Poljarnyj vestnik: Norwegian Journal of Slavic Studies, vol. 26, 2023, pp. 1–23
https://doi.org/10.7557/6.6880

Eye movements unveil that iconicity facilitates semantic recognition of onomatopoeic words*

Tuomo Häikiö and Oksana Kanerva

1. Introduction

The object of this study is a group of onomatopoeic words in Russian such as *bac* [bˈat͡s] 'bang', *pyx* [pˈɨx] 'puff', and *ščelk* [ɕːˈɵlk] 'click'. Traditionally, these linguistic units are simultaneously attributed to interjections and onomatopoeias (Karcevskij, 1984; Švedova, 1980 I, § 1700) or even to special deverbal formations derived from verbs (Nikitina, 2012). They have been labelled as "verbal-interjectional forms" that have onomatopoeic qualities (Mchitar′jan, 2016; Prokopovič, 1969; Xvostunova, 2000), "onomatopoeic interjections" that often function as predicates (Viimaranta & Vihervä, 2019), "verboids" that are also onomatopoeic (Nikitina, 2012) or "onomatopoeic verbal interjections" that combine all three features (Kanerva, 2018), mostly due to their unique capacity to capture auditory aspects of various events and actions. Despite the variability of views on how to label them, there is a consensus among scholars that Russian words such as *bac, pyx,* and *ščelk* are instances of form-meaning iconicity that are used for depiction of various sounds and indication of action causing these sounds (cf. Karcevskij, 1984; Švedova, 1980 I, § 1700; Wierzbicka, 2003). In this paper we will treat them as onomatopoeic forms as we believe that iconic sound depiction is their primary quality. Even though iconicity is a well acknowledged property of onomatopoeias, depiction of environmental sounds by means of language requires the use of phonemes and their combinations. These sequences of phonemes are often typical for the given language and shaped by a linguistic convention. Therefore, there are reasons to believe that generally iconic lexical forms such as onomatopoeias have conventional elements in themselves as well.

1.1 Iconicity between form and meaning and its connection to arbitrariness of language

Generally speaking, the symbolic nature of linguistic signs is the basic characteristic of all languages. Language users build conventional sequences of discrete symbols with essentially arbitrary connection between form and meaning in order to compose utterances. In linguistic theory, it has been accepted almost axiomatically that

* This research was supported by Ella and Georg Ehrnrooth Foundation, Kone Foundation, and Strategic Research Council.

arbitrariness is an inherent feature of human linguistic systems (Hockett, 1960; Saussure et al., 1960). Arbitrariness, as defined by Thompson et al. (2012, p. 1443), refers to the absence of a "meaningfully motivated link between the meaning and form of a word or sign". For example, there is nothing in the form of the English word *cow* that indicates that it is used to label a domestic animal. Thus, the meaning is assigned to this word by an agreement between language users and the connection between the form and meaning is arbitrary. In this respect, iconicity then refers to "the resemblance-based mapping between aspects of form and meaning" (Dingemanse et al., 2015, p. 604). For instance, English *moo* attempts to depict the sound produced by a cow with the help of its form, thus the way this word sounds resembles the sound that exists in the nature.

Although arbitrariness of languages has long been considered their fundamental feature, a growing body of research data suggests that iconicity is a wide-spread phenomenon, potentially universal (Voeltz & Kilian-Hatz, 2001). By offering scaffolding material for connecting linguistic forms and sensory experience, iconicity provides an insight into how languages establish a connection between words and extra-linguistic phenomena (Imai & Kita, 2014; Perniss & Vigliocco, 2014). Some studies have even shown that speakers of other languages are sensitive to the iconicity of real words in languages they have never learned before. For instance, Iwasaki et al. (2007) have found that English native speakers with no knowledge of Japanese and Japanese native speakers assigned similar meanings to Japanese ideophones used for depicting laughter.

Iconic correspondences between formal expression and meaning take various shapes and are revealed at different levels of the linguistic system. The most transparent form of iconicity is found in onomatopoeias. While onomatopoeic units resemble their referent fairly directly, there are other word classes that carry more subtle resemblances between form and meaning. In particular, ideophones (Estonian *tinga-tinga* 'to move taking short steps', *sussa-vussa* 'to move slowly') are known to map sounds and their combinations with different aspects of perceptual input or even temporal and spatial characteristics of events (Voeltz & Kilian-Hatz, 2001). Ideophones or artificially created words resembling ideophones (canonical nonce words *bauba/kiki* that maps round/jagged shapes) have been used to test how certain combinations of sounds evoke associations with different aspects of meaning coming from other modalities (Dingemanse, 2012; Revill et al., 2018). In particular, sounds/graphemes and their combinations have been linked with other types of sensual input such as taste (Pathak et al., 2020), colour (Hubbard et al., 2005), and shape (Maurer et al., 2006). In addition to that, even intonational patterns of utterances have been mapped with emotions (Cosmides, 1983) and attitudes (Bryant & Fox Tree, 2002).

The term embracing various instances of direct (iconic, one-on-one) linkage between sound and meaning is usually referred to as "sound symbolism" (Hinton et al., 1994). It has been defined based on "the proposal that linguistic sounds such as phonemes, features, syllables, or tones can be meaningful" (Nuckolls, 1999, p.225). Hinton, et al. (1994) suggest different types of sound symbolism. "Corporeal" sound symbolism comprises indications of emotional and physiological states of people such as interjections, involuntary bodily sounds such as sneezing or hiccups, and communicative exclamations. "Imitative" sound symbolism embraces onomatopoeic words such as English *buzz*, *bang*, and *click* that

imitate various acoustic events. "Synesthetic" sound symbolism refers to sounds, intonational and repetitive patterns imitating such non-linguistic phenomena as size, shape, colour, and taste. It is often observed in clusters of sounds, called phonesthemes. For instance, *gl-* in such English words as *glitter*, *glass,* or *glow* is often associated with light (Bergen, 2004). The most "conventional" or arbitrary type of sound symbolism has been found between phonemes and some aspects of meaning such as grammatical categories. For instance, cross-linguistic comparison between English, Dutch, French and Japanese showed that the presence of bilabial consonants *b, p,* and *m* could reliably predict that the word in question is a noun, while velar consonants *k* and *g* were good indicators of verbs (Monaghan et al., 2007). Having analysed a vast corpus of language data, Monaghan et al. (2014) admitted the presence of subtle sound symbolism, not always detectable by speakers, in the conventional lexicon as well. The study confirmed systematic regularities between measures of sound and measures of semantics on a large-scale sample of English monosyllabic words. In more detail, this research showed that phonologically similar words exhibited an above-chance tendency to appear in similar contexts that serves as an indication of their similarity in meaning. These results showed that form-meaning mapping went far beyond particular instances of imitative or synesthetic sound symbolism and confirmed the existence of conventional systematic sound symbolism (cf. Cuskley & Kirby 2013; Dingemanse et al., 2015).

To sum up, iconicity between form and meaning can be expressed to a different degree at different levels of language. Onomatopoeias are in the focus of this paper as they represent an almost perfect example of form-meaning iconicity and constitute cases of imitative sound symbolism. Indeed, experimental studies have shown that language users of English (Perry et al., 2015; Winter et al., 2017) and Spanish (Perry et al., 2015; Hinojosa et al., 2021) tend to rate the connection between form and meaning of onomatopoeic words as more iconic than words from other classes. Similar results were obtained for Japanese ideophones that were also rated as highly iconic by Japanese speakers (Thompson et al., 2020). These results indicate graded pervasiveness of iconicity in different layers of vocabulary. What is more, studies that looked into non-arbitrary form-meaning correspondences in basic vocabulary across thousands of languages found that languages tend to use or avoid the same speech sounds or their combinations to name the same objects or concepts hinting on more subtle universal regularities in form-meaning iconicity (Wichmann et al., 2010; Blasi et al., 2016). The question is whether words from an inherently iconic class such as onomatopoeias can have different degrees of iconicity between form and meaning.

We are aware that arbitrariness and iconicity in language collaborate to ensure communicative efficiency. In particular, a certain degree of arbitrariness increases semantic discrimination between phonologically similar units (Monaghan & Christiansen, 2006). Some onomatopoeic words indeed may sound similar (cf. Russian *bac, bam, bom,* and *bum* that all can be translated as 'bang'), but based on the agreement between language users that certain words are assigned to have a certain meaning, the difference even in one phoneme is sufficient to indicate that each of them is a distinct unit carrying a set of semantic features. Thus, arbitrariness in this particular case helps native speakers

successfully map each word with the shades of meaning they refer to. In the present study, we will test how the connection between iconicity and arbitrariness of Russian words such as *bac, pyx,* and *ščelk* is perceived and processed by those who do not know the language in question, in this case by Finnish L1 speakers with no knowledge of Russian.

1.2 Iconicity of Russian onomatopoeic words

A broader perspective on onomatopoeic words in a cross-linguistic sense suggests different degrees of conventionality for "wild" and "tame" onomatopoeias (Rhodes, 1994). In this connection, wild forms are intended to depict acoustic events as accurately as possible and often contain unusual combinations of sounds to serve this purpose best (*xrrr* that stands for cracking or *pšš* for steam let out). Tame forms in turn belong to more conventionalized items composed of sequences of phonemes typical for a particular language. This generalization casts light on the long history of attributing Russian words such as *bac, pyx,* and *ščelk* to semi-iconic entities (Karcevskij, 1984; Švedova, 1980 I, § 1700; Wierzbicka, 2003). Similar to wild onomatopoeias, words such as *bac, pyx,* and *ščelk* can exhibit fairly direct mapping between form and meaning, but employ more usual Russian combinations of phonemes for this purpose. In particular, there are units depicting very similar acoustic events related to hitting that differ by one phoneme only (e.g., *bax, bac, bam*; *bam, bom, bum*). All of them can be translated into English as 'bang'. In particular, *bax* refers to dull sounds with great force, *bac* to sharp and unexpected hits, *bum* to hits with moderate force, *bom* to hits against a bell, and *bam* to hits with moderate force, but also to a clock strike (Mchitar'jan, 2016). These semantic intricacies suggest that even naturally iconic linguistic units such as onomatopoeias contain conventionalized elements in order to be understood by language users. For Russian speakers this difference in one phoneme promotes differentiation between certain characteristics of the actual event of hitting. For a foreigner, however, these conventional variations might be inaccessible, whereas the semantics in relation to banging can be delivered by general iconic correspondence between form and meaning.

Additionally, among Russian onomatopoeic words there are ones that depict sounds produced by humans (e.g., *apčxi, čmok, am*). According to Hinton et al. (1994), involuntary utterances, including interjections and human bodily sounds, show a direct linkage between sound and meaning, i.e., they serve as symptoms of a physiological and emotional state of the speaker. These words are claimed to be highly recognizable due to the universal component they have. This assumption suggests their low conventionality and high iconicity. On the other hand, an experimental study by Perry et al. (2015) has established that interjections, being indications of emotional states, are perceived by native speakers of English and Spanish as less iconic in comparison with onomatopoeic words. These considerations indicate a need to investigate form-meaning mapping for different onomatopoeic words in Russian using experimental methods.

1.3 The present study

The central aim of this study is to determine how the difference in form-meaning iconicity affects semantic recognition of Russian onomatopoeic words by Finnish L1 speakers who marked in the questionnaire that they had never learned Russian before. This study sets out to find answers to the following research questions: 1) whether linguistic units from a

rather iconic class of words can exhibit form-meaning mapping to a greater or lesser extent; 2) whether variations in iconicity ratings are connected with how the semantics of stimulus words is recognised.

We believe that contemporary experimental methods may provide a valuable insight into the possible sensitivity of people to variations in iconicity and whether these have an effect on how the connection between form and meaning of words is established. We predict that different Russian onomatopoeic words have different levels of iconicity. We also hypothesise that the connection between more iconic units and their referents will be established more efficiently by Finnish L1 speakers with no knowledge of Russian as such words a priori should have a stronger link between form and meaning.

1.4 Materials

This study used 50 onomatopoeic words in Russian (*apčxi, bultyx, svis', ljap, šarax, bac, grox, bren', trax, čux-čux, tres', brjak, f'ju- f'ju, čeburax, din'-din', čmok, tuk-tuk, xlobys', xlop, pljux, skrip, bax, don, šmjak, čix, zvjak, ux-ux, šlep, bul'-bul', bux, tik-tak, xlest', om-om, kap, cvirk, bom, xrust', ščelk, cok-cok, porx, xljup, vžik, pyx, am, švark, tu-tu, čirk, top-top, pif-paf, šu-šu*), the same as in Kanerva and Häikiö (2022).

2. Iconicity ratings

The idea we consider to be central for our analysis is that different onomatopoeic words exhibit different degrees of iconicity. The iconicity rating test was used to establish perceived iconicity degrees for each token used in this study.

2.1 Participants

In total 28 participants took part in the iconicity rating test (2 males, 25 females, 1 other; mean age 51.07, SD 15.64), all adult native speakers of Finnish who marked in the questionnaire that they had never learned Russian before. The participants were recruited from a Facebook group. Informed consent was acquired from all participants prior to testing. The participants were tested in early spring of 2021.

2.2 Stimuli and Procedure

The iconicity rating test was designed as an online questionnaire using the QuestionPro.com platform. This tool allows creation of surveys that include both audio recordings and various types of response options. Stimulus material comprised 50 onomatopoeic words in Russian selected as material for this study (see section Material). The participants were asked to listen to the stimuli pronounced one by one by a native speaker of Russian. The respondents had to rate each onomatopoeic word on a scale from -5 to 5, where -5 indicated that the connection between form and meaning was perceived as assigned by convention and 5 indicated that it was perceived as iconic. The response anchors were numbered from -5 to 5 and the extreme points were labelled in Finnish as *sovinnainen* ‘conventional’ and *ikoninen* ‘iconic’. The scale was represented by function buttons for marking the selection from -5 to 5 about whether the connection between form and meaning for each token was perceived as conventional or iconic. Iconicity was defined to participants in terms of direct and straight forward resemblance between form and

meaning, and conventionality in terms of a shared agreement about the meaning of particular words between the language users.

Before the test, the participants were given instructions to listen to Russian words imitating natural sounds, to read the dictionary entry offered in Finnish describing the meaning of these words, and to rate each word on the iconicity scale. As for the possible overlap between the test items in Russian and Finnish a Finnish native speaker unfamiliar with Russian who did not take part in any of the experiments indicated that very few tokens, to their mind, resembled their counterparts (namely, Russian *apčxi* and Finnish *atsii* – 'sneezing', Russian *čux-čux* and Finnish *tsuku-tsuku* – 'train', Russian *f'ju-f'ju* and Finnish *viuh* – 'flying', and Russian *tik-tak* and Finnish *tik-tak* – 'clock'). The semantics of each token used in this study is presented in the Appendix. The participants were also provided with examples of similar linguistic units in Finnish exhibiting a different degree of iconicity. As examples, we used Finnish *kukkokiekuu*, which imitates sounds made by a rooster as an animal cry and is usually perceived as iconic, Finnish *naputinap*, which imitates keyboard typing and is more conventional, and Finnish *räiskis*, which depicts the sound of cracking and would be somewhere in the middle of this scale as it refers to sounds produced by something breaking but is not a unique sound and language specific such as typing, for instance. We did not want to prime the participants with giving human bodily sounds as examples of iconic entities, and as a result we offered another animate sound instead.

2.3 Results

Iconicity scores rated by the participants from -5 to 5 had a mean of 0.61 (SD = 2.35, Min = -3.43, Max = 4.68). The total amount of missing data was 1.6%. These cases comprised instances when the participants did not mark the answer. The iconicity rating for each token is given in the Appendix.

Our results show that iconicity between phonetic expression of words and their meaning varies within one and the same class of onomatopoeic words. Even though onomatopoeic entities are generally considered to be almost perfect examples of highly iconic linguistic units, some of them exhibit less motivated mapping between form and meaning. Thus, such words have a noticeably weaker connection with their referents as compared to those rated as more iconic.

3. Eye tracking: the visual world paradigm

In this study, the eye-tracking method was used to test whether Finnish L1 speakers with no knowledge of Russian could grasp the meaning of onomatopoeic words by hearing them and whether the degrees of iconicity between form and meaning of each word had any effect on their semantic accessibility. Our prediction was that the semantics of words rated as more iconic in the iconicity rating test would be recognized faster in comparison with less iconic items, as indexed by gaze directed to the corresponding picture faster for high-iconicity than low-iconicity words. The "eye-mind hypothesis" (Just & Carpenter, 1984) lies at the very core of eye tracking. According to this hypothesis, people direct their gaze at the objects they are processing at the moment, i.e., people look at what they are thinking of. In particular, this research employed the visual world paradigm. Huettig et al. (2011) present a detailed account on the application of this method in studying language

processing. In a typical setup, participants are presented with audio stimuli while looking at the experimental display containing the target object that is either described or mentioned in the audio stimuli. An experimental display usually contains a 2 x 2 grid of pictures. Either before or after the spoken expression is played, the participants have some time to view the pictures. Their eye movements are recorded for further analysis in order to gain an insight into the cognitive mechanisms involved in understanding the spoken expressions.

For example, in Allopenna et al. (1998) spoken word recognition was tested with the visual word paradigm. After being shown an experimental display the participants were presented with a spoken utterance "Pick up a candle" naming the stimulus word (candle). The experimental display contained the target image (candle), a phonological competitor that had a cohort overlap (candy), a phonological competitor that had a rhyme overlap (handle), and an irrelevant distractor (dollar). The results showed that lexical activation triggered the direction of the gaze to the target image depicting a candle. Additionally, it indicated that phonological similarity between the spoken word and the two competitors (candy and handle) also attracted the looks. In particular, when the word "candle" started to play the likelihood of fixating the candle and candy was higher, but towards the end of the word "candle", the likelihood of fixating the handle started to increase. Furthermore, semantic similarity between the target image and the item used as a competitor drives fixations (Cooper, 1974; Huettig & Altmann, 2005). For instance, in a classic study (Cooper, 1974), when simultaneously presented with a spoken word (Africa) and a visual item (zebra) sharing a semantic connection with it, the participants looked at that item. Finally, it has been found that visual attention is attracted to objects that visually resemble the target image, but are not related to it semantically (Huettig & Altmann, 2007). For instance, when hearing the target word (snake), the participants directed their gaze at a visually competing image (cable) that was similar in shape to the target.

In general terms, these studies show that different pieces of information, e.g., phonological, semantic, or visual, that are available about the target activate the fixations. In our design we used semantic and visual competitors to the target. The eye tracking data for the prospective analysis is the same as in Kanerva and Häikiö (2022) that also used eye tracking as method. However, said study did not study iconicity in relation to eye movements.

3.1 Participants

In total 27 native speakers of Finnish (26 female and 1 male, mean age 23.83 years, SD 4.83), who had never learned Russian before, took part in this experiment. None of them participated in the iconicity rating test. All of the participants had normal or corrected-to-normal eyesight and had no hearing deficits. Informed consent was acquired from all participants prior to testing. They received course credits for the participation. They were tested in late fall of 2020.

3.2 Stimuli and Procedure

We used the visual world paradigm for our eye-tracking experiment. It contained a direct action task, the equivalent instruction in English to which would be "look at the source of

the sound depicted by the word you hear". On each trial, the respondents heard an onomatopoeic word in Russian while looking at the visual display that consisted of four pictures (cf. Huettig & McQueen, 2007). The stimuli were the same as in the iconicity rating test of this study, pronounced by a native speaker of Russian and recorded. The experimental display contained four pictures presented on the screen (Figure 1). The pictures were drawings taken from the Papunet.net picture bank (https://kuvatyokalu.papunet.net). One of them was the target image that illustrated the same acoustic event depicted by the stimulus word in Russian. Another picture was a visual distractor that resembled the target image visually but contained an object or phenomenon completely semantically unrelated to the stimulus word. The third picture was a semantic distractor. The image in it was related semantically to the target image, but not to the acoustic event itself. The object in it could not produce a similar sound or perform similar action connected to that sound. Therefore, a semantic distractor could not be confused with the target. The fourth picture contained an irrelevant object to the target image. We selected objects that would be naturally rather static, silent, and unattractive for the participants, e.g., there were very few foods, in case someone comes hungry for the experiment; few animals, because they may have been preferred over things and few luxury or positional goods which people might want to have. A total of 50 picture quadruplets were constructed. The pictures were edited so that the color schemes within the quadruplets were similar. Some pictures were edited slightly to remove any texts or numbers, and to make the visually similar pictures as visually similar to the targets as possible. Each picture was sized to fit a 400*400 pixel box. None of the pictures were used twice.

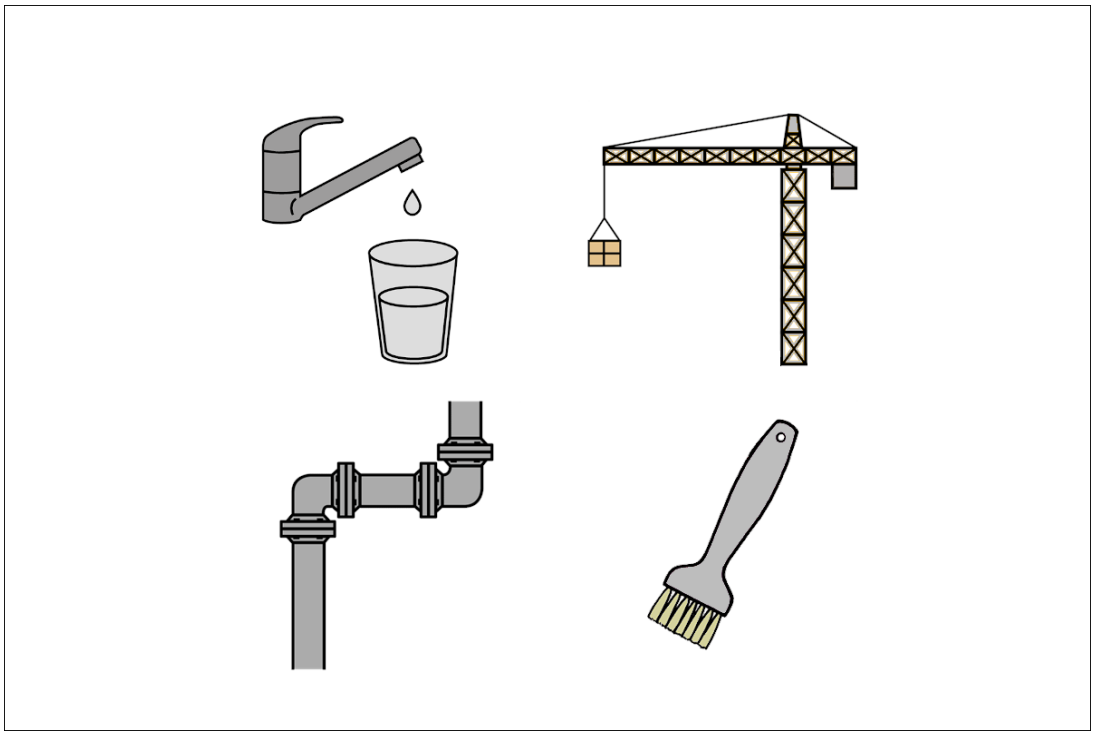

Figure 1. Experimental display for the target word *kap* 'dripping' (top left). The visual distractor is on top right, semantic distractor on bottom left, and irrelevant control picture on bottom right. Pictures are taken from the Papunet picture bank, papunet.net, Kuvako (top left), Sergio Palao/ARASAAC (top right and bottom left), Elina Vanninen (bottom right); all are edited versions of the original pictures. Pictures are shared with the Creative Commons BY-NC-SA licence

Before running the experiment, we needed to confirm that semantic distractors were related to the target pictures, but visual distractors and irrelevant pictures were not perceived as semantically similar with the target images. For this purpose, we asked 30 participants (19 female, mean age 43.60 years, SD 14.36) to rate how semantically similar each distractor and each irrelevant picture was to the target. The survey was conducted online using Webropol. It included 150 picture pairs in which each target image was paired with the corresponding semantic and visual distractors and the irrelevant image. Semantic similarity between images in each picture pair was rated on a 7-point Likert scale ranging from no connection at all to strong connection. This test allowed us to make sure that all semantic distractors had a mean rating above 4, and all visual distractors and irrelevant images had a mean rating below 4. The pictures that did not score within the designated margin were substituted with more suitable ones and the rating for those picture pairs was run again. The new picture pairs were rated by 49 respondents (37 female, mean age 39.10 years, SD 8.62) who did not participate in the first rating. One picture had to be substituted once more. This picture pair was rated by 9 respondents (4 female, mean age 41.56 years, SD 12.05). In the end, the mean rating for each semantic distractor was above 4 (mean 5.76, SD 0.71), and for visual distractors and irrelevant images below 4 (mean 2.38, SD 0.66, and mean 1.91, SD 0.52, respectively).

Furthermore, the visual similarity of the distractors in relation to the target picture was assessed by another rating task. For this, another group of participants (23 participants, 17 females, mean age 46.17 years, SD 17.79) was recruited to rate the visual similarity of each distractor in comparison to the target. The survey was conducted online using Webropol. It included 150 picture pairs in which each target image was paired with the corresponding semantic and visual distractors and the irrelevant image. Visual similarity between images in each picture pair was rated on a 7-point Likert scale ranging from no similarity at all to strong similarity. The mean rating of visual similarity to the target picture was 3.17 (SD 0.74), 2.09 (SD 0.52), and 1.56 (SD 0.36) for visual, semantic, and irrelevant distractors, respectively.

The eye-tracking experiment took place after an unrelated lexical decision task. The participants, who did not take part in any of the previous experiments or ratings of this study, were seated at the EyeLink 1000 tracker attached to a computer. A 1000 Hz sampling rate was used and the eye movements of the right eye were tracked. The stimuli were presented on a 24-inch BenQ XL2411 monitor with a resolution of 1920*1080 and refresh rate of 144 Hz. The participants sat approx. 60 cm from the screen and placed their head on a chin-and-forehead rest. They were instructed that they would hear Russian words and would need to look at the picture that could produce the sound they heard. A nine-point calibration procedure was conducted prior to the experiment. All but one participant had an average error below 0.50 degrees. The remaining participant had an average error of 0.85 degrees. Before each trial, the participant had to fixate on a circle in the center of the screen. After a stable fixation, the four pictures appeared in a 2x2 grid. After 2000 ms, the sound was played. After the sound onset, the pictures stayed on screen for 3500 ms.

3.3 Results

The eye tracking data was analyzed with the R statistical software (version 4.2.1, R Core Team, 2022). The data was preprocessed with the VWPre package (version 1.2.3, Porretta et al., 2016). For 2.21% of the data, the eye fixations were either off screen or not detected. However, every trial (a total of 1350) had more than 75% data intact that were thus included in the analysis. The proportions of the eye position samples falling within and outside the four interest areas were calculated using 40 ms bins separately for each interest area. The resulting proportions were converted to empirical logits. The analysis period started at 200 ms from the word onset as this is the time it takes to program a saccade (Fischer, 1992) and lasted until the end of the trial. The resulting time series were analyzed with generalized additive mixed modeling using the mgcv package (version 1.8-40, Wood, 2011). The visualizations have been done with the itsadug package (version 2.4, van Rij et al., 2017).

In order to test the effect of iconicity over time, the interaction between time and iconicity was added to the model as a surface using a tensor product. Moreover, simple random intercepts of the event (combination of subject and item) were included in the model in order to allow a unique intercept for each time series. Factor smooths of subjects and items over time were also included in the model. Finally, for the weights in the model, the inverse of the empirical logit variance estimates was used. From the first model, $\rho = 0.285$ was extracted in order to compensate for the autocorrelation. The model was refitted with the ρ included. There were no problems with the residuals of the model. The final model explained 50.5% of the deviance. The final model is presented in Table 1.

Parametric coefficient	Estimate	SE	t value	p value
Intercept	0.8467	0.1289	6.568	< 0.001
Smooth terms	**Edf**	**Ref. df**	**F value**	**p value**
Tensor product: Time by Iconicity	10.13	10.39	20.37	< .0001
Random effect: Time, subject	189.42	242.000	229.19	< .0001
Random effect: Time, item	373.01	448.000	248.50	< .0001
Random effect: Event	1260.80	1348.000	37.83	< .0001

Note. Edf = effective degrees of freedom, Ref. df = reference degrees of freedom.

Table 1. Generalized additive mixed model for the empirical logit of looks to the target picture as a function of time. The upper panel reports the parametric coefficient (intercept) and the lower panel the smooths and random effects

The proportions of looks to the target are plotted as a function of time in Figure 2. It can be seen from Figure 2 that as the trial proceeded, the likelihood of fixating the correct

image kept growing. Furthermore, the conditional tensor product over time showed that this effect was dependent on the iconicity score.

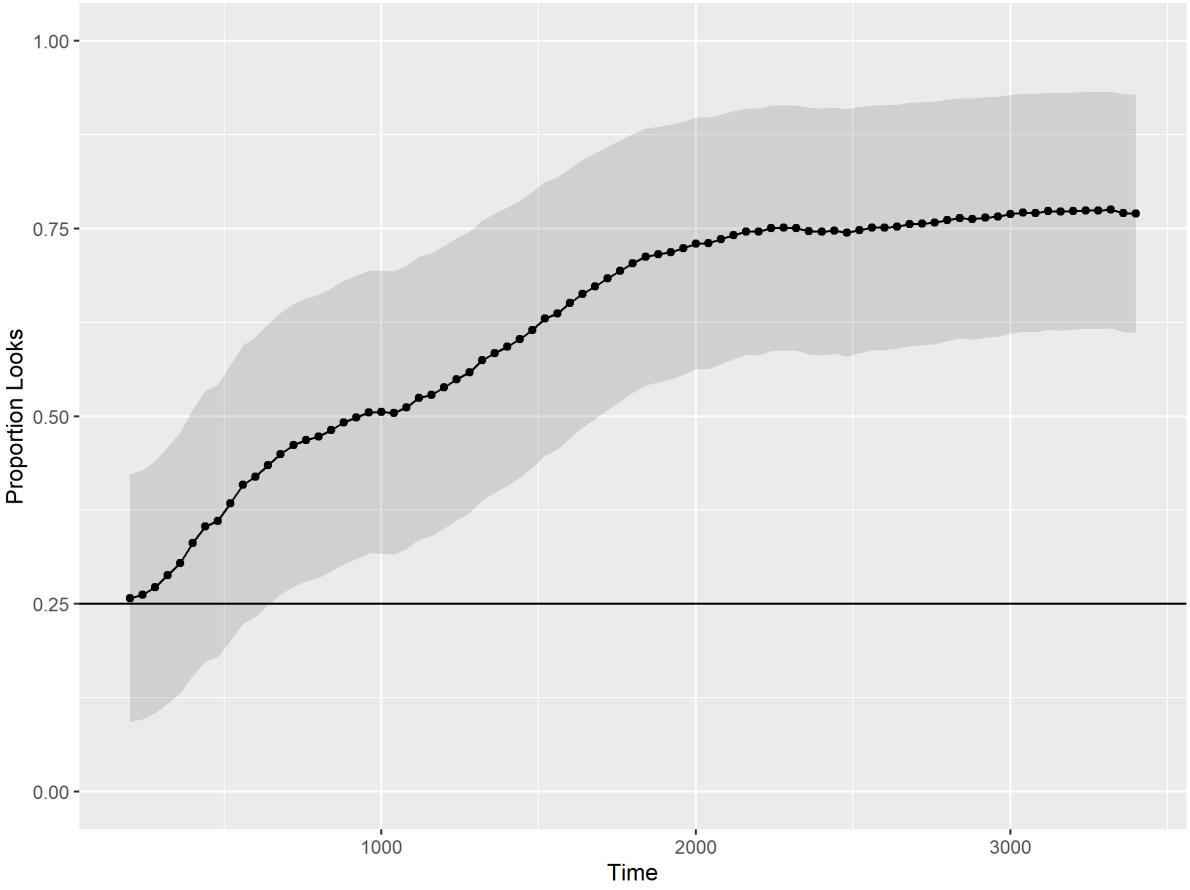

Figure 2. The fixation proportions to the target as a function of time. The shaded area denotes the 95% confidence interval. The horizontal line at 25% denotes the chance level as there were 4 images in each quadruplet

Next, the partial effects are visualized in Figure 3. This type of contour plot has been adopted in recent years for visualizing time series data analyzed with generalized additive mixed models, such as those used in visual world paradigm (e.g., Porretta & Kyröläinen, 2019; Porretta et al., 2020). It is included in the r package itsadug developed for evaluating and interpreting GAMM models (van Rij et al., 2017). Using a contour plot, three continuous variables can be plotted in the same two-dimensional graph. In Figure 3, we have Time on X axis, Iconicity on Y axis and the t value for any given point in said two-dimensional space colored with different shades of color. The t value refers to the likelihood of fixation to the target in relation to the average of fixating the target over the whole trial. The areas with red/orange hues (medium grey in print) are associated with fewer looks than average over the whole trial (i.e., negative t values), and the areas with yellow hues (light grey in print) with more (i.e., positive t values). The statistically non-significant regions are masked in the figure (i.e., appear in darkest colors).

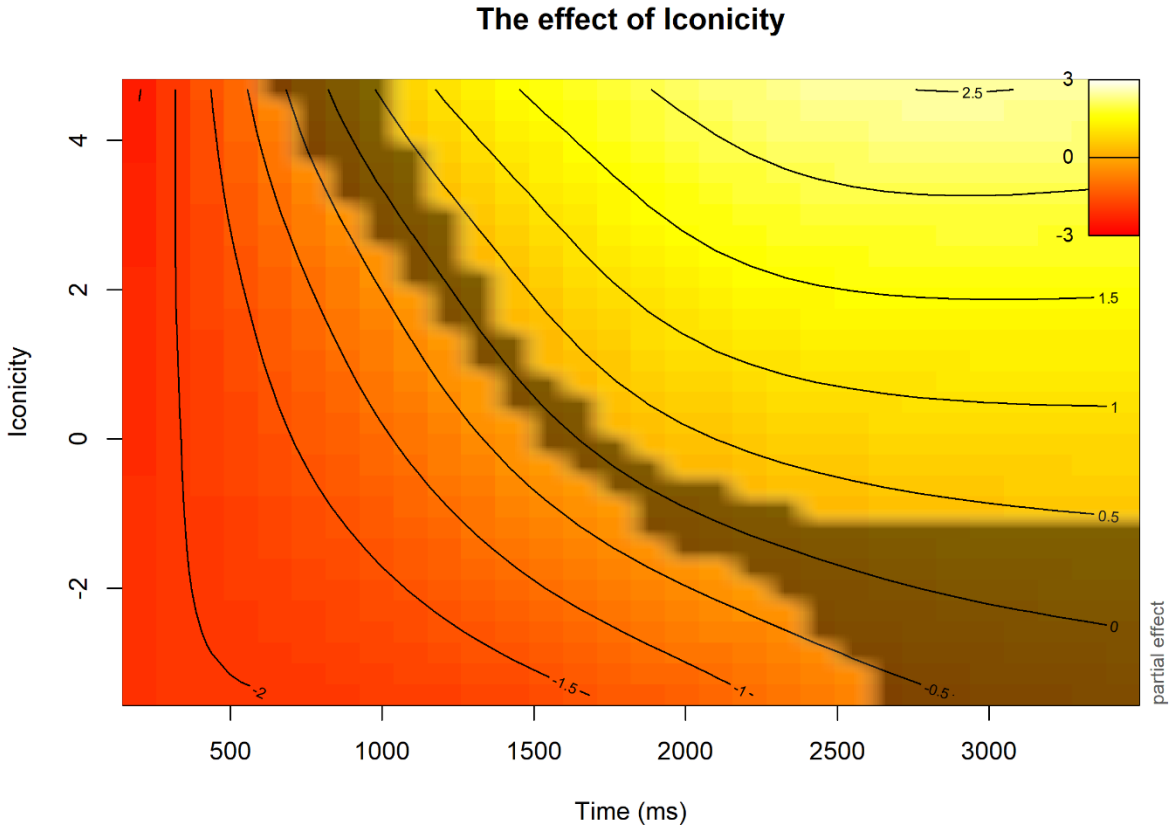

Figure 3. Contour plot of Time by Iconicity on the looks on the target picture. The masked regions indicate area that includes zero within the 95% CI

Interestingly, Figure 3 contains topographic contour lines. Areas around each contour line have attracted the same proportion of fixations. Negative values refer to fewer looks than average over the whole trial, and positive values refer to more looks. For instance, the line associated with the value -2 is at the beginning of the trial and areas close to this line have attracted fewer looks than on average over the whole trial. Furthermore, the line is vertical, which means that the proportion of fixations at that given time point is similar across the whole range of iconicity scores. Next, the line associated with the value 0 crosses the graph. As the value is 0, around this area the proportion of fixations is at average. Moreover, around this line, the items with the highest iconicity score attracted as large proportion of fixations around 800 ms after sound onset as the items with the lowest iconicity scores did only after 2700 ms. Finally, lines associated with positive values, i.e., attracting more looks than average, do not reach the bottom of the graph. This means that items with higher iconicity scores attracted high likelihood of looks already 1000 ms after sound onset unlike those with the lowest iconicity scores that did not do so even at the end of the trial. To sum up, higher iconicity scores were related to a higher likelihood of fixations to the target than lower iconicity scores. Furthermore, for items with higher iconicity scores, the target pictures were likely to be fixated earlier than those with lower iconicity scores.

In addition, we ran separate analyses for the visual and semantic competitors by modeling the difference in looks (on empirical logit scale) between the competitor and distractor. This allowed us to quantify how much visual and semantic competition was

induced relative to a baseline, i.e., distractor in the case of our study. For both types of competitors, the predictors were the same as for the Target model with the difference that the weights could not be included in the models as the difference scores of empirical logits were used. From the first models, ρ was extracted in order to compensate for the autocorrelation (ρ = 0.283 and ρ = 0.297 for visual and semantic competitors, respectively). The interaction between time and Iconicity was not significant for either model and was dropped from the final models. The final models are presented in Tables 2 and 3 for visual and semantic competitors, respectively. Crucially, the main effect of Iconicity was not significant for either model.

Parametric coefficient	Estimate	SE	t value	p value
Intercept	0.23353	0.06232	3.748	< 0.001
Smooth terms	**Edf**	**Ref. df**	**F value**	**p value**
Smooth: Iconicity	1.001	1.001	1.272	.259
Random effect: Time, subject	175.994	242.000	6.270	< .0001
Random effect: Time, item	360.649	448.000	22.733	< .0001
Random effect: Event	1266.009	1348.000	22.444	< .0001

Note. Edf = effective degrees of freedom, Ref. df = reference degrees of freedom.

Table 2. Generalized additive mixed model for the difference between the empirical logits of looks to the visual distractor versus baseline as a function of time. The upper panel reports the parametric coefficient (intercept) and the lower panel the smooths and random effects

Parametric coefficient	Estimate	SE	t value	p value
Intercept	0.2335	0.0623	3.747	< 0.001
Smooth terms	**Edf**	**Ref. df**	**F value**	**p value**
Smooth: Iconicity	1.001	1.001	1.275	.259
Random effect: Time, subject	176.210	242.000	6.369	< .0001
Random effect: Time, item	360.697	448.000	22.024	< .0001
Random effect: Event	1265.865	1348.000	22.462	< .0001

Note. Edf = effective degrees of freedom, Ref. df = reference degrees of freedom.

Table 3. Generalized additive mixed model for the difference between the empirical logits of looks to the semantic distractor versus baseline as a function of time. The upper panel reports the parametric coefficient (intercept) and the lower panel the smooths and random effects

4. Conclusions

In this research, we tested 50 Russian onomatopoeic words that imitate inanimate environmental sounds and human bodily sounds, and can also depict an action connected to these acoustic events. The primary objective of this study was to check whether there is a variation in form-meaning iconicity across linguistic units by default regarded as iconic. Findings of this research suggest that "absolute iconicity" that "involves a fairly straightforward one-to-one resemblance between aspects of form and meaning" (Dingemanse et al., 2015, p. 606), generally considered typical of all onomatopoeias, nevertheless varies among tested tokens from the perspective of Finnish L1 speakers with no knowledge of Russian. From the theoretical perspective, onomatopoeias tend to demonstrate a fairly direct mapping between phonological features of sounds that constitute them and the acoustic features of their referents (Dingemanse et al., 2015; Rhodes, 1994). At the same time, the phonological expression of onomatopoeic words that have entered dictionaries resembles original sounds by convention (Švedova, 1980; Šaronov, 2008). Therefore, words imitating sounds produced by objects and human bodily sounds are likely to contain both language specific and universal elements (cf. Wierzbicka, 2003). This suggested a possible variation in iconicity levels within the group of Russian onomatopoeic words. Our results are in line with this assumption as some of the tokens were perceived as highly iconic by speakers of another language with no familiarity with Russian, while others were rated as rather conventional. There is one interesting observation emerging from the qualitative analysis of the data. Inspection of the mean iconicity ratings for different words made it evident that human bodily sounds such as *čmok, am, top-top, šu-šu, apčxi,* and *čix* on average scored high in iconicity. This detail offers new evidence to the claim that involuntary human vocalisations should have stronger linkage between from and meaning (cf. Hinton et al., 1994). The reason for

human bodily sounds to be marked as more iconic may be explained by the view that they serve as symptoms of physical and emotional states that are universal, because of their instinctive or involuntary nature that makes them common for all humans (Goffman, 1978). Some other expressions of emotional reactions or attitudes, for example English *yuk*, German *pfui*, Dakota *xox*, and Russian *fuu* have a language-specific form, but serve as universal gestures of revulsion (Haiman, 1998). Another example would be the word *Huh?* indicating grunting that is universal because it has been shaped by adaptive context of interactional environment, conditions of which are similar across cultures and are directed at initiating repair in conversation (Dingemanse et al., 2013).

Another intriguing tendency concerned *bac, bax, bux* that were rated as somewhere in the middle of the scale, while *din', don, bom* were regarded as very iconic. Even though the words within both groups are phonologically quite similar to one another, *bac, bax* and *bux* are semantically broader than *din', don* and *bom* that mostly refer to sounds made by bells of different sizes. This observation supports evidence from the study by Lupyan and Winter (2018) that found that iconicity is inversely related to contextual diversity. At this point, this is one example of an interesting trend, but more research on a broader range of items is needed to test this pattern.

Another objective of this study was to determine how the difference in form-meaning iconicity affects semantic recognition of Russian onomatopoeic words by Finnish L1 speakers who had never learned Russian before. In this study, the use of the visual world paradigm allowed closer examination of processes underpinning semantic recognition of unknown onomatopoeic words based on how these words sound. Participants' eye movements demonstrated a robust processing advantage in mapping onomatopoeic words to their referents offered by form-meaning iconicity. These results also align with the outcomes of another experimental eye-tracking study by Laing (2017), in which infants recognized onomatopoeias faster than the corresponding words from conventionalized vocabulary. Furthermore, Iconicity was not related with the fixation likelihoods to visual and semantic competitors, highlighting the mapping between the sound and the referent being correctly established. The most important outcome emerging from this research is that iconicity promotes semantic recognition of words. It confirmed the assumption that people with reported no prior knowledge of the languages can map the form and meaning of more iconic entities faster. These findings offer support to the hypothesis that more iconic entities are more accurate in depicting sounds they stand for. Admittedly, the generalizability of these results has certain limitations as we only tested one language on native speakers of another language. Thus, further research on universality of form-meaning correspondences in onomatopoeias would offer more evidence of how iconicity works.

References

Allopenna, P. D., Magnuson, J. S., & Tanenhaus, M. K. (1998). "Tracking the Time Course of Spoken Word Recognition Using Eye Movements: Evidence for Continuous Mapping Models." *Journal of Memory and Language*, 38(4), 419−439. https://doi.org/10.1006/jmla.1997.2558

Bergen, B. K. (2004). "The Psychological Reality of Phonaesthemes." *Language*

(Baltimore), 80(2), 290–311. https://doi.org/10.1353/lan.2004.0056

Blasi, D. E., Wichmann, S., Hammarström, H., Stadler, P. F., & Christiansen, M. H. (2016). "Sound–meaning association biases evidenced across thousands of languages". *Proceedings of the National Academy of Sciences – PNAS*, 113(39), 10818–10823. https://doi.org/10.1073/pnas.1605782113

Bryant, G. A., & Fox Tree, J. E. (2002). "Recognizing Verbal Irony in Spontaneous Speech." *Metaphor and Symbol*, 17(2), 99–117. https://doi.org/10.1207/S15327868MS1702_2

Cooper, R. M. (1974). "The control of eye fixation by the meaning of spoken language: A new methodology for the real-time investigation of speech perception, memory, and language processing." *Cognitive Psychology*, 6, 84−107. https://doi.org/10.1016/0010-0285(74)90005-X

Cosmides, L. (1983). "Invariances in the Acoustic Expression of Emotion During Speech." *Journal of Experimental Psychology*, 9, 864–881. https://doi.org/10.1037/0096-1523.9.6.864

Cuskley, C., & Kirby, S. (2013). "Synesthesia, Cross-modality, and Language Evolution." In J. Simner and E. M. Hubbard (Eds.), Oxford Library of Psychology. *The Oxford Handbook of Synaesthesia*, (pp. 869–899). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199603329.013.0043

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., Blasi, D. E., Lupyan, G., Christiansen, M. H., & Monaghan, P. (2015). "Arbitrariness, Iconicity, and Systematicity in Language." *Trends in Cognitive Sciences*, 19(10), 603–615. https://doi.org/10.1016/j.tics.2015.07.013

Dingemanse, M., Torreira, F., & Enfield, N. J. (2013). "Is "Huh?" a universal word? Conversational infrastructure and the convergent evolution of linguistic items". PloS One, 8(11), e78273–e78273. https://doi.org/10.1371/journal.pone.0094620

Fischer, B. (1992). "Saccadic Reaction Time: Implications for Reading, Dyslexia, and Visual Cognition." In K. Rayner (Ed.), *Eye movements and visual cognition*, (pp. 31–45). New York, NY: Springer. https://doi.org/10.1007/978-1-4612-2852-3_3

Goffman, E. (1978). "Response Cries". *Language*, 54, 787–815. https://doi.org/10.2307/413235

Haiman, J. (1998). *Talk is Cheap: Sarcasm, Alienation, and the Evolution of Language.* Oxford University Press, Incorporated.

Hinojosa, J. A., Haro, J., Magallares, S., Duñabeitia, J. A., & Ferré, P. (2021). "Iconicity ratings for 10,995 Spanish words and their relationship with psycholinguistic variables". *Behavior Research Methods*, 53(3), 1262–1275. https://doi.org/10.3758/s13428-020-01496-z

Hinton, L., Nichols, J., & Ohala, J. J., eds. (1994). *Sound Symbolism*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511751806

Hockett, C. F. (1960). "The Origin of Speech." Scientific American, 203(3), 88–96. https://doi.org/10.1038/scientificamerican0960-88

Hubbard, E. M., Arman, A. C., Ramachandran, V. S., & Boynton, G. M. (2005). "Individual Differences Among Grapheme-color Synaesthetes: Brain-behavior

Correlations." *Neuron*, 45, 975–985. https://doi.org/10.1016/j.neuron.2005.02.008

Huettig, F., & Altmann, G. T. M. (2005). "Word meaning and the control of eye fixation: Semantic competitor effects and the visual world paradigm." *Cognition*, 96(1), B23−B32. https://doi.org/10.1016/j.cognition.2004.10.003

Huettig, F., & Altmann, G. T. M. (2007). "Visual-shape competition during language-mediated attention is based on lexical input and not modulated by contextual appropriateness." *Visual Cognition*, 15(8), 985−1018. https://doi.org/10.1080/13506280601130875

Huettig, F., Rommers, J., & Meyer, A. S. (2011). "Using the visual world paradigm to study language processing: a review and critical evaluation." *Acta Psychologica*, 137 (2), 151–171. https://doi.org/10.1016/j.actpsy.2010.11.003

Huettig, F., & McQueen, J. M. (2007). "The tug of war between phonological, semantic, and shape information in language-mediated visual search." *Journal of Memory and Language*, 54, 460−482. https://doi.org/10.1016/j.jml.2007.02.001

Imai, M., & Kita, S. (2014). "The Sound Symbolism Bootstrapping Hypothesis for Language Acquisition and Language Evolution." *Philosophical Transactions. Biological Sciences*, 369(1651), 20130298–20130298. https://doi.org/10.1098/rstb.2013.0298

Iwasaki, N., Vinson, D. P., & Vigliocco, G. (2007). "How does it Hurt, kiri-kiri or siku-siku? Japanese Mimetic Words of Pain Perceived by Japanese Speakers and English Speakers". In M. Minami (Ed.), *Applying Theory and Research to Learning Japanese as a Foreign Language*, (pp. 2–19). Cambridge Scholars Publisher.

Just, M.A., & Carpenter, P.A. (1984). "Using eye fixations to study reading comprehension". In D.E. Kieras, & M.A. Just (Eds.), New methods in reading comprehension research (pp. 151–182). Hillsdale, NJ: Erlbaum.

Kanerva, O. (2018). "Correlation Between Expressiveness and Syntactic Independence of Russian Onomatopoeic Verbal Interjections." *Poljarnyj Vestnik: Norwegian Journal of Slavic Studies*, 21, 15–30. https://doi.org/10.7557/6.4435

Kanerva, O., & Häikiö, T. (2022). "Sound symbolic potential of Russian onomatopoeias: Evidence from eye-tracking." *Journal of Experimental Psychology. Learning, Memory, and Cognition*, 48(3), 432–445. https://doi.org/10.1037/xlm0001114

Karcevskij, S. O. (1984). "Vvedenie v izučenie meždometij." *Voprosy jazykoznanija*, 6, 127-137.

Laing, C. E. (2017). "A Perceptual Advantage for Onomatopoeia in Early Word Learning: Evidence from Eye-Tracking." *Journal of Experimental Child Psychology*, 161, 32–45. https://doi.org/10.1016/j.jecp.2017.03.017

Lupyan, G., & Winter, B. (2018). "Language is more abstract than you think, or, why aren't languages more iconic?" *Philosophical Transactions of the Royal Society of London. Series B. Biological Sciences*, 373(1752), 20170137–20170137. https://doi.org/10.1098/rstb.2017.0137

Maurer, D., Pathman, T., & Mondloch, C. J. (2006). "The Shape of Boubas: Sound-shape Correspondences in Toddlers and Adults." *Developmental Science*, 9(3), 316–322. https://doi.org/10.1111/j.1467-7687.2006.00495.x

Mchitar'jan, G. S. (2016). *Grammatičeskij status i funkcionirovanie glagol'no-meždometnych form v sovremennom russkom jazyke*. Pjatigorsk.

Monaghan, P., & Christiansen, M. H. (2006). "Why Form-meaning Mappings Are Not Entirely Arbitrary in Language." In R. Sun (Ed.), *Proceedings of the 28th Annual Conference of the Cognitive Science Society*, (pp. 1838–1843).

Monaghan, P., Christiansen, M., & Chater, N. (2007). "The Phonological-distributional Coherence Hypothesis: Cross-linguistic Evidence in Language Acquisition." *Cognitive Psychology, 55(4), 259–305. https://doi.org/10.1016/j.cogpsych.2006.12.001*

Monaghan, P., Shillcock, R., Christiansen, M., & Kirby, S. (2014). "How Arbitrary Is Language?" Philosophical Transactions. Biological Sciences, 369(1651), 20130299–20130299. https://doi.org/10.1098/rstb.2013.0299

Nikitina, T. (2012). "Russian Verboids: A Case Study in Expressive Vocabulary". *Linguistics*, 50(2), 165–189. https://doi.org/10.1515/ling-2012-0007

Nuckolls, J. B. (1999). "The Case for Sound Symbolism." *Annual Review of Anthropology*, 28(1), 225–252. https://doi.org/10.1146/annurev.anthro.28.1.225

Pathak, A., Calvert, G., & Motoki, K. (2020). "Long Vowel Sounds Induce Expectations of Sweet Tastes." *Food Quality and Preference*, 86, 104033. https://doi.org/10.1016/j.foodqual.2020.104033

Perniss, P., & Vigliocco G. (2014). "The Bridge of Iconicity: From a World of Experience to the Experience of Language." *Philosophical Transactions. Biological Sciences*, 369 (1654), 20140179. https://doi.org/10.1098/rstb.2013.0300

Perry, L. K., Perlman, M., & Lupyan, G. (2015). "Iconicity in English and Spanish and Its Relation to Lexical Category and Age of Acquisition." *PloS One*, 10(9), e0137147–e0137147. https://doi.org/10.1371/journal.pone.0137147

Porretta, V., Buchanan, L. & Järvikivi, J. (2020). "When processing costs impact predictive processing: The case of foreign-accented speech and accent experience." *Attention, Perception, & Psychophysics*, 82, 1558–1565. https://doi.org/10.3758/s13414-019-01946-7

Porretta, V., & Kyröläinen, A.-J. (2019). "Influencing the time and space of lexical competition: The effect of gradient foreign accentedness." *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 45(10), 1832–1851. https://doi.org/10.1037/xlm0000674

Porretta, V., Kyröläinen, A., van Rij, J., & Järvikivi, J. (2016). *VWPre: Tools for Preprocessing Visual World Data*. Version 1.2.3, updated 2020-03-08, URL https://CRAN.R-project.org/package=VWPre

Prokopovič, E. N. (1969). *Stilistika častej reči (glagol'nye slovoformy).* Moscow: Prosvešenie.

R Core Team. (2022). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. Vienna, Austria. URL https://www.R-project.org/

Revill, K. P., Namy, L. L., & Nygaard, L. C. (2018). "Eye Movements Reveal Persistent Sensitivity to Sound Symbolism During Word Learning." *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 44(5), 680–698. https://doi.org/10.1037/xlm0000476

Rhodes, R. (1994). "Aural Images." In L. Hinton, J. Nichols, & J. Ohala (Eds.), *Sound Symbolism*, (pp. 276–292). Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511751806.019

Šaronov, I. A. (2008). *Meždometija v reči, tekste i slovare*. Moscow: Izdatel'skij centr Rossijskogo gosudarstvennogo gumanitarnogo universiteta.

Saussure, F. d., Bally, C., Baskin, W., Reidlinger, A., & Sechehaye, A. (1960). *Course in General Linguistics*. London: Peter Owen. https://doi.org/10.2307/538001

Švedova, N. J., ed. (1980). *Grammatika russkogo jazyka*. Akademija Nauk SSSR, Moscow.

Thompson, A. L., Akita, K., & Do, Y. (2020). "Iconicity ratings across the Japanese lexicon: A comparative study with English". *Linguistics Vanguard: Multimodal Online Journal*, 6(1). https://doi.org/10.1515/lingvan-2019-0088

Thompson, R. L., Vinson, D. P., Woll, B., & Vigliocco, G. (2012). "The Road to Language Learning Is Iconic: Evidence from British Sign Language." *Psychological Science*, 23(12), 1443–1448. https://doi.org/10.1177/0956797612459763

van Rij, J., Wieling, M., Baayen, R. H., & van Rijn, H. (2017). *itsadug: Interpreting Time Series and Autocorrelated Data Using GAMMs*. R package version 2.4. URL https://cran.r-project.org/package=itsadug

Viimaranta, J., & Vihervä, M. (2019). "The Use of Onomatopoeic Interjections in Predicate Function in Russian and Other Languages: A Perspective from the Corpus of Parallel Texts of the Russian National Corpus". *Scando-Slavica*, 65(2), 239–262. https://doi.org/10.1080/00806765.2019.1672092

Voeltz, F. K. E., & Christa Kilian-Hatz, C. eds. (2001). *Ideophones*. Amsterdam, Philadelphia: John Benjamins. https://doi.org/10.1075/tsl.44

Wichmann, S., Holman, E. W., & Brown, C. H. (2010). "Sound Symbolism in Basic Vocabulary". *Entropy* (Basel, Switzerland), 12(4), 844–858. https://doi.org/10.3390/e12040844

Wierzbicka, A. (2003). "Interjections across Cultures". In A. Wierzbicka (Ed.), *Cross-Cultural Pragmatics: The Semantics of Human Interaction*, (pp. 285–340). Mouton de Gruyter. https://doi.org/10.1515/9783110220964.285

Winter, B., Perlman, B., Perry, L. K. & Lupyan, G. (2017). "Which words are most iconic? Iconicity in English sensory words". *Interaction Studies*, 18(3), 430–451. https://doi.org/10.1075/is.18.3.07win

Wood, S. N. (2011). "Fast Stable Restricted Maximum Likelihood and Marginal Likelihood Estimation of Semiparametric Generalized Linear Models." *Journal of the Royal Statistical Society, Series B (Statistical Methodology)*, 73 (1), 3–36. https://doi.org/10.1111/j.1467-9868.2010.00749.x

Xvostunova, T. I. (2000). *Funkcionirovanie glagol'no-meždometnyx form v sovremennom russkom literaturnom jazyke*. Saransk.

Appendix

	Token	Transcription	Meaning	Mean iconicity rating
1.	*čeburax*	t͡ɕɪbʊrˈax	hitting, falling clumsily or with rolling	-3.43
2.	*tres′*	trʲˈesʲ	hitting, bursting or splitting into parts, for example, of something wooden	-2.81
3.	*ščelk*	ɕːˈɵlk	mechanical clicking, also cracking of a nutshell, snapping with fingers	-2.79
4.	*xlobys′*	xlɐbˈɨsʲ	hitting, falling, also a massive volume of liquid being poured	-2.48
5.	*bultyx*	bʉltˈɨx	an object falling into water	-2.43
6.	*porx*	pˈorx	flying away or wings flapping	-2.39
7.	*švark*	ʂvˈark	hitting or throwing	-2
8.	*cvirk*	t͡sʲvʲˈirk	water trickling, also a thin sheet of water pouring	-1.79
9.	*zvjak*	zvʲˈæk	metallic or glass objects hitting against one another	-1.68
10.	*šarax*	ʂɐrˈax	hitting, throwing, exploding	-1.64
11.	*skrip*	skrʲˈip	friction or compression	-1.29
12.	*trax*	trˈax	cracking or rumbling	-1.26
13.	*ux-ux*	ˈux-ˈux	hitting, falling, exploding or cannon fire	-1.18
14.	*pif-paf*	pʲˈif-pˈaf	shooting with a pistol or gun	-1.18
15.	*xlest′*	xlʲˈɵstʲ	whipping, wind slashing	-1.14
16.	*pljux*	plʲˈʉx	something liquid or muddy falling	-1.11
17.	*bac*	bˈat͡s	hitting or shooting	-0.96

	Token	Transcription	Meaning	Mean iconicity rating
18.	*grox*	gr'ox	falling with noise or hitting the ground	-0.89
19.	*čirk*	t͡ɕ'irk	striking or rubbing against something dry or wooden	-0.89
20.	*xlop*	xl'op	falling, hitting, bursting, shooting or throwing	-0.79
21.	*kap*	k'ap	water dripping	-0.57
22.	*brjak*	brʲ'æk	hitting metallic or glass objects against one another, falling or hitting something with solid objects	-0.54
23.	*bax*	b'ax	hitting, shooting or exploding	-0.43
24.	*xljup*	xlʲ'ʉp	splashing, a munching sound of something liquid or muddy	0.07
25.	*xrust'*	xr'ustʲ	something fragile cracking or breaking	0.29
26.	*bux*	b'ux	hitting, exploding or falling	0.33
27.	*om-om*	'om-'om	eating or gulping	0.89
28.	*cok-cok*	t͡s'ok-t͡s'ok	clatter of hoofs or metallic heel tips	0.96
29.	*vžik*	vʐ'ik	buzzing or flying past	1.04
30.	*ljap*	lʲ'æp	hitting, slapping, also something muddy or wet falling down	1.36
31.	*čmok*	t͡ɕm'ok	kissing	1.36
32.	*am*	'am	eating or devouring	1.57
33.	*šlep*	ʂlʲ'ɵp	hitting, also slapping with a hand or walking clumsily, in a lazy manner or on mud	1.79
34.	*pyx*	p'ix	steam being let out or something burning down completely	1.86
35.	*šmjak*	ʂmʲ'æk	hitting, falling with a flopping noise	1.89

	Token	Transcription	Meaning	Mean iconicity rating
36.	*tuk-tuk*	tˈuk-tˈuk	sharp hits or knocking, usually on wood	2
37.	*bren′*	brʲˈenʲ	string musical instrument, e.g., guitar	2.14
38.	*top-top*	tˈop-tˈop	stomping, walking	2.18
39.	*šu-šu*	ʂˈu-ʂˈu	whispering or gossiping	2.19
40.	*svis′*	svʲˈisʲ	whistling, also an object flying through the air or items rubbing against one another	2.5
41.	*f′ju- f′ju*	fjˈʉ-fjˈʉ	an object flying past, also whistling	3.25
42.	*din′-din′*	dʲˈinʲ-dʲˈinʲ	small bell	3.64
43.	*tu-tu*	tˈu-tˈu	train horn	3.96
44.	*bom*	bˈom	big bell ringing	4.04
45.	*čux-čux*	t͡ɕˈux-t͡ɕˈux	train	4.21
46.	*don*	dˈon	big bell ringing or metallic clutter	4.29
47.	*tik-tak*	tʲˈik-tˈak	clock	4.33
48.	*bul′-bul′*	bˈulʲ-bˈulʲ	water, bubbles	4.57
49.	*apčxi*	ɐpt͡ɕxʲˈi	sneezing	4.57
50.	*čix*	t͡ɕˈix	person sneezing	4.68

Note. The mean iconicity ratings for each word were calculated based on the responses of 28 participants in the iconicity rating test. They provided their answers on the Likert scale from -5 to 5, where -5 indicated that the connection between form and meaning was perceived as assigned by convention and 5 indicated that it was perceived as iconic.

Table A1. Target words, their meanings, and mean iconicity ratings

author: Tuomo Häikiö
affiliation: Department of Psychology and Speech-Language Pathology, University of Turku, Turku, Finland
email: tuilha@utu.fi

author: Oksana Kanerva
affiliation: Department of Psychology and Speech-Language Pathology, University of Turku, Turku, Finland; Faculty of Arts, University of Helsinki, Helsinki, Finland
email: okskan@utu.fi