

Possible Cognitive Mechanisms for Identifying Visually-presented Sound-Symbolic Words

Liubov O. Tkacheva, Yulia G. Sedelkina*, Andrey D. Nasledov

St. Petersburg State University, St. Petersburg, Russia

* Corresponding author. E-mail: yulsedspsb@gmail.com

Background. Sound symbolism (SS) refers to the direct association between the sound and the meaning of a word. The results of cross-linguistic research prove that SS is universal for different languages and cultures. Thirty percent of all natural languages consist of SS words. But despite the large number of psychosemantic studies which have been conducted, the cognitive mechanisms of the perception of SS words still remain unclear.

Objective. The aim of this study was to determine how Russian-speaking subjects perceive visually presented English and Russian words, as measured by the Lexical decision method.

Design. The study sample consisted of 148 subjects of ages ranging from 13 to 78. The study was conducted in two stages. During the first stage, the perception of visually-presented English SS words by Russian learners of English, with three different levels of language proficiency, was studied. During the second stage, the perception of visually-presented Russian SS words by Russian native speakers from three different age groups was studied.

The stimuli material was selected according to the following criteria: 1) each word was monosyllabic; 2) Each SS word corresponded to a single arbitrary (non-SS) word of the same pronunciation type; and 3) Each word corresponded to a non-word, formed from it by replacing letters according to the phonotactic rules of English and Russian. At each stage of the study, each subject was given 80 stimuli consisting of 20 SS words, 20 non-SS words, and 40 non-words. An analysis of contingency tables (Chi-square test), comparison of averages (Student's t-test), and analyses of variances (ANOVA) were applied to the data.

Results. The visually-presented SS words were identified more slowly and with more errors than the non-SS words, regardless of the language (Russian or English), the subjects' age, and their English language proficiency.

Conclusions. The observed delay effect in the cognitive processing of visually-presented SS words is due to the cognitive complexity of the task, which leads to the activation of cross-modal interaction system, besides, interfering systems of information processing are assumed to exist.

Keywords:
phonosemantics,
psychosemantics,
sound-iconicity,
sound symbolism
(SS), lexical
decision task

Introduction

Despite the abstract nature of natural languages, 30% of the words in all of them retain an explicit or implicit link between the signified object and the signifying word (Armoskaite, 2017). Such words, which feature close proximity between their acoustic form and their meaning, are called sound-symbolic (SS). Amongst them, onomatopoes and ideophones can be distinguished, depending on whether they signify acoustic or non-acoustic sensations (Voronin, 2006). Thus, SS words evoke ideas of sensual images which can be represented in a wide range of perceptions within all sensory modalities—from sound to movement or texture, from appearance to internal feeling (Dingemanse, 2018).

For a long time most of the research on SS has focused on its phonological, morpho-syntactic, and symbolic aspects (Dingemanse, 2012; Sidhu, 2017), as well as its universal character for all natural languages (Kazuko, 2010; Svantesson, 2017). Lesser works studied the meaning and practical use of SS. Only in the 2000s did the rich sensory meaning of SS words and cross-modal interaction in the process of their decoding begin to be studied (Ameka, 2001). It has been found that, in addition to the direct link with the auditory system of perception, SS extends to other sensory modalities, such as visual, motor, and tactile systems; they sometimes may include internal visceral sensitivity and trigger psychological conditions (Akita, 2009). Moreover, this hierarchy of the sensory impact of SS can be characterized by a multifaceted semantic map, with several possible trajectories of semantic expansion and grooves (Van der Auwera, 2006).

Neurolinguistic studies have set a new milestone, marked by the search for physiological correlates of the cognitive processes involved in decoding SS. In the early 2000s, data was obtained on the high plasticity of the sensory areas of the cerebral cortex, and the propensity for various sensory modalities to interact at the cortical level (Shimojo & Shams, 2001; Ramachandran & Edward, 2001). The most pronounced manifestation of such cross-modal interaction can be found in the studies of how young children perceive SS. It has been shown that, due to the close connections between the sensory areas of the brain, various sensory modalities for speech sounds are spontaneously activated in infants (Walker, 2010). Cross-modal interaction facilitates the understanding of words, allowing the child to concentrate on the referents built into the complex scene (Imai, Kita, 2014). It has been suggested that this early capability for cross-modal interaction may later turn into a more abstract system of symbols (Cytowic & Eagleman, 2009).

Furthermore, it is recognized that SS plays a significant role in teaching languages (Imai, 2008; Laing, 2014; Sedelkina, 2016) and in natural spontaneous communication (Perniss & Vigliocco, 2014; Clark, 2016). Experiments with MRI proved that mirror (mimic) neurons are responsible for receiving onomatopoeia (Osaka, 2006), and that they are also activated in the auditory perception of the sounds easily associated with the recognition task (Rizzolatti & Craighero, 2005). In addition, it has been confirmed that the multisensory interaction in the auditory perception of SS words, is associated with the emotion of laughter (Osaka, 2003).

Taking into account the lack of a clear description of the cognitive mechanisms involved in the processing of visually presented SS words, we decided to compare the speed and accuracy of identification of visually presented SS words, in contrast to non-SS words.

Method

Materials

To collect the data, the “Lexical decision” method was used (Ratcliff, Gomez, & McKoon, 2004) as part of the software complex for longitudinal research (Miroshnikov, 2001). The research was carried out in two stages. During the first, the perception of visually presented English SS words by Russian learners of English with three different levels of language proficiency, was studied. During the second, the perception of visually presented Russian SS words by Russian native speakers from three different age groups, was studied.

Table 1

The stimuli of the first stage

SS words	Arbitrary words	Non-words	
peak	deep	heep	feep
clap	luck	clatt	claff
knock	map	moff	nak
click	pink	stim	pimk
crash	trash	prash	grash
wow	hour	bout	vout
pump	stamp	tunk	pank
bat	cat	pab	cag
tap	top	dod	taf
wind	band	wint	bant
kick	sick	kif	tith
bell	bill	gell	pell
flow	low	fow	lau
glance	chance	lunce	hunce
fly	life	thly	gly
scream	cream	rean	reang
slide	side	lide	shide
slip	pill	silp	siple
snake	save	smake	snate
jump	just	junt	chunt

In the first stage, stimuli material was selected according to the following criteria:

- 1) All semantic stimuli were monosyllabic English words from the PET vocabulary list (PET Vocabulary List, 2011), representing a B1 level of language proficiency according to the Common European Framework of References for Languages: Learning, Teaching, Assessment (CEFR). This was the level of the majority of the subjects;
- 2) Each SS word corresponded to a single non-SS word similar in quality and quantity of vowels and consonants; and
- 3) Each word corresponded to a non-word formed from it by replacing letters according to the phonotactic rules of English, so that it was similar in composition and quality of vowels and consonants.

The SS stimuli consisted of 11 onomatopes and 9 ideophones. Onomatopes were collected from the lexical lists presented in S.V. Voronin's thesis (Voronin, 1969). Ideophones were collected from the phonosemantic dictionary of M. Mag-

Table 2

The stimuli of the second stage

SS words	Arbitrary words	Non-words	
plyukh (splash)	slukh (hearing)	flyukh	khlus
bukh (bounce)	buk (beech)	buj	bun
chmok (peck)	srok (term)	kmok	ksor
shcholk (click)	sholk (silk)	shchokl	shlyok
khlop (clap)	klop (bug)	khlok	klap
gav (woof)	rov (ditch)	vag	rav
pisk (squeak)	risk (risk)	sipk	skipr
bakh (wham)	bar (bar)	khab	rap
skrip (creak)	krest (cross)	skirb	sterk
plesk (splash)	press (press)	pleks	spers
khlyup (squelch)	klub (club)	khluk	bluk
tresk (crack)	trest (group)	tersk	stret
khруп (crunch)	trup (corpse)	prukh	rupt
vizg (scream)	disk (disk)	zvig	ksid
lyazg (clank)	glaz (eye)	zyagl	zagl
khrip (groan)	khrom (chrome)	prikh	mokhr
chirk (strike)	tsirk (circus)	krich	kritis
svist (whistle)	tvist (twist)	stisv	svitt
stuk (knock)	kust (bush)	tusk	skut
pshik (puff)	shpik (pork fat)	piksh	shipk

nus (Magnus, 2017). The meaning of the sound combinations with a symbolic sense was checked according to the tables of data from statistical research on SS (Drellishak, 2006). The stimuli used in the first stage of the research are presented in *Table 1*.

In the second stage, the stimuli material was selected according to similar criteria:

- 1) All words were monosyllabic, collected by the method of continuous sampling from etymological (Fasmer, 1986) and phonosemantic dictionaries (Shlyachova, 2004);
- 2) Each SS word corresponded to a single non-SS word of the same acoustic type, e.g. bukh (SS word, meaning bounce) to buk (arbitrary, non-SS word, meaning beech); and
- 3) each word corresponded to a non-word, formed from it by replacing letters according to the phonotactic rules of Russian, so that it is similar in quality and quantity of vowels and consonants, e.g. slukh (a word meaning hearing) to flyukh (non-word).

The SS stimuli consisted of 20 onomatopoeic words which presented all types of phonosemantic sounds: instant (bukh, bakh, stuk, khlop, khlyup), continuant (vizg, gav, pisk, svist, chmok, pshik), frequentative (skrip, tresk, khrip, khруп, chirk), and integrated sounds (lyazg, plesk, plyukh, shcholk) that combine the characteristics of instant, continuant, and frequentative sounds. The transliterated Russian stimuli used in the second stage of the research, along with their English equivalents in brackets, are presented in *Table 2*.

Procedure

The study was carried out according to the classical “Lexical decision” method (Meyer & Schvaneveldt, 1971; Ratcliff, Gomez, & McKoon, 2004) during both stages of the research. A subject received instructions (first orally, and then visually on the screen) explaining the task sequence and telling him/her to make a decision as quickly as possible. Then, 20 SS words, 20 non-SS words, and 40 non-words were presented on the screen in random order one by one (see *Tables 1* and *2*). The subject’s task was to identify the presented stimulus as a word or non-word by pressing the button which corresponded to the type of the stimulus. Identification time was restricted to no more than 1000 ms. We collected data on the time required for identification, the number of errors, and the number of delays. The experimental session was preceded by a training one, where 10 words and 10 non-words were presented in random order.

The sample

In total, 148 persons were surveyed. The first stage of the study involved 90 participants, among them 25 male and 65 female, aged 17 to 20 years, who were divided into four groups according to the level of their English language proficiency: 1) 0–A1: 9 people; 2) A2–B1: 15 people; 3) B1–B2: 54 people; and 4) B2+: 12 people. All participants were Russian-speaking first-year B.A. students studying at the faculty

of Asian and African Studies, the faculty of Philology, and the faculty of Psychology of St. Petersburg State University who had studied English as a foreign language.

The second stage of the study involved 58 Russian-speaking subjects, divided into three groups according to their age: 1) 15 years old and younger: 4 people; 2) 15-50 years old: 31 people; and 3) 51 year old and older: 23 people. The entire sample included 23 males and 35 females.

Statistical data analysis

In total, 5902 target stimuli were presented. The distributions of correct answers, errors, and delays for SS and non-SS words were compared using the chi-square test. To compare the reaction times for word recognition, the average reaction time for SS and non-SS words was calculated for each subject; hereafter the sequence of stimulus was presented as repeated measures. The comparison was made using the Student's t-test for dependent samples. To check the influence of the level of language proficiency/age on the time taken to recognize the words, a 2-factor analysis of variance with repeated measures ANOVA 2x4 (type of stimuli) and the dependent variable "time" was carried out (Nasledov, 2013). All statistical analysis was performed using IBM SPSS software version 24.

Results

The distributions of delays, correct answers, and mistakes for SS and non-SS English words are presented in *Table 3*.

Table 3

Contingency table of the "Type of stimuli" x "Accuracy" for English words

Words		Accuracy			total
		delay	correct	mistake	
SS	Amount	40	1496	264	1800
	%	2.2%	83.1%	14.7%	100.0%
Non-SS	Amount	34	1595	171	1800
	%	1.9%	88.6%	9.5%	100.0%
Total	Amount	74	3091	435	3600
	%	2.1%	85.8%	12.2%	100.0%

The differences are statistically significant (chi-square=22.606; df=2; p <0.001). The number of correct answers is statistically significantly lower for SS words (83.1%) than for non-SS (88.6%), due to the increase in the number of delays and mistakes.

The times taken for word recognition were compared using the Student's t-tests for dependent samples. The results are shown in *Table 4*. Differences were found at a high level of statistical significance (t=4.542; df=89; p<0.0001; R2 =0.188): the

average time for SS words recognition is more than that for recognizing non-SS words. The difference in the type of stimulus (SS words /non-SS words) explains 18.8% of the differences in time recognition.

Table 4
Descriptive statistics of reaction time for English words

Words	Mean (MS)	N	Standard deviation	Error of mean
SS words	642.4328	90	72.29333	7.62039
non-SS words	625.4398	90	68.87652	7.26022

According to the results of a 2-factor analysis of variance with repeated measures ANOVA 2x4 (type of stimuli) and the dependent variable “time”, the effect of the interaction of factors was statistically insignificant ($F(3; 86) = 2.017; p < 0.118$). Thus, the difference in the time required for recognition of SS and non-SS words is manifest regardless of the level of a subject’s language proficiency.

The average values of the word recognition time, depending on the level of language proficiency and the type of stimulus (SS word /non-SS word), are shown in *Figure 1*.

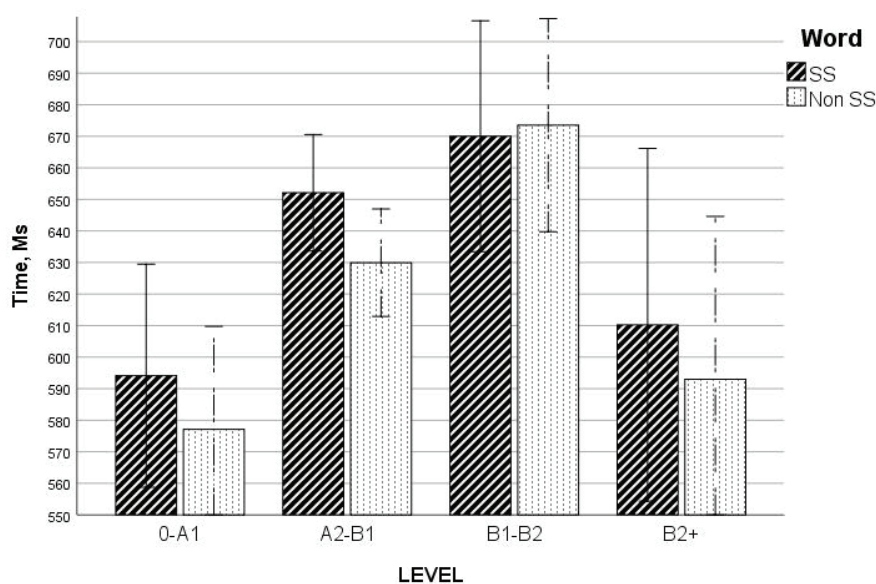


Figure 1. Average values of the recognition time for English words.

In *Table 5* the distributions of delays, correct answers, and mistakes for SS and non-SS Russian words are presented.

The differences are statistically significant ($\chi^2 = 25.253; df = 2; p < 0.0001; \phi = 0.105$). The number of correct answers is statistically significantly lower for SS words (70.4%) than for non-SS words (79.5%). The effect was tested for each of the three age groups. For the group “up to 15 years,” the effect was statistically

unreliable, likely due to the small number of the sample (chi-square= 0.144; df=2; $p < 0.931$). However, for the other two groups, the effect is statistically significant. For the sample “from 15 to 50 years” the proportion of correct answers for SS words was 73.6%, and for non-SS words, it was 85.5% (chi-square=30.062; df=2; $p < 0.0001$). For the sample “over 50 years,” for SS words it was 70.4%, and for non-SS, 79.5% (chi-square=6.851; df=2; $p < 0.033$).

Table 5

Contingency table of the “Type of stimuli” x “Accuracy” for Russian words

Word		Accuracy			Total
		delay	correct	mistake	
SS	Amount	139	831	210	1180
	%	11.8%	70.4%	17.8%	100.0%
non-SS	Amount	101	938	141	1180
	%	8.6%	79.5%	12.0%	100.0%
Total	Amount	240	1769	351	2360
	%	10.2%	74.8%	15.0%	100.0%

Table 6

Descriptive statistics of reaction times for Russian words

Words	Mean (MS)	N	Standard deviation	Error of mean
SS word	743.73	58	63.154	8.292
non-SS word	713.22	58	65.876	8.650

The time for word recognition was compared using the Student’s t-test for dependent samples; the results are shown in *Table 6*. Differences were found at a high level of statistical significance ($t=5.460$; $df=57$; $p < 0.0001$; $R^2=0.343$): the average time for SS words recognition is more than that for non-SS words. The difference in the type of stimulus (SS words /non-SS words) explains 34.3% of the difference in time recognition.

To check the influence of age on the time of word recognition, a 2-factor analysis of variance with repeated measures ANOVA 2x3 (type of stimuli) was undertaken, with the dependent variable “time” (ms).

Statistically significant effects of the factor “type of stimuli” ($F(1; 55)=17.883$; $p < 0.0001$) and the factor “age” ($F(2; 55)=5.978$; $p < 0.004$) were found. The effect of the interaction of factors was statistically insignificant ($F(2; 55)=0.953$; $p < 0.392$). Thus, the difference in time recognition of SS and non-SS words is manifested re-

ardless of age. The average values of the word recognition time depending on the age group and the type of stimulus (SS words / non-SS words) are shown in *Figure 2*.

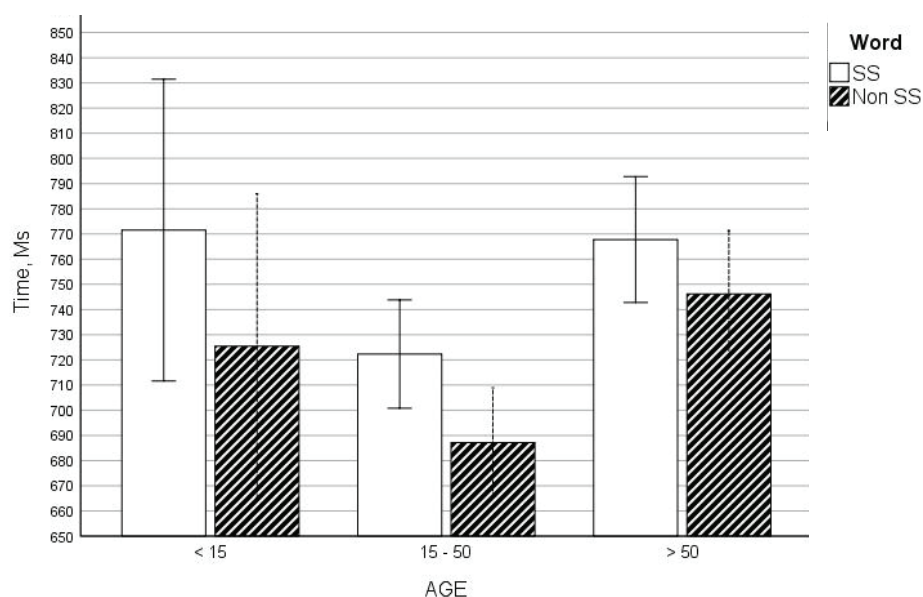


Figure 2. Average values of the recognition time for Russian words.

Discussion

The results show that Russian native speakers recognize both Russian and English SS words more slowly and less correctly than non-SS words when they are presented visually. Furthermore, the magnitude of this effect for Russian SS words ($R^2=0.343$) is higher than for English ones ($R^2=0.188$). The result is also interesting because, in total, 40 pairs of stimuli were used in the experiment, while in the majority of studies of SS, no more than eight pairs of stimuli were used (Westbury, 2018).

It would appear that the observed delay in the identification of words is caused by the cognitive complexity of the task, since, in addition to the cross-modal interaction system, two interfering systems of information processing are supposed to be activated. One is the system of decoding semantic information automated in ontogeny, associated with the left-brain contour of functional dominance with dominant right-handedness, and other is the figural system of decoding information that requires the activation of right hemisphere resources.

In the EEG experiment on perception of words and non-words, consistent coherence in the beta range in the left hemisphere was recorded when the subjects perceived the words (von Stein, 1999). Similarly, in EEG research on understanding visually presented texts with the increasing completeness of information, high-frequency activity with the predominant involvement of the left hemisphere in the phase of idea generation was recorded as well (Tkacheva, 2015).

What is also significant is that the delay in recognizing SS words remained regardless of age. But it is also necessary to note that the youngest age group in our study corresponded to the age period of 13-15 years (7th to 9th grades), an age when the system of decoding semantic information has long ago been automated.

Ideas in support of the cross-modal activation theory underlying the process of identification of SS words, have been repeatedly voiced (Ramachandran, 2001).

According to this theory, the search for correspondence between the sound and form of the word can be explained by the presence of sensory connections between the auditory and visual zones of the cortex (Kovic, 2010). EEG experiments on the perception of SS has proved that infants of 11 and 14 months-old process SS information faster than non-SS (Asano, 2015; Miyazaki, 2013). It has also been shown that synesthesia, arising in the process of perceiving SS words and associated with activation of cross-modal integration, contributes to better intuitive understanding of information (Revill, 2014; Bankieris, 2015).

At the same time, in experiments with adults using the method of Event Related Potential (ERP), a delay in cognitive processing of SS information was detected (Lockwood, 2016), as well as a late negative component in the composition of the ERP as an indicator of audio-visual integration (Molholm, 2002).

It is very likely that in the initial stages of ontogeny, SS information is an integral part of speech development. The analysis of a single case cannot directly confirm this assumption; however, in a longitudinal study, it was found that SS was the basis for a bootstrapping mechanism of an infant's speech, both at the lexical and phonological levels (Laing, 2014). In many studies SS is regarded as a stage of the early evolution of language as a linguistic system (Pleyer, 2017; Blasi, 2016). Therefore the analogy between ontogenesis and phylogenesis suggests we should see SS as an indispensable early stage in the formation of the speech system both at the individual and global levels.

The idea of a bi-directional, competing link, semantic and phonological, interfering with the processing of SS words, has already been expressed (Pexman, 2012), but has not been proved experimentally. It is interesting that, when the SS word is perceived by hearing, it is identified more accurately (Revill, 2014), but when it is presented visually, the probability of error is significantly increased. It turns out that the task of auditory identification of SS is much easier than the task of its visual identification. Apparently, with years of continuous training of the decoding system of semantic information using verbal-logical codes and left-hemisphere strategies (under dominant right-handedness), the latter becomes a priority and is activated at each meeting with semantic information. In this case, if the word contains not only a semantic but also a figurative message, the resources of one semantic system are not enough to decode the information properly. Thus, it is necessary to involve additional processing circuits, that can decipher a figurative sensory message which provokes cross-modal interaction

Conclusion

Our data show that, in comparison with non-SS words, visual perception of SS words of both native and foreign language, causes a delay in cognitive processing in adult subjects. This is probably because SS words carry information in two dimensions at the same time: semantic and sensual. For this reason, visual cognitive processing of such a stimulus is more complex and slower, and involves activation of at least two processing loops.

To further this research, we are planning to conduct a psychophysiological experiment with the help of EEG and ERP, registering the bioelectric activity of the

brain when it is in the process of decoding SS words, presented both visually and audibly. We hope to get information about the reorganization of the brain's systemic activity in processing SS words. In addition to contributing to the solution of the psychophysiological problem at the methodological level, this will allow us to approach an understanding of neurocognitive mechanisms of perception of SS.

Limitations

This study would have been more complete if it had been possible to collect information about the cognitive processing of visually presented words by primary school pupils who had only formed, but not yet automated, the decoding system for semantic information. It would also be useful to collect information on the contour of interhemispheric functional interaction, and to render this factor as a dependent variable when taking into account the quantitative indices of perception of SS.

References

- Akita, K. (2009). A grammar of sound-symbolic words in Japanese: theoretical approaches to iconic and lexical properties of Japanese mimetics. PhD dissertation, Kobe University. Retrieved from: <http://www.lib.kobe-u.ac.jp/repository/thesis/d1/D1004724.pdf>
- Ameka, F.K. (2001). Ideophones and the nature of the adjective word class in Ewe. In F.K. Erhard Voeltz and Ch. Kilian-Hatz (Eds.), *Ideophones* (pp. 25–48). Amsterdam: John Benjamins Publishing Company. <https://doi.org/10.1075/tsl.44.04ame>
- Armoskaite, S. & Koskinen, P. (2017). Structuring sensory imagery: Ideophones across languages and cultures/La structuration de l'imagerie sensorielle: les idéophones dans diverses langues et cultures. *Canadian Journal of Linguistics/Revue canadienne de linguistique*, 62, 2, 149–153. <https://doi.org/10.1017/cnj.2017.12>
- Asano, M., et al. (2015). Sound symbolism scaffolds language development in preverbal infants. *Cortex: Research report*, 63, 196–205. <https://doi.org/10.1016/j.cortex.2014.08.025>
- Bankieris, K. & Simner, J. (2015). What is the link between synaesthesia and sound symbolism? *Cognition*, 136, 186–195. <https://doi.org/10.1016/j.cognition.2014.11.013>
- Blasi, et al. (2016). Sound-meaning association biases evidenced across thousands of languages. *Proceedings of the National Academy of Sciences of the United States of America*, 113, 10818–10823. <https://doi.org/10.1073/pnas.1605782113>
- Clark, H.H. (2016). Depicting as a method of communication. *Psychological Review*, 123, 3, 324–347. <https://doi.org/10.1037/rev0000026>
- Cytowic, R.E. & Eagleman, D. (2009). *Wednesday is indigo blue: discovering the brain of synesthesia*. Cambridge, UK: The MIT Press.
- Dingemanse, M. (2012). Advances in the Cross-Linguistic Study of Ideophones. *Language and Linguistics Compass*, 6(10), 654–672. <https://doi.org/10.1002/lnc3.361>
- Dingemanse, M. (2018). Redrawing the margins of language: Lessons from research on ideophones. *Glossa: a journal of general linguistics*, 3(1), 4. <https://doi.org/10.5334/gjgl.444>
- Drellishak, S. (2006). *Statistical Techniques for Detecting and Validating Phonesthemes*. University of Washington. Seattle, WA.
- Fasmer, M. (1986). *Etimologichesky slovar russkogo yazyka* [Etymological dictionary of the Russian language]. Russia, Moscow: Progress.
- Imai, M. et al. (2008). Sound symbolism facilitates early verb learning. *Cognition*, 109, 1, 54–65. <https://doi.org/10.1016/j.cognition.2008.07.015>

- Imai, M. & Kita, S. (2014). The sound symbolism bootstrapping hypothesis for language acquisition and language evolution. *Philosophical Transactions of the Royal Society, B* 369: 20130298. <https://doi.org/10.1098/rstb.2013.0298>
- Kazuko, S. & Shigeto, K. (2010). A cross-linguistic study of sound symbolism: the images of size. *Berkeley Linguistics Society*, 396–410. <https://doi.org/10.3765/bls.v36i1.3926>
- Kovic V., Plunkett K., & Westermann, G. (2010). The shape of words in the brain. *Cognition*, 114, 1, 19–28. <https://doi.org/10.1016/j.cognition.2009.08.016>
- Laing, C.E. (2014). A phonological analysis of onomatopoeia in early word production. *First Language*, 34, 5, 387–405. <https://doi.org/10.1177/0142723714550110>
- Lockwood, G., Hagoort, P., & Dingemanse, M. (2016). How iconicity helps people learn new words: neural correlates and individual differences in sound-symbolic bootstrapping. *Collabra*, 2(1), 7, 1–15. <https://doi.org/10.1525/collabra.42>
- Magnus, M.A. *Dictionary of English Sound*. Retrieved from: <http://www.trismegistos.com>
- Meyer, D.E. & Schvaneveldt, R.W. (1971). Facilitation in recognizing pairs of words: Evidence of a dependence between retrieval operations. *Journal of Experimental Psychology*, 90, 227–234. <https://doi.org/10.1037/h0031564>
- Miroshnikov, S.A. (2010). *Ekspertnaya sistema Longitud. Experimentalno-dagnostichesky complex (EDC)* [Expert system longitudinal research data. Experimental-diagnostic complex (EDC)] Russia, Sankt-Petersburg: LEMA.
- Miyazaki, M., et al. (2013). The facilitatory role of sound symbolism in infant word learning. *Proceedings of the Annual Meeting of the Cognitive Science Society*, 35, 3080–3085. Retrieved from <https://escholarship.org/uc/item/5zt40388>.
- Molholm, S. et al. (2002). Multisensory auditory-visual interactions during early sensory processing in humans: a high-density electrical mapping study. *Brain Research: Cognitive Brain Research*, 14(1), 115–128. [https://doi.org/10.1016/S0926-6410\(02\)00066-6](https://doi.org/10.1016/S0926-6410(02)00066-6)
- Nasledov, A.D. (2013). IBM SPSS 20 i AMOS: professionalniy statisticheskiy analiz dannih. [IBM SPSS 20 and AMOS: professional statistical data analysis.]. Russia, Sankt-Petersburg: Piter.
- Osaka, N., et al. (2003). An emotion-based facial expression word activates the laughter module in the human brain: a functional magnetic resonance imaging study. *Neuroscience Letters*, 340(2), 127–130. [https://doi.org/10.1016/S0304-3940\(03\)00093-4](https://doi.org/10.1016/S0304-3940(03)00093-4)
- Osaka, N. (2006). Human anterior cingulate cortex and affective pain induced by mimic words: A functional magnetic resonance imaging study. *Psychoanalysis and neuroscience*, 257–268. https://doi.org/10.1007/88-470-0550-7_11
- Perniss, P. & Vigliocco, G. (2014). The bridge of iconicity: from a world of experience to the experience of language. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 369(1651). <https://doi.org/10.1098/rstb.2013.0300>
- PET Vocabulary List (2011). UCLES, Retrieved from: <http://www.cambridgeenglish.org/images/84669-pet-vocabulary-list.pdf>
- Pexman, P.M. (2012). Meaning-based influences on visual word recognition. In J.S. Adelman (Ed.), *Current issues in the psychology of language. Visual word recognition: Meaning and context, individuals and development* (pp. 24–43). New York, NY, US: Psychology Press.
- Pleyer, M., et al. (2017). Interaction and iconicity in the evolution of language: Introduction to the special issue. *Interaction Studies*, 18(3), 305–315. <https://doi.org/10.1075/is.18.3.01ple>
- Ramachandran, V.S. & Edward, M.H. (2001). Synaesthesia: a window into perception, thought and language. *Journal of Consciousness Studies*, 8(12), 3–34. Retrieved from: <https://philpapers.org/rec/RAMSA-5>
- Ratcliff, R., Gomez, P., & McKoon, G.A. (2004). Diffusion Model Account of the Lexical Decision Task. *Psychological Review*, 111(1), 159–182. <https://doi.org/10.1037/0033-295X.111.1.159>

- Revoll, K.P., et al. (2014). Cross-linguistic sound symbolism and crossmodal correspondence: Evidence from fMRI and DTI. *Brain and Language*, 128(1), 18–24. <https://doi.org/10.1016/j.bandl.2013.11.002>
- Rizzolatti, G. & Craighero, L. (2005). The mirror-neuron system. *Annual Review of Neuroscience*, 27(1), 169–192. <https://doi.org/10.1146/annurev.neuro.27.070203.144230>
- Sedelkina, Yu.G. (2016). Zapominaniye i usvoyeniye angliyskikh frazeologizmov v zabisimosti ot nalichiya v nikh fonosemanticheskogo komponenta [Memorizing and learning of English idioms depending on the presence of phonosemantic element] *Nauka i obrazovaniye segodnya*, 10(11), 65–66.
- Shimojo, S. & Shams, L. (2001). Sensory modalities are not separate modalities: plasticity and interactions. *Current Opinion in Neurobiology*, 11, 505–509. [https://doi.org/10.1016/S0959-4388\(00\)00241-5](https://doi.org/10.1016/S0959-4388(00)00241-5)
- Shlyachova, S.S. (2004). *Drebezgi yazika: Slovar russkikh fonosemanticheskikh anomalii* [Language clash: Russian Dictionary phonosemantic anomalies] Russia, Perm: publishing house of Perm state pedagogical University.
- Sidhu, D.M. & Pexman, P.M. (2017). Five mechanisms of sound symbolic association. *Psychonomic Bulletin & Review*, 1–25. <https://doi.org/10.3758/s13423-017-1361-1>
- Svantesson, J.O. (2017). Sound symbolism: The role of word sound in meaning. *Wiley Interdisciplinary Reviews: Cognitive Science*, 8:e1441, 1–12. <https://doi.org/10.1002/wcs.1441>
- Tkacheva, L.O., Gorbunov, I.A., & Nasledov, A.D. (2015). Reorganization of system brain activity while understanding visually presented texts with the increasing completeness of information. *Human Physiology*, 41(1), 11–21. <https://doi.org/10.1134/S0362119714060127>
- Van der Auwera, J. & Ceyhan, T. (2005). Semantic maps. In K. Brown (Ed.), *Encyclopedia of language & linguistics* (pp. 131–134). Oxford: Elsevier.
- von Stein, A., et al. (1999). Synchronization between temporal and parietal cortex during multimodal object processing in man. *Cereb. Cortex*, 9, 137–150. <https://doi.org/10.1093/cercor/9.2.137>
- Voronin, S.V. (1969). Angliyskie onomatopy (typi i stroenie). [English onomatopes (types and structure).] PhD dissertation. Soviet Union, Leningrad: Leningrad State University.
- Voronin, S.V. (2009). *Osnovi fonosemantiki* [The basics of phonosemantics] Russia, Moscow: Lenand.
- Walker, P., et al. (2010). Preverbal Infants' Sensitivity to Synaesthetic Cross-Modality Correspondences. *Psychological Science*, 21(1), 21–25. <https://doi.org/10.1177/0956797609354734>
- Westbury, C. (2018). Weighing up the evidence for sound symbolism: Distributional properties predict cue strength. *Journal of Memory and Language*, 99, 122–150. <https://doi.org/10.1016/j.jml.2017.09.006>

Original manuscript received September 28, 2018

Revised manuscript accepted January 10, 2019

First published online March 15, 2019

To cite this article: Tkacheva, L.O., Sedelkina, Y.G., Nasledov, A.D. (2019). Possible cognitive mechanisms for identifying visually-presented sound-symbolic words. *Psychology in Russia: State of the Art*, 12(1), 188–200. DOI: 10.11621/pir.2019.0114