* 24(2), 145–163. https://doi.org/10.1007/BF02143960

Wierzbicka, A. (1997) *Understanding Cultures through Their Key Words: English, Russian, Polish, German, and Japanese.* New York: Oxford University Press.

Yoon, C., Feinberg, F., Hu, P., Gutchess, A.H., Hedden, T., Chen, H., … & Park, D.C. (2004) Category norms as a function of culture and age: Comparisons of item responses to 105 categories by American and Chinese adults. *Psychology and Aging,* 19, 379–393. https://doi.org/10.1037/0882-7974.19.3.379

Van Overschelde, J.P., Rawson, K.A., & Dunlosky, J. (2004). Category norms: An updated and expanded version of the Battig and Montague (1969) norms. *Journal of Memory and Language,* 50, 289–335. https://doi.org/10.1016/j.jml.2003.10.003

Voorspoels, W., Storms, G., Longenecker, J., Verheyen, S., Weinberger, D.R., & Elvevåg, B. (2013). Deriving semantic structure from category fluency: Clustering techniques and their pitfalls. *Cortex,* 55. https://doi.org/10.1016/j.cortex.2013.09.006

Original manuscript received October 29, 2018
Revised manuscript accepted January 10, 2019
First published online March 15, 2019

To cite this article: Marchenko, O.P. (2019). Comparing the generation of words from different semantic categories in native and foreign languages. Psychology in Russia: State of the Art, 12(1), 163–171. DOI: 10.11621/pir.2019.0112