Sound symbolic potential of Russian onomatopoeias: evidence from eye-tracking

Oksana Kanerva[1, 2], Tuomo Häikiö[1]

1 Department of Psychology, University of Turku, Finland

2 Faculty of Arts, University of Helsinki, Finland

Author Note

Oksana Kanerva https://orcid.org/0000-0001-7309-0623

Tuomo Häikiö https://orcid.org/0000-0001-6272-0255

We have no conflicts of interest to disclose.

Correspondence concerning this article should be addressed to Oksana Kanerva, Assistentinkatu 7, 20500, Turku, Finland, oksana.kanerva@helsinki.fi

Abstract

We investigated whether native Finnish speakers can grasp the meaning of Russian onomatopoeic words without any prior knowledge of the Russian language. In Experiment 1, elicitation test, naïve listeners generated associations for the acoustic events depicted by onomatopoeic words they heard. A cluster analysis suggested presence of different types of cues that affect the elicitation of associations. In the Facilitating cluster, associations were mostly correct; in the Counteracting cluster, they were predominantly incorrect. Worthy of note, many of the incorrect associations were systematic. In the Mixed cluster, there was a combination of cues; and in the Undefined cluster, no discrete cues affecting the formation of common associations were found. In Experiment 2, the same stimulus words were used in an eye-tracking experiment using visual world paradigm. It was shown that the participants have even better chances to map the onomatopoeic words to the correct semantic domain when extra-linguistic information is available, in this case target images presented on the experimental display. The availability of both audio and visual inputs substantially boosted this process in all four clusters. Our findings support the view that imitative sound symbolism offers a scaffolding material for connecting onomatopoeias to their referents when words are pronounced in isolation. Cross-linguistic sound symbolism offers a good explanation to the presence of different cues that affect semantic recognition of unknown onomatopoeic words. On a larger scale, cross-linguistic similarities in onomatopoeias may be part of a broader phenomenon, universal sound symbolism, form-meaning mapping shared by a wide array of languages.

Keywords: sound symbolism, imitative, onomatopoeia, elicitation test, visual world paradigm

This paper focuses on the type of sound-meaning relations that is based on direct imitation of natural sounds and is found in onomatopoeias. It sets out to answer the question whether it is possible to grasp the meaning of onomatopoeic words without any prior knowledge of the language in question. Onomatopoeias (Finnish *pam*, German *bumm*, Swedish *pang*, all equivalents of English *boom*) are generally considered to be shaped by the environmental sounds they represent. In this study, an onomatopoeic word is regarded as "a word based on an approximation of some non-linguistic sound but adapted to the phonemic system of the language" (Oswalt, 1994, p. 293). Onomatopoeias are believed to demonstrate fairly "direct mapping between the acoustic features of the sound itself and the phonological features of the word that label that sound" (Rhodes, 1994, p. 279). They are regarded as examples of true (Lyons, 1968; Sapir, 1970) or imitative sound symbolism (Hinton et al., 1994).

Sound symbolism, in its turn, has been perceived as an extensive term embracing various instances of "non-arbitrary mapping between sound and meaning" (Childs, 2015, p. 284). While in arbitrariness meaning is assigned to words by convention and often in an unpredictable manner, non-arbitrariness concerns all those cases when "aspects of a word's meaning or grammatical function can be predicted from aspects of its form" (Dingemanse et al., 2015, p. 604). Two distinct, but not mutually exclusive, forms of non-arbitrariness are iconicity and systematicity. Iconicity involves "resemblance-based mapping between aspects of form and meaning" (Ibid.), while systematicity refers to "statistical regularities in form that allow one to make predictions about meaning" (Motamedi et al., 2019, p. 190).

Sound symbolism is a phenomenon that manifests itself at different levels of a linguistic system. It embraces various kinds of non-arbitrary relations between sound and meaning, ranging from instances that are iconic to the ones showing systematic regularities in sound-meaning mapping. In more detail, it concerns entities that involve one-to-one and iconic resemblances between form and meaning such as onomatopoeias. It also touches upon more subtle relations between certain aspects of form and meaning that are assigned to sounds and their combinations and are found across sensory modalities. Sound symbolism even involves systematic regularities for sound patterns and aspects of meaning, as phonological similarity of words or intonational patterns used with certain words can statistically correlate with their semantic similarity or similarity in linguistic functions (cf. Cuskley & Kirby, 2013; Hinton et al., 1994; Childs, 2015).

In natural languages, sound symbolism is robustly exemplified by ideophones (e.g., Japanese *gorogoro* “multiple heavy objects rolling”, Ewe *potopoto* “sound of a big drum”, Zulu *shishilizi* “glide along”), a class of marked words that vividly represent events, ideas or sensory imagery through sound (Dingemanse et al., 2015; Voeltz & Kilian-Hatz, 2001). Sound-meaning relations have been tested on ideophones, for instance, by Dingemanse et al. (2016) who showed that native Dutch speakers guessed the meaning of stimulus words from Japanese, Korean, Siwu and Ewe with above chance accuracy having no prior knowledge of these languages. Similarly, Iwasaki et al. (2007) showed that naïve listeners, namely native English speakers, attributed similar meanings to some groups of Japanese ideophones, such as the ones depicting laughing, as native speakers of Japanese. Ideophones also facilitate learning; Imai et al. (2008) demonstrated that 3-year-olds learned novel verbs better when they were ideophonic.

The meaning of unknown non-ideophonic words can also be successfully guessed based on how these words sound (e.g., Brown et al., 1955; Tsuru & Fries, 1933). For instance, Tsuru and Fries (1933) tested English speaking participants in Japanese, an unfamiliar language to them. Although Japanese is known to be rich in ideophones, descriptive words from general vocabulary such as *mushi* “worm”, *tori* “bird”, *fune* “boat”, *tako* “kite” were used as stimuli. It was shown that the participants selected English counterparts for antonym pairs in Japanese with above chance expectancy. Furthermore, in Brown et al. (1955), monolingual, native English speakers matched antonym pairs presented to them in Hindi, Czech, and Chinese with English equivalent pairs with a higher than chance accuracy for each language. Other studies have been conducted using English speaking participants with a variety of languages, including Japanese and Croatian (Maltzman et al., 1956), Hindi and Chinese (Weiss, 1963), Hebrew, Japanese and Chinese (Brackbill & Little, 1957) with similar results.

While the studies mentioned above have tested languages that differ substantially, Leskinen (1998) tested semantic recognition of real onomatopoeic words in a cross-linguistic association test among native speakers of two closely related languages, Finnish and Hungarian. The respondents, naïve listeners of the target language (Finnish speakers had no knowledge of Hungarian and vice versa) recognized the semantics of animate (Fin. *bää-bää*, Hung. *be-e-e* for sheep; Fin. *röh-röh*, Hung. *röf-röf* for pigs) onomatopoeias well. However, the participants had poor performance in eliciting the correct meaning of inanimate onomatopoeias (Hung. *tyhú*, Fin. *viuh* for whipping) or words depicting human bodily sounds (Hung. *trr*, Fin. *niisk* for blowing the nose). There was a low number of correct associations generated by naïve listeners in connection with words imitating inanimate sounds or human bodily sounds.

Another study, conducted by Laing (2017), employed eye-tracking to demonstrate processing advantage for infants for onomatopoeic forms (e.g., *moo, baa*) over corresponding conventional words (e.g., *cow, sheep*). It is worth mentioning that this experimental study used predominantly animate onomatopoeias with a few inanimate words as stimuli in order to test the role of the "sound symbolism bootstrapping hypothesis" (Imai & Kita, 2014) in learning new words. This hypothesis claims that onomatopoeias have advantage over more conventional words in language acquisition as they help associate speech sounds with their referents and in this way establish the lexical meaning of unknown words. According to this hypothesis, when infants hear onomatopoeic words and see objects those words depict, sound symbolism helps them to integrate two types of inputs and connect the words to their referents. Indeed, Laing (2017) found that infants performed better in matching target images to onomatopoeic forms than to corresponding conventional words.

In the current study, we tested whether individual onomatopoeic words are capable of evoking associations with the corresponding acoustic events when the listener has no knowledge of the language. In particular, this research aimed at checking the sensitivity of naïve listeners, native Finnish speakers, to imitative sound symbolism naturally occurring in Russian onomatopoeic words (e.g., *bac* “bang”, *pljukh* “splosh” and *zvjak* “tinkle”). Although we did not compare the performance of infants and adults in form-meaning mapping in learning tasks, the sound symbolism bootstrapping hypothesis by Imai and Kita (2014) may be a good explanation of how adult naïve listeners access the meaning of iconic entities such as onomatopoeias.

As indicated by Nielsen and Dingemanse (2021) in connection with the sound symbolism bootstrapping hypothesis by Imai and Kita (2014), the empirical evidence for form-meaning mapping being instrumental in establishing referentiality is scarce and does not seem to account for any research on adults. Nonetheless, we believe that imitative sound symbolism may serve as a scaffolding material in connecting speech sounds and their referents, especially when two types of input are available, i.e., the word is pronounced and the picture demonstrating its semantics is presented.

Important for our setup, onomatopoeias may differ substantially between languages. Viimaranta et al. (2016) examined words typically used by Finnish and Russian native speakers for depicting sounds pertaining to water. In their study, the participants were expected to generate real onomatopoeic words from their native vocabulary in response to video recordings of acoustic events connected to water. It was shown that the words associated with each particular water-related sound were very different by their phonological structure, frequency of use and semantic implications in the two languages.

The material for this study is composed of 50 Russian onomatopoeic words. The central common feature for these words is that they can iconically depict sounds and indicate action connected to these sounds, i.e., point at instances of events when these sounds occured (cf. Kanerva, 2018, 2019). Therefore, in Russian it is quite common to use them not only as regular onomatopoeias (Rus. *Bakh! – kniga upala na pol*, Eng. “Bang! – the book fell on the floor”), but also similarly to verbs in predicate function (Rus. *što-to bakh v steklo/ što-to bakhnulo*-V *v steklo*, Eng. “something banged into the window”). The class of words tested in this study is special in a sense that both features are easily traceable. This sets them apart from such words as *morg* “blink”, *nyrk* “plunge into water”, *tolk* “push” that can indicate kinetic movements, but seem to lack sound depiction capacity. The stimulus words chosen for this study also differ from such onomatopoeic words as *gav* “woof”, *kukareku* “cock-a-doodle-do” and *mu* “moo” that imitate sounds produced by animals for communication. Animal cries do not depict action, but reproduce the act of ‘saying’ itself and are intended to affect the recipient of that signal. As this research focuses on semantic recognition of onomatopoeic words that imitate both natural sounds as well as action connected to these sounds, we do not include animal cries as stimuli.

Experiment 1

In the first experiment, we assessed the types of associations naïve listeners generate when they hear onomatopoeic words in a foreign language, namely in Russian. In this way, we tested whether it is possible to elicit accurate guesses about the semantics of Russian onomatopoeias without any advanced knowledge of the language. For this purpose, we conducted an elicitation test. This method allows checking whether the meaning of unknown words can be accessed when the number of potential referents is substantially large, if not unlimited, and no other cues, except for the way the token sounds, are available.

The elicitation test (also known as association test) is a rather unbiased approach to measure the associative semantics of stimulus words (Deese, 1965). The central idea behind this method is that response words represent the associative meaning of the stimulus words. The results of such tests reflect the participants' conscious or subconscious understanding of the common core of meaning between the stimulus concepts and the concepts generated in response.

Given the fact that elicitation tests can be used to retrieve meaning of individual words from familiar vocabulary (Kiss, 1975; McRae et al., 2005; Vinson & Vigliocco, 2008) as well as to assess word-association behaviour in foreign language acquisition (Fitzpatrick & Izura, 2011; Meara, 2009), we decided to use this method to measure the semantic accessibility of unknown words with imitative sound symbolic potential. Word associations created by the respondents were expected to mirror how strong the relations between form and meaning of individual onomatopoeic words are. We were interested in the accuracy of immediate, close to subconscious associations generated by naïve listeners.

Method

Participants

Twenty-three students participated in the experiment (1 male, 22 females, mean age 25.57, SD 6.19). All of them reported themselves as native speakers of Finnish with no prior Russian language studies and no knowledge of Russian. All participants were volunteers. As the participants were volunteering adults and no potential harm was inflicted, no ethical review was necessary.

Material

We selected 50 onomatopoeic word forms in Russian as stimuli. The semantics of 49 tokens was verified with dictionaries (Efremova, 2000; Šljakhova, 2004). We included *om* "gulping", mentioned by Kanerva (2020), as a synonym to *am* "eating" into the stimuli.

Semantically, the tokens ranged from depicting banging (*bac, trakh*), explosion (*šarakh, khlobys'*), and falling (*bukh, grokh*); sounds pertaining to natural elements, e.g., water (*pljukh, bultykh*), fire (*pykh*), and air (*vžik, f'ju-f'ju*); sounds produced by mechanisms, e.g., lock (*ščelk*), clock (*tik-tak*), and train (*tu-tu)*; human bodily parts, e.g., whistling (*svis'*) and footsteps (*top-top*); and tools/other tangible objects, e.g., matches (*čirk*), chain (*zvjak*), door (*khlop, švark*).

A recording was made of all fifty items in a randomized order, spoken by a native speaker of Russian who also speaks Finnish as a foreign language. Every participant listened to this same recording. The presentation of each token was preceded by an index number (1 to 50) in Finnish which matched the numbers in the answer sheet. The word was repeated after a pause of approximately one second. After each token there was a pause of approximately 12 seconds during which the respondents wrote down their associations.

The meaning of each token subsequently used in analyzing the results is given in Appendix A. It reflects the order in which the participants heard the tokens in the recording.

Procedure

The experiment was conducted as a group testing at the beginning of a lecture. The participants were instructed that they were going to hear Russian words which are used to imitate natural sounds. They were asked to write down their associations with what could have produced the sounds depicted by each word in the recording. The participants were instructed to write only one association per each item during the pause between the words. The instruction suggested such possible sound sources as tangible objects, natural elements like water, air or metal, or events like falling or hitting. Two typical Finnish onomatopoeic words referring to splashing and banging were given as examples.

Results

The analysis of answer types

The definitions of the words (see Appendix A) were used to rate the answers on the binary scale as correct or incorrect. The response was considered to be correct if any of the key points in the definition was mentioned in the association given by the respondent. For example, the answer was marked as correct if the participants used the word “water”, “liquid”, “to splash”, etc. in connection with words such as *khljup* “squish”, *ljap* “plop”, *bultykh* “plunge” that depict water-related sounds, or words such as “hit” or “break” for *khlobys'* “strong hit” and *brjak* “metallic objects falling/something beaking” depicting the corresponding acoustic events. Some responses were more detailed than others. For instance, *khrust'* “crack” was not only associated with something “wooden” or “dry” as some participants wrote longer answers such as “biting bread”, “raking leaves”, “breaking a thin piece of wood”, and “moving through reeds”. All these associations were regarded as correct, because all of the instances mentioned were dry, crunchy and could make a cracking sound. As for the single source objects like trains (*čukh-čukh*, *tu-tu*) or clocks (*tik-tak*) and concrete human bodily sounds like sneezing (*apčkhi*, *čikh*) or eating (*om-om*, *am*), the instances had to be named explicitly. Furthermore, missing answers were coded separately. The associations of 10 respondents were coded independently by two raters. The inter-rater agreement as measured by Cohen’s kappa was 0.77. The rest of the associations were coded by one rater. The numbers of correct, incorrect and missing answers for each onomatopoeic word are given in Appendix A.

The answers were further analyzed in order to find repeated patterns in the formation of associations with the stimulus words. First of all, the correct responses were tagged Matching as they matched the lexical meaning given in Appendix A. Second, when the incorrect answers were examined, we noticed that for some stimulus words, there were repeated answers which consistently mismatched the established semantics of those tokens. If answers of three or more participants referred either to the exact same sound source or the same semantic field, they were regarded as not occurring by coincidence and were tagged as Mismatching as they mismatched with the definition in Appendix A. For instance, participants often associated *kap* “water dripping” with metallic hoofs/heeltaps using such expressions as “hoofs”, “horse clops”, “taps of high heels” or “hoofs clop”, *zvjak* “tinkle” with electricity using expressions such as “electric shock” or “electric wave”, *trakh* “hit” with dragging, and *grokh* “fall” with human bodily sounds. It is worth emphasizing that although the answers did not match the semantics of these words in Russian, they still followed certain patterns. Third, the number of incorrect answers not showing a shared pattern across associations was calculated together with missing data and regarded as Random answers. It is important to highlight that the missing data comprises all instances where the respondents did not write anything. The absence of discrete regularities in these answers was the central reason for joining them together into the Random answers.

Cluster analysis

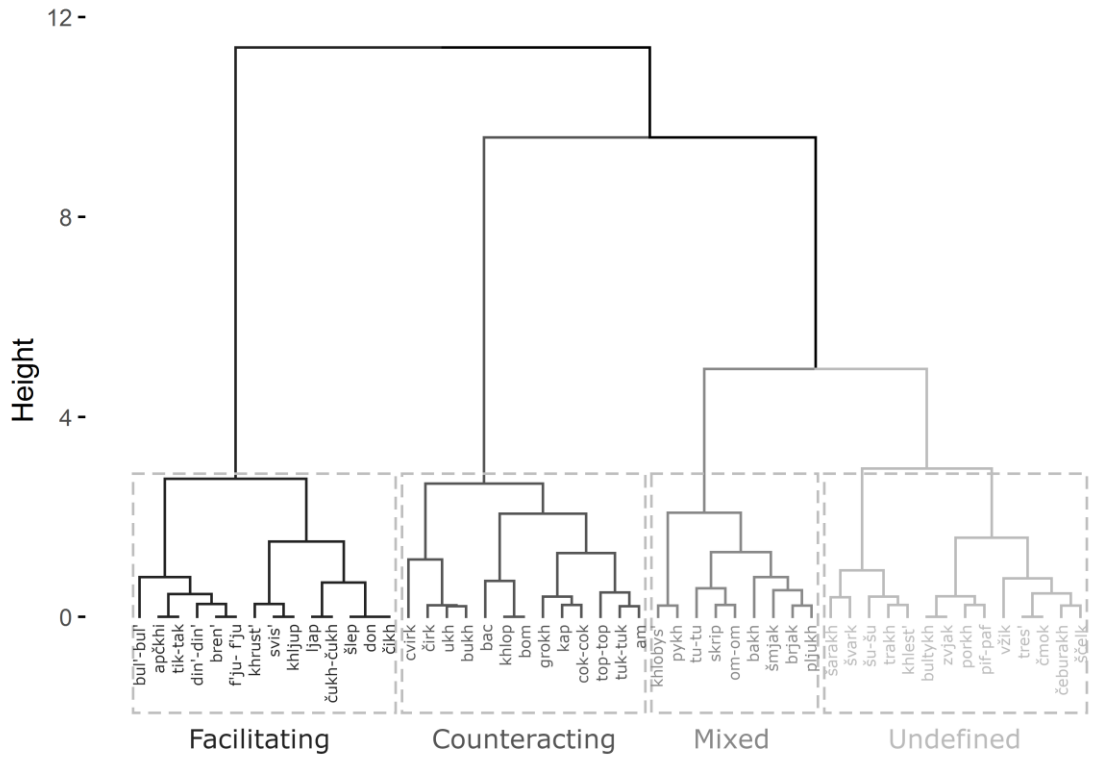
After dividing the words to three groups (Matching, Mismatching, and Random), cluster analysis was conducted using the cluster package (version 2.1.0; Maechler et al., 2019) for R statistical software (version 3.6.3; R Core Team, 2020). The visualization was done with the factoextra package (version 1.0.6; Kassambra & Mundt, 2019). To see whether the selected words can be reliably put in different clusters on the basis of the type of answer (Matching, Mismatching, and Random), we used hierarchical clustering. The dissimilarity matrix was computed using the Euclidean method and the clustering was computed using Ward’s minimum variance method (D2). The data was scaled and centered prior to the analysis. The Hopkins statistics was 0.818, denoting that the dataset was significantly clusterable.

The resulting dendrogram with 4 clusters can be seen in Figure 1. The clusters for individual tokens are reported in Appendix A. The gap statistics for K means [fviz_nbclust()]

suggested that the optimal number of clusters is 3. However, looking at the solutions, both 3-cluster solution and 4-cluster solutions are possible, as reasoned below. In both 3- and 4-cluster solutions, clusters 1 and 2 (two clusters on the left of Figure 1) are the same while the cluster on the right is divided into two groups for the 4-cluster solution. Because of this, we decided to take a closer look at the answer types in each possible cluster.

[Insert Figure 1 about here.]

Figure 1.
The dendrogram for Facilitating/Counteracting/Mixed/Undefined clusters.



Cluster 1 consisted of 14 tokens (*bul'-bul'* “glug”, *apčkhi* “sneeze”, *tik-tak* “clock ticking”, *din'-din'* “small bell tolling”, *bren'* “strings vibrating”, *f'ju- f'ju* “flying/blowing”, *khrust'* “cracking”, *svis'* “whistle”, *khljup* “splash”, *ljap* “plop”, *čukh-čukh* “train”, *šlep* “slapping or walking clumsily”, *don* “big bell tolling”, *čikh* “sneeze”) that elicited lots of Matching answers and almost no Mismatching or Random answers. The formation of this cluster suggests the existence of facilitating cues that helped participants generate semantically similar associations with ones already established in Russian. This cluster is referred to as Facilitating.

Cluster 2 consisted of 13 tokens (*cvirk* “water trickling”, *čirk* “match strike”, *ukh* “cannon fire”, *bukh* “falling”, *bac* “hitting”, *khlop* “bursting/hitting”, *bom* “big bell tolling”, *grokh* “falling”, *kap* “water dripping”, *cok-cok* “clatter”, *top-top* “stomping”, *tuk-tuk* “knock”, *am* “eating”) that elicited a high number of Mismatching answers. This cluster points at the existence of cues which prevented naïve listeners from guessing the meaning of Russian onomatopoeias correctly. This cluster is referred to as Counteracting.

Cluster 3 consisted of 9 tokens (*khlobys'* “hitting”, *pykh* “burning”, *tu-tu* “train”, *skrip* “squeak”, *om-om* “gulping”, *bakh* “bang”, *šmjak* “falling”, *brjak* “breaking”, *pljukh* “splosh”) that elicited a considerably high number of both Matching and Random answers. This suggests the presence of a combination of different types of cues, which led us to call this cluster Mixed.

Finally, Cluster 4 consisted of 9 tokens (*šarakh* “explosion”, *švark* “hitting”, *šu-šu* “whisper”, *trakh* “hitting/breaking”, *khlest'* “lashing”, *bultykh* “plunge”, *zvjak* “tinkle”, *porkh* “flying”, *pif-paf* “shooting”, *vžik* “buzzing”, *tres'* “cracking”, *čmok* “kissing”, *čeburakh* “falling”, *ščelk* “click”) that elicited a high number of Random answers while the other types of answers were considerably low. This suggests absence of discrete cues affecting the formation of common associations which led us to call this cluster Undefined.

The means and standard deviations of Matching, Mismatching, and Random responses by the 23 participants for the words in each of the four clusters are presented in Table 1.

[Insert Table 1 about here.]

Table 1.
The mean number of words from each response type that were grouped into each cluster (SDs in parentheses). Columns sum up to 23, since there were 23 participants.

Response type	Cluster			
	Facilitating	Counteracting	Mixed	Undefined
Matching	17.93 (3.00)	2.61 (3.01)	8.78 (1.72)	2.14 (1.29)
Mismatching	0.00 (0.00)	14.69 (3.40)	3.78 (2.54)	3.50 (3.20)
Random	5.07 (3.00)	5.69 (2.43)	10.44 (2.88)	17.36 (2.82)

As stated before, in both 3-cluster and 4-cluster solutions, cluster 1 (Facilitating) and cluster 2 (Counteracting) are the same, but the 4-cluster solution further distinguishes cluster 3 (Mixed) and cluster 4 (Undefined). Because clusters 3 and 4 seem to consist of considerably different answer types, we think the 4-cluster solution is more plausible than the 3-cluster solution.

Discussion

In Experiment 1, Finnish participants completed an elicitation test in which they wrote down what they thought was the sound source of the Russian onomatopoeic words. The cluster analysis showed that there were four clusters with distinct proportions of answers that matched the correct meaning, had a mismatch with the correct meaning, but a certain pattern between the answers, or seemed completely random.

In the Facilitating cluster, the participants very accurately elicited the meaning of the words. The existence of this cluster suggests sensitivity to imitative sound symbolism among naïve listeners. We believe that the words in this cluster offer facilitating cues for mapping the form with the correct semantic domain even without any prior knowledge of the tested language.

The Counteracting cluster was characterized by a high number of Mismatching answers. The participants had associations with other concepts which did not correspond with the established semantics of these words. For 6 words from this cluster, the participants elicited responses connected to living beings as options for the possible sound sources. In particular, *am* “eating” was associated with cows, *čirk* “match strike” and *cvirk* “water trickling” – with chirping of birds or crickets, and *ukh* “cannon fire”, *bukh* “falling” and *grokh* “falling” – with dogs, pigs and other animals. Obviously, the associations with animate sound sources induced by the counteracting cues were strong enough for many participants to prefer them. Other words had the following shared mismatching associations: *bac* “hitting/bursting” was incorrectly associated with water, *khlop* “hitting” with jumping, *bom* “big bell tolling” with bouncing, *kap* “water dripping” with hoofs and knocking, *cok-cok* “clutter” with scissors and chopping, *top-top* “stomping” with knocking, and *tuk-tuk* “knocking” with tractors, trains and other vehicles.

The Mixed cluster was characterized by a high number of both Matching and Random answers with some mismatching answers also present. This suggests that imitative sound symbolism of these words gave mixed cues to what the meaning of them could be. Some participants elicited correct guesses, some of them had shared associations with other objects. The rest of the responses did not follow a distinct pattern, i.e., they belonged to Random answers. It is interesting that from this cluster only *bakh* “bang” had a considerably high amount of shared responses connected to animal cries (4 associations with dogs and barking). The other words had a relatively low amount of shared mismatching associations. Among them, *tu-tu* “train” was associated with cars (the highest number of shared responses among the words listed here at the level of 5), *om* “eating” with meditation, *bakh* “bang” with snorting, *šmjak* “falling” with smacking, and *pljukh* “splosh” with blowing.

Finally, the Undefined cluster had a high number of Random answers. These words constituted a major difficulty in linking their form with the correct semantic domain, suggesting the absence of easily defined cues to fulfill the objective, i.e., undefined cues prevailed. In particular, the mean for Matching answers was the lowest and the mean for

Mismatching responses was the second lowest.

The elicitation test results do not confirm the ones reported in Leskinen (1998) that naïve listeners are (almost) incapable of guessing the meaning of words imitating inanimate sounds or human bodily sounds correctly. We will come back to this issue in General Discussion.

Our findings of Experiment 1 align with the assumption that non-arbitrary names reduce the cognitive effort to “learn, retain, and actively employ" onomatopoeic words (Berlin & O’Neill, 1981, p. 259). On a general level, they support the claim that language users are able to grasp general semantics of words without knowing their meaning. Some of our stimulus words belonging to inanimate onomatopoeias can be regarded as partial synonyms to one another (e.g., *trakh, bakh*, *švark*, all depicting strong hitting), while human bodily sounds (*apčkhi* “sneezing”, *top-top* “stomping”, *om-om* “gulping”, *čmok* “kissing”) belonging to animate onomatopoeias seem to be more specific. However, human bodily sounds could be found in each cluster. This observation suggests that it is the phonological structure of the word itself that facilitates guessing or intervenes with this process but not necessarily how specific the semantics of the given word is.

In conclusion, the results of Experiment 1 revealed that people with no knowledge of a language can guess the general semantics of some unknown words even in the absence of contextual cues. We believe that the partial word-knowledge in case of onomatopoeias originates from the form-meaning mapping these words naturally exhibit.

Experiment 2

In general terms, elicitation tests and multiple-choice tests provide information whether the choice or association are correct or not. Even though they do not offer any information about the amount of time and degree of doubt involved in making the decision, the results of

Experiment 1 were crucial in establishing the direction of our further investigation and the focus of prospective data analysis. In Experiment 2, we used the visual world paradigm to assess how naturally naive listeners could guess the semantics of unfamiliar words based on auditory and visual inputs. As far as the research on eye movements provides real time experimental data, it is typically used to accurately examine cognitive processes involved in language perception. The visual world experiments allow assessing how the listeners perceive and understand spoken expressions. The scenes that are presented on the experimental display usually contain objects referred to either in the task (Allopenna et al., 1998) or in the sentences linked to the scenes (Altmann & Kamide, 2007). In a typical setup, the participants are shown a 2 x 2 grid of pictures. The participants have a short period of time to examine the pictures, after which the spoken utterance is played on the speakers. While the participants observe visual stimuli, their eye movements are recorded for further analysis of the processing mechanisms involved in the perception of those objects. The accessibility of an extra-linguistic referent on the screen induces the activation of conceptual and lexical identification, which, in its turn, can be witnessed in the subsequent eye movement patterns.

For instance, in their classic study, Allopenna et al. (1998) showed the participants a target picture (e.g., a beaker), a phonological competitor with an onset overlap (a beetle), a phonological competitor with a rhyme overlap (a speaker), and an unrelated distractor (a carriage). The spoken utterance had an instruction "Pick up the beaker". The results showed that after the onset of the target word in the sentence (beaker) the participants were more likely to direct their gaze at the beaker and beetle than the other two pictures. Towards the end of the word "beaker", the likelihood of fixating the beetle started to decrease while the likelihood of looking at the speaker started to increase. This experiment demonstrated how fixations to pictures are triggered by lexical activation, and how the phonological competitors were also activated when the word was heard. Furthermore, the time span of the activation was readily witnessed in the fixation patterns.

In later studies, it has been shown that not only phonological information but also other information drives fixations towards certain pictures. For example, Huettig and Altmann (2005) used semantically related objects (e.g., a trumpet) in the same grid as the target object (a piano) and witnessed an increased likelihood of fixating the semantically related object in comparison to an unrelated distractor (a hammer). Furthermore, Huettig and Altmann (2007) demonstrated that the participants were more likely to fixate a visually similar object (e.g., a cable) when they heard the target word ("snake") even though the grid included the target object and the participants had a five-second period of viewing the grid prior to the word onset. To sum up, these findings show that when the participants hear a spoken utterance, different types of information about the target word are activated, and thus driving the fixations.

In the current study, we were interested in seeing whether the participants can recognize the sound source of the onomatopoeic words. As long as using phonological competitors would be detrimental to the recognition of the correct picture, we used semantically and visually similar competitors. Semantic competitors to the target pictures were chosen in a similar way as in Huettig and Altmann (2005). In particular, we selected them based on the real-life relation between the depicted objects. For example, Russian *bultykh* stands for a big object coming into contact with the surface of water. For the target picture, we chose a boy jumping into the swimming pool. Semantic competitor to the target was supposed to have real-life connection with it, i.e., appear in similar circumstances or be functionally attributed to the target. As a result, we chose a life ring for the semantic distractor since they are generally associated with swimming and safety instructions require them in swimming pools. At the same time, it was presented in a static manner as if attached to the wall in order to exclude associations with it being thrown into water. As for the visual distractor, we selected a female practicing martial arts sitting on a blue tatami of the same shade of colour as the water surface in the target picture. The girl was sitting in the same position as the boy who was jumping into the water. Additionally, her hair colour matched with the boy's. The unrelated control picture (the fourth image in the quadruplet) was supposed to be unrelated to any of the other three pictures in terms of semantic, visual, or phonological similarity. A hairbrush was chosen for this purpose.

Method

Participants

Twenty-seven Finnish adults (26 female, mean age 23.83 years, SD 4.83) participated in the study. They were students and received course credit for the participation. All of the participants had Finnish as their L1 and none of them had taken any Russian studies. All of the participants had normal or corrected-to-normal vision and had no hearing deficits. All participants were volunteers. As the participants were volunteering adults and no potential harm was inflicted, no ethical review was necessary.

Apparatus

Eye movements were recorded with the EyeLink 1000 (SR Research, Canada), using the 1000Hz sampling rate. Viewing was binocular, but only the movements of the right eye were recorded. The pictures and sounds were presented on a 24 inch BenQ XL2411 monitor with a resolution of 1920*1080 pixels and refresh rate 144Hz.

Material

The sound stimuli were the same as in Experiment 1. The pictures were selected from the Papunet.net picture bank (https://kuvatyokalu.papunet.net). We selected 50 pictures depicting the sound source of the Russian target words (targets), 50 pictures depicting an object with a semantic relationship to the target word (semantic competitors), 50 pictures depicting an object that visually resembled the target picture (visual competitors), and 50 pictures depicting an object that had nothing to do with the target word (control distractors). This resulted in 50 picture quadruplets. None of the competitors or distractors depicted an object that could produce a similar sound as the target. None of the pictures were used twice. Each quadruplet included pictures that shared the colour scheme: all four pictures were either black-and-white or color drawings. The pictures were edited to tone down the saturation and unify the color scheme. Slight edits were made to make sure the pictures did not contain text or numbers, and that the visual competitors and targets were as visually similar as possible. The sizes of the pictures were edited to the resolution of 400*400, subtending 10.6 degrees of visual angle.

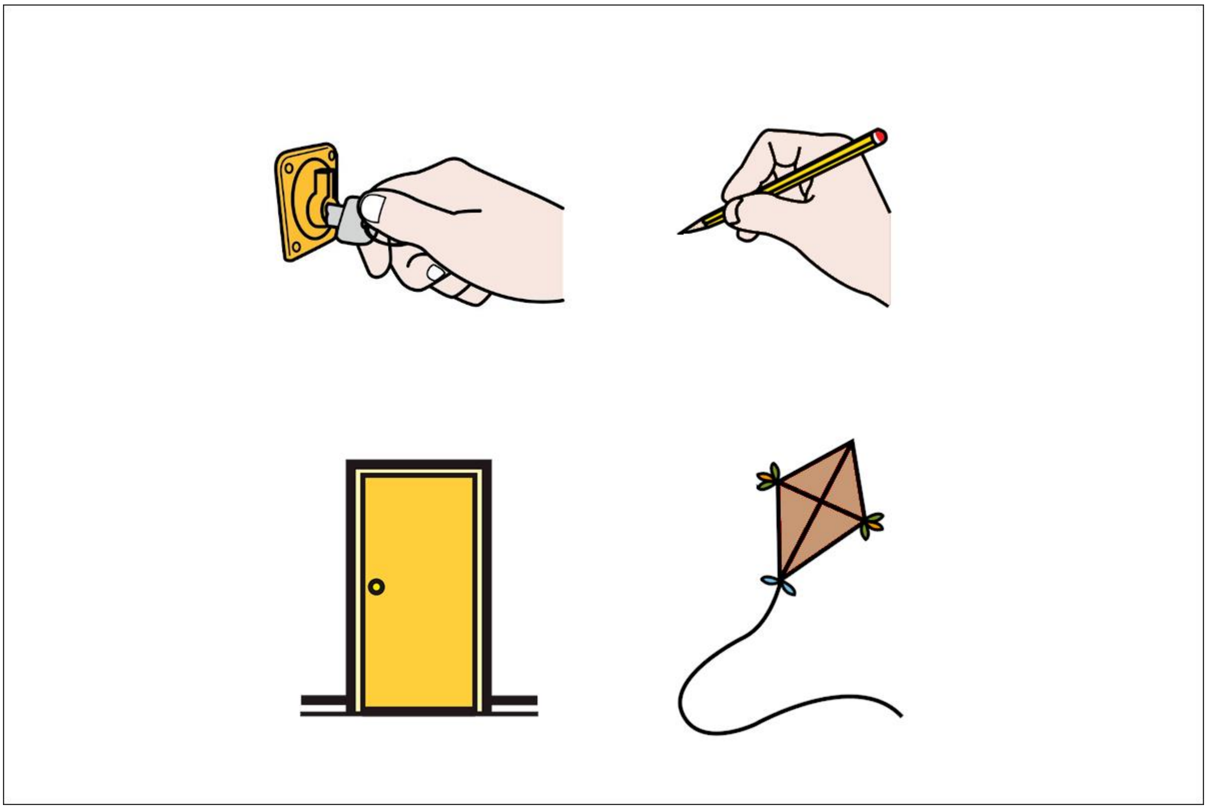
In order to make sure that the semantic pictures were related to the targets while the visual and control pictures were not, 30 participants (19 female, mean age 43.60 years, SD 14.36) rated the perceived similarity of picture pairs in a web-based Webropol survey. The survey included 150 picture pairs (each target paired with the semantic, visually similar, and control picture, respectively), and the task was to rate the semantic connection between the picture pairs on a 7-point Likert scale (no connection at all – strong connection). The Intra-class correlation between the raters was ICC = .568. Five semantic pictures had a mean rating below 4 (lower than medium connection), and seven visually similar pictures had a mean rating above 4 (higher than medium connection). All of the control pictures had a mean rating below 4. New pictures were searched to replace the 12 pictures that had either too low or too high ratings for the semantic and visually similar condition, respectively. The resulting new picture pairs were rated by 50 participants (12 female, mean age 39.40 years, SD 8.46) who had not taken part in the first rating (ICC = .462). Now all of the semantic pictures had a mean rating above 4 and visually similar pictures below 4. The mean semantic ratings were 5.77 (SD 0.70), 2.38 (SD 0.66), and 1.90 (SD 0.52) for the semantic, visually similar and control pictures, respectively.

Each trial (50 in total) consisted of the picture quadruplets arranged in a 2 by 2 grid around the center of the screen (see Figure 2). The picture order was predetermined using a latin square design to counterbalance the different order combinations. The target word started sounding out at 2000ms from the picture onset. After the target word, the trial lasted for an additional 3500ms.

[Insert Figure 2 about here.]

Figure 2.

An example of a stimulus screen. Upper left (A) is the target picture for the sound ščelk that depicts mechanical clicking of a lock, for instance, upper right (B) is the visually similar competitor, lower left (C) is the semantic competitor, and lower right (D) is the control distractor. Pictures are taken from the Papunet picture bank, papunet.net, Sergio Palao/ARASAAC (A, B, and D), Paxtoncrafts Charitable Trust (C); B and D are edited versions of the original pictures. Pictures are shared with the Creative Commons BY-NC-SA license.



Procedure

The eye-tracking experiment was conducted in the same session as a lexical decision experiment that was not related to the present study in any way. In the eye tracking experiment, the participants were told that they would hear Russian words that are used to imitate natural sounds. They were further told that they would see four pictures on each screen. The task was to listen to the word carefully, followed by looking at the picture they thought depicted the spoken word. The participants sat approximately 60cm from the screen. A chin-and-forehead rest was used to stabilize the head. A nine-point calibration was deemed successful if the average error did not exceed 0.50 degrees. However, for one participant, the calibration resulted in 0.85 degrees average error. As the visual data inspection indicated that the accuracy was satisfactory in order to differentiate eye fixations on different pictures, we included this data in the analyses. Before each trial, the participant fixated on the circle at the center of the screen, after which the pictures appeared.

Data preparation and statistical considerations

The preprocessing of the data was done using the R package VWPre (version 1.2.3, Porretta et al., 2016). 2.21% of the data was marked off-screen or trackloss. All of the 1350 trials had more than 75% data intact and were included in the analysis. We calculated the proportions of samples that fell within and outside each of the four interest areas using 40 ms bins. The proportions were converted to empirical logits. The data was modelled from 200 ms from the word onset in order to take the time to program a saccade into account (Fischer, 1992). The time series were analyzed with R version 4.0.3 (R Core Team, 2020) using generalized additive mixed modeling (mgcv, version 1.31-31; Wood, 2011). The visualizations were done with the itsadug package (version 2.3, van Rij et al., 2017).

Results

We modeled the empirical logit of looks to the target picture as a function of time by cluster (Facilitating, Mixed, Counteracting vs. Undefined) to test the differences between the clusters over time. For the parametric coefficients, the Facilitating cluster acted as the baseline. Simple random intercepts were included for the event (i.e., the combination of subject and item), which allows a unique intercept for each time series. Factor smooths were included for the subjects and items over time. The inverse of the empirical logit variance estimates was used as the weights in the model. To compensate for the autocorrelation, we extracted the Rho from the first model (0.285) and fitted a new model including the Rho. This model is presented in Table 2.

[Insert Table 2 about here.]

Table 2.

Generalized additive mixed model for the empirical logit of looks to the target picture as a function of time by cluster. The upper part reports the parametric coefficients and the lower part the smooths and random effects. For the parametric coefficients, the Facilitating cluster is the baseline.

Parametric coefficients	Estimate	SE	t value	p value
Intercept	1.5447	0.2291	6.743	< .0001
ClusterMixed	-0.3300	0.3357	-0.983	.3256
ClusterCounteracting	-1.0412	0.3021	-3.447	.0006
ClusterUndefined	-1.4895	0.2935	-5.075	< .0001
Smooth terms	Edf	Ref. df	F value	p value
Smooth: Time, ClusterFacilitating	5.163	5.623	16.316	< .0001
Smooth: Time, ClusterMixed	6.106	6.605	10.688	< .0001
Smooth: Time, ClusterCounteracting	5.479	5.957	6.619	< .0001
Smooth: Time, ClusterUndefined	3.575	3.883	5.629	.0002
Random effect: Time, subject	189.108	242.000	307.207	< .0001
Random effect: Time, item	357.338	446.000	386.737	< .0001
Random effect: Event	1263.749	1346.000	37.891	< .0001

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

The parametric coefficients showed that the clusters Counteracting and Undefined differed significantly from Facilitating. Furthermore, the conditional smooths for time revealed significantly different curves for different clusters. The proportions of looks to the target for each cluster are plotted as a function of time in Figure 3. As can be seen in Figure 3, the likelihood of fixating the target picture reaches above chance level (25%) for each cluster. The estimated difference curves are presented in Figure 4, showing the statistically significant areas between different clusters. Only the Facilitating and Mixed clusters did not differ significantly from each other at any point of time. We will return to this result in the Discussion of Experiment 2. The likelihood of fixating the target picture in the Facilitating cluster was higher than in the Counteracting or Undefined clusters, starting from 920ms and 658ms after the word onset, respectively. The likelihood of fixating the target picture in the Mixed cluster was higher than in the Counteracting or Undefined clusters within 1935-3145ms and 756-3309ms after the word onset, respectively. Finally, the likelihood of fixating the target picture in the Counteracting cluster was higher than in the Undefined cluster within 1345-1902ms after the word onset.

[Insert Figure 3 about here.]

[Insert Figure 4 about here.]

Figure 3.

The fixation proportions to the target for the four clusters as a function of time. The horizontal line denotes the chance level.

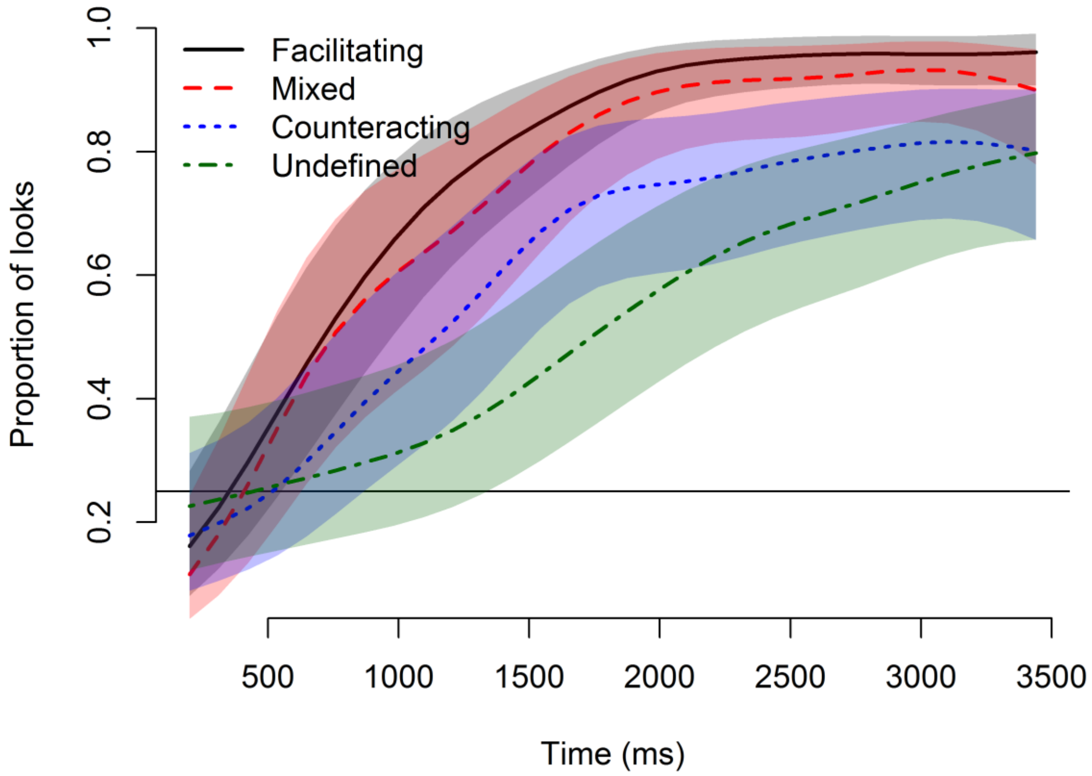
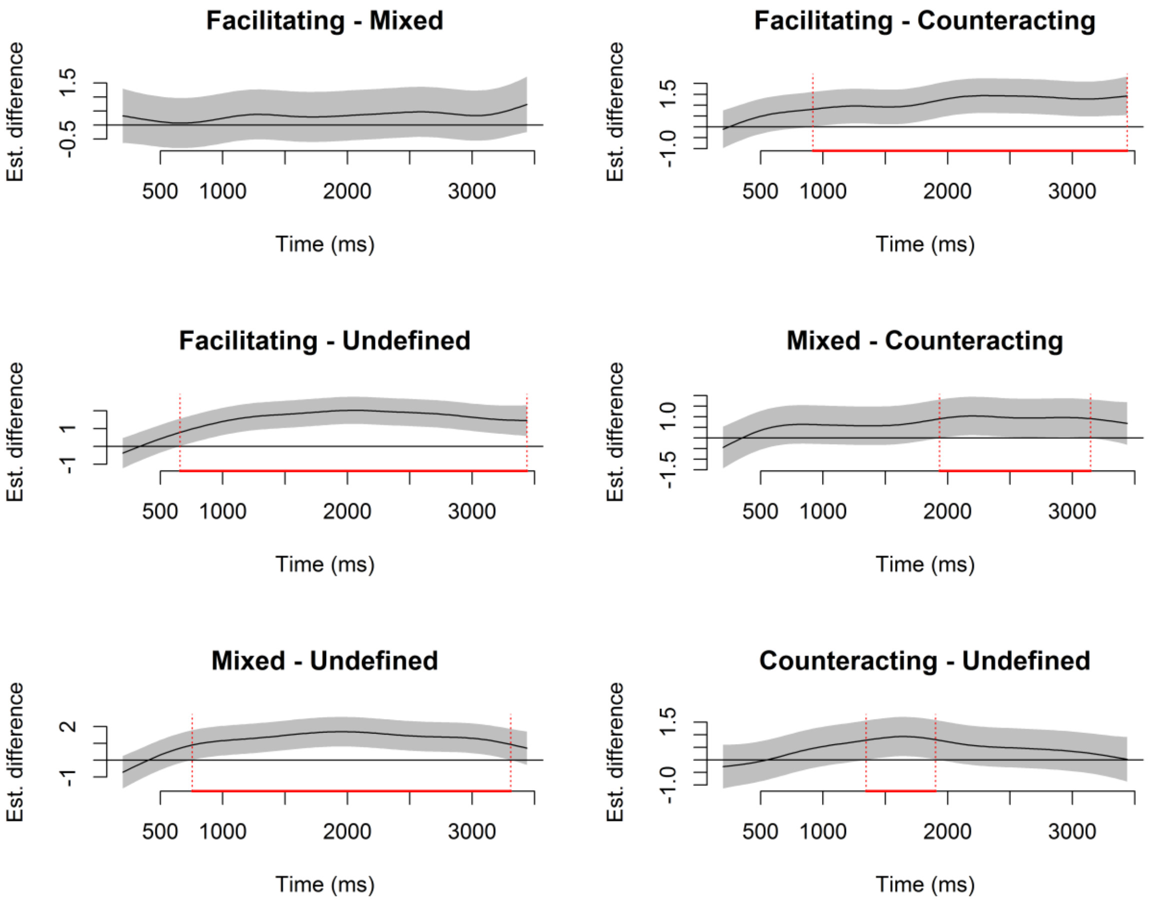


Figure 4.

The estimated differences of looks to target between the four clusters as a function of time. The statistically significant areas are within the two dotted lines.



To test the influence of the visual competitor and the semantic competitor between the Clusters, we modeled the difference in looks on empirical logit scale between the visual competitor and distractor, and the semantic competitor and distractor. The other details of the models were the same as for the main analysis. However, there were no significant differences between the clusters, either in the parametric coefficients or the conditional smooths over time. The full models of the competitor analyses are presented in the Appendix B.

Discussion

In Experiment 2, the Finnish participants took part in a visual world paradigm, in which they heard the spoken Russian onomatopoeic word and in which their task was to look at the target picture depicting the sound source of the heard onomatopoeic word. It was shown that the participants were able to recognize and thus fixate the target picture above chance level even with no prior knowledge of the language. This was true for all of the clusters established in Experiment 1, even though there were differences between the clusters in this respect, with the Facilitating cluster exhibiting the highest proportion of looks to the target.

While the Facilitating cluster showed rather expected behavior with regard to the proportion of looks, the most intriguing observations concerned the Mixed cluster. As stated in relation to Experiment 1, the cluster analysis based on the elicitation test results showed that this cluster contained a relatively high amount of Matching answers, i.e., second highest after the Facilitating cluster, combined with a high amount of Random answers, i.e., second highest after the Undefined cluster (cf. Table 1). As for the eye tracking experiment, extra-linguistic cues, i.e., the target picture depicting the acoustic event in question, allowed the participants to identify the target picture successfully even in the presence of three distracting pictures. In other words, these cues were instrumental in revealing the semantics of the stimulus words, approximating the results for the Mixed cluster to the ones demonstrated by the Facilitating cluster. This seems to be the primary reason why the Facilitating and Mixed clusters did not differ significantly from each other at any point of time in the eye tracking experiment.

Although all four clusters had above chance likelihood for the participants to fixate the target picture, both the Counteracting and Undefined clusters differed from the Facilitating and Mixed clusters in the fixation proportions to the target. In particular, the proportions of looks to the target was distinctly lower in the Counteracting cluster in comparison with the Facilitating cluster, and the Undefined cluster had effectively the lowest proportion of looks than any other cluster.

As stated in relation to Experiment 1, the words from the Counteracting cluster generated a significantly high amount of associations with animal cries. In the eye tracking experiment, distractor pictures did not feature animals similar to ones mentioned by the participants in the elicitation test, which could have attracted attention to themselves. Nonetheless, both Counteracting and Undefined clusters significantly differed from the Facilitating and Mixed clusters showing a noticeably lower likelihood of fixating the target. Such results point out that even the absence of cues that could generate responses mismatching with the meaning of the stimulus words (animals, in particular) and the presence of the correct target picture did not put the chances of the naïve listeners to grasp the general semantics of these tokens to the same level as the tokens from the Facilitating and Mixed clusters. Likewise, the absence of corresponding animals in the distractors for words from the Mixed cluster could not be responsible for its remarkable performance in the eye-tracking experiment for the same reason.

In the elicitation test, the Undefined cluster generated the highest number of Random responses with the lowest amount of Matching associations. These results indicate that the participants were unable to successfully map the form and meaning of unknown onomatopoeic words in this cluster. The analysis of the eye-tracking experiment revealed that the Undefined cluster had the lowest proportion of looks at the target picture, but still above chance likelihood of fixating the correct picture. In other words, in the presence of extra-linguistic cues, naïve listeners were able to integrate these two types of input, e.g., unknown words perceived by ear and the pictures depicting their semantics.

It needs to be noted, though, that not every single word was recognized in the eye-tracking experiment. The visual inspection showed that for three words in the Counteracting cluster (*grokh* “falling with noise”, *khlop* “hitting or bursting”, and *ukh* “cannon fire”) and two words in the Undefined cluster (*čeburakh* “falling clumsily” and *zvjak* “clinking”) the likelihood of fixating the target was not distinguishable from the competitors or the control distractor at any point in the trial.

General Discussion

The outcomes of this study demonstrate that naïve listeners can elicit correct semantic associations of unknown onomatopoeic words when the number of response options for the source of the sound in question is practically unlimited as in Experiment 1. The participants had even better chances to pinpoint the right semantic domain when extra-linguistic cues were available, i.e., pictures with the target pictures were presented on the experimental display in the eye tracking test in Experiment 2.

The outcomes of this study seem to be in discordance with those of Leskinen (1998), who showed that naïve listeners were (almost) incapable of guessing the meaning of words imitating inanimate sounds or human bodily sounds correctly. We believe this is due to the range of stimuli and the approach of presenting them. More specifically, the current study examined a substantially wider range of stimuli and used a different approach to present them. In Leskinen (1998), 23 Finnish and 30 Hungarian inanimate onomatopoeias were tested, but, in fact, some trials contained several words depicting the same sound. For instance, Hungarian *loccs, platy, plotty* depicting splashing were presented together to the Finnish speaking participants, and Fin. *klumps, plumps, läts* were presented together to the Hungarian speaking respondents. Thus, the experiment on Finnish-Hungarian imitative words had 16 trials for each language in which the stimulus words depicting similar sounds were presented together, while in the current study 43 Russian inanimate onomatopoeias were tested one by one. It needs to be mentioned that the essence of imitative sound symbolism lies in the ability of individual words to demonstrate form-meaning correspondences. It may be the case that in the study of Leskinen (1998), the presentation of several words together has disturbed the association processes that might have otherwise occurred.

A major factor affecting the performance in our study pertains to the cross-linguistic similarities. Even though there seem to be very few tokens among Russian onomatopoeias that rather apparently resemble their Finnish counterparts (Russian *apčkhi* “sneezing”, *čukh-čukh* “train”, *f'ju-f'ju* “flying”, and *tik-tak* “clock” are *atsii*, *tsuku-tsuku*, *viuh*, and *tik-tak* in Finnish, respectively), a closer examination of the stimuli and the associations generated by the participants indicated cases of more subtle cross-linguistic correspondences in sound-meaning mapping.

In the Facilitating cluster, in which the responses coincided with the target meaning of Russian stimuli, for several of the tokens, some of the associations given in Finnish belonged to general vocabulary and were phonetically similar to Russian onomatopoeias. For instance, Rus. *khljup* “splash”, and *bul'-bul'* “glug” accounted for such expressions as Fin. *vesi pulppuaa* “water gushes/glugs”, *pulppuaminen* “gushing/glugging”, which are phonetically similar to Fin. *plup* or *pulp* “gushing/glugging”. Another example of this would be Rus. *khrust'* “cracking” which was associated with Fin. *rapina/rutina* “rustle-N”, *kuiva(t) lehdet/leipä* “dry leaves/bread”, in association with *raps* or *ruti-ruti* both referring to a crackling sound.

While there were several associations in the Facilitating cluster that shared the meaning with the Russian counterpart, in the Counteracting cluster there were several associations that exhibited cross-linguistic similarities in the way these words sounded but resulted in wrong answers. For instance, Rus. *kap* “water dripping” was associated with Fin. *kop* “knocking” as the respondents gave such derivatives from general vocabulary as *kopina, kopahdus* and *koputus*, nouns referring to knocking. Similarly, Fin. *tsirp* “chirping of birds” could have been the reason why the respondents wrote Fin. *lintu* “bird” in association with both Rus. *cvirk* “water trickling” and *cirk* “match strike”.

Many interesting parallels could be found in the Mixed cluster. For instance, on one hand, Rus. *skrip* “squeak” was associated with Fin. *kriik* “squeak” that has similar semantics, as there were such answers as Fin. *joku jarruttaa tiukasti* “someone is breaking strongly”, *oven saranan narina* “door hinge creaks”, and *liitu liitutaululla* “chalk on chalkboard”. On the other hand, there were such answers as Fin. *raapaisu* and *repeäminen,* nouns referring to scratching and ripping, and *raapia* “to scratch-V”, which all hint on the association with Fin. *raps* that depicts crackling and scratching, differing from the target meaning in Russian.

Finally, as there were no clear patterns in the associations given for the Undefined cluster, we cannot draw conclusions about such similarities between the languages. At any rate, the cross-linguistic similarities for the correct associations are likely a factor for higher performance in the Facilitating and Mixed clusters.

It is worth noticing that the very existence of the Counteracting and Undefined clusters supports the findings from an experimental study by Viimaranta et al. (2016) which aimed at finding overlaps in onomatopoeic vocabulary between Finnish and Russian. It showed that speakers of two different languages, e.g., Finnish and Russian, not only name the same acoustic events using phonetically different words, but, presumably, even perceive these extra-linguistic sounds in a different way.

Even though our results show that there are similarities between Russian and Finnish that drive the process of semantic recognition of unknown onomatopoeic words, the present study cannot pinpoint the exact source of such similarities. On one hand, a prolonged language contact between Russian and Finnish could have introduced some cross-linguistic correspondences in sound-meaning mapping. On the other hand, linguistic sounds (i.e., phonemes), available in any particular language, may be used to compose onomatopoeic words that are quite successful in depicting natural sounds. Despite strong evidence that onomatopoeias vary across languages, universal elements are likely to be present in each language. Sound symbolism may offer a viable explanation to the factors that drive the process of semantic recognition of unknown words by naïve listeners. Our findings broadly support the assumption that onomatopoeias contain phonemes in such combinations that the sounds produced by referents can be successfully unveiled to naïve listeners.

In the Experiment 2 of this study, the participants recognized the target images of 47 stimulus words with above chance likelihood. It shows that they managed to integrate the two types of input and establish the sound-referent pairing. In more detail, when adult naïve listeners heard the onomatopoeic words and saw the objects those words depict in the presence of three distracting images, they identified the target accurately enough. This indicates that the sound symbolism bootstrapping hypothesis (Imai & Kita, 2014) seems to be relevant for adults with no prior knowledge of the tested language. It suggests that imitative sound symbolism offers a scaffolding material for integrating audio and visual inputs and connecting words to their referents. It is likely that in more natural conditions, when speech sounds are presented in connection with events or visual imagery, imitative sound symbolism characteristic of onomatopoeias may help establish referential insights for connecting unknown words with their meaning. Finally, the results of the eye-tracking experiment show that other cues, except the phonology of words itself, contribute to unveiling the meaning of unknown onomatopoeias.

Related to our finding, Thompson and Do (2019) used short video clips in English to study whether native speakers of English could unambiguously interpret arbitrary speech depicting speaker's cognitive states and emotional reactions (e.g., Eng. *I'm gonna fight* referring to being determined), onomatopoeias with conventionalized meaning (Eng. *ah-ha-ha* "laughter", *doot-doot-doodle-duh-doo* "carnival music") and phonologically anomalous unconventional sound imitations (e.g. *prfff* "something rushing by", *fšw* "something flying over the speaker's head"), depending on whether additional information coming from hand gestures, facial expression, intonation and context was available or not. Relevant to our results, unconventional sound imitations strongly relied on co-speech iconic hand gestures or contextual information uttered by the speaker in order to be understood correctly. This suggests that when the listeners are unfamiliar with the meaning of sound imitations, additional information substantially aids accessing their semantics. The results of our study are in line with this finding by Thompson and Do (2019). Indeed, naïve listeners, the participants of the eye-tracking experiment, successfully grasped the semantics of 47 out of 50 onomatopoeic words, when congruent visual imagery was available, despite the presence of three distracting images.

To sum up, sound-meaning mapping is instrumental in establishing referentiality when unknown onomatopoeias are presented in isolation. The availability of both audio and visual inputs substantially boosts this process. Furthermore, the present study demonstrates that eye tracking can be used to study the processing of onomatopoeia in a detailed way. The future research should use a wider range of languages to spot universal similarities involved in establishing sound-meaning mapping in onomatopoeia.

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

Altmann, G. T. M., & Kamide, Y. (2007). The real-time mediation of visual attention by language and world knowledge: Linking anticipatory (and other) eye movements to linguistic processing. *Journal of Memory and Language*, *57*(4), 502−518. https://doi.org/10.1016/j.jml.2006.12.004

Berlin, B., & O'Neill, J. P. (1981). The pervasiveness of onomatopoeia in Aguaruna and Huambisa bird names. *Journal of Ethnobiology, 1*(2), 238–61.

Brackbill, Y., & Little, K. B. (1957). Factors determining the guessing of meanings of foreign words. *Journal of Abnormal and Social Psychology*, *54*(3), 312–318. https://doi.org/10.1037/h0042411

Brown, R. W., Black, A. H., & Horowitz, A. E. (1955). Phonetic symbolism in natural languages. *Journal of Abnormal and Social Psychology, 50*(3), 388–393. https://doi.org/10.1037/h0046820

Childs, G. T. (2015). Sound symbolism. In J. R. Taylor (Ed.), *The Oxford handbook of the word* (Vol. 1). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199641604.013.030

Cuskley, C., & Kirby, S. (2013). Synesthesia, cross-modality, and language evolution. In J. Simner & E. M. Hubbard (Eds.), *The Oxford handbook of synesthesia* (pp. 869–899). Oxford University Press.

Deese, J. (1965). *The structure of associations in language and thought.* Baltimore, MD: Johns Hopkins Press.

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., Schuerman, W., Reinisch, E., Tufvesson, S., & Mitterer, H. (2016). What sound symbolism can and cannot do: Testing the iconicity of ideophones from five languages. *Language (Baltimore)*, *92*(2), e117–e133. https://doi.org/10.1353/lan.2016.0034

Efremova, T. F., (2000). *Novyj Slovar' Russkogo jazyka: Tolkovo-slovoobrazovatel'nyj [The new Russian dictionary: Explanatory-derivational]*. Moscow, Russia: Russkij jazyk.

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

Fitzpatrick, T., & Izura, C. (2011). Word association in L1 and L2: An exploratory study of response types, response times, and interlingual mediation. *Studies in Second Language Acquisition*, *33*(3), 373–398. https://doi.org/10.1017/S0272263111000027

Hinton, L., Ohala, J. J., & Nichols, J. (1994). *Sound symbolism*. Cambridge: Cambridge University Press.

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

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

Imai, M., Kita, S., Nagumo, M., & Okada, H. (2008). Sound symbolism facilitates early verb learning. *Cognition*, *109*(1), 54–65. https://doi.org/10.1016/j.cognition.2008.07.015

Iwasaki, N, Vinson, D. P., & Vigliocco, G. (2007). What do English speakers know about gera-gera and yota-yota?: A cross-linguistic investigation of mimetic words for laughing and walking. *Japanese-Language Education around the Globe, 17*, 53–78.

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. (2019). On differences between homonymy and polysemy of Russian onomatopoeic verbal interjections. *Die Welt der Slaven, 64*(2), 252–261.

Kanerva, O. A. (2020). Russian onomatopoeic verbal interjections. Why use ‘non-words’ instead of ordinary ones? *RSUH/RGGU Bulletin. ‘Literary Theory. Linguistics. Cultural Studies’ Series* 7, 130-146. https://doi.org/10.28995/2686-7249-2020-7-130-146

Kassambra, A., & Mundt, F. (2019). *factoextra: Extract and visualize the results of multivariate data analyses*. R package version 1.0.6. https://CRAN.R-project.org/package=factoextra

Kiss, G. R. (1975). An associative thesaurus of English: Structural analysis of a large relevance network. In: A. Kennedy & A. Wilkes (Eds.), *Studies in long term memory* (pp. 103-121). London: Wiley.

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

Leskinen, J. (1998). Miten vieraskielisiä imitatiiveja tunnistetaan? [How are foreign language imitatives recognised?]. *Oekeeta asijoo. Commentationes Fenno-Ugricae in honorem Seppo Suhonen sexagenarii 16.5.1998*, 310–315. Helsinki: Suomalais-Ugrilainen Seura.

Lyons, J. (1968). *Introduction to theoretical linguistics*. Cambridge, UK: Cambridge University Press.

Maechler, M., Rousseeuw, P., Struyf, A., Hubert, M., & Hornik, K. (2019). *cluster: Cluster analysis basics and extensions*. R package version 2.1.0.

Maltzman, I., Morrisett, L., & Brooks, L. O. (1956). An investigation of phonetic symbolism. *Journal of Abnormal and Social Psychology*, *53*(2), 249–251. https://doi.org/10.1037/h0048406

McRae, K., Cree, G., Seidenberg, M., & Mcnorgan, C. (2005). Semantic feature production norms for a large set of living and nonliving things. *Behavior Research Methods*, *37*(4), 547–559. https://doi.org/10.3758/BF03192726

Meara, P. M. (2009). *Connected words: Word associations and second language vocabulary acquisition*. Amsterdam: John Benjamins Pub. Co.

Motamedi, Y., Little, H., Nielsen, A., & Sulik, J. (2019). The iconicity toolbox: Empirical approaches to measuring iconicity. *Language and Cognition*, 11(2), 188-207. https://doi.org/10.1017/langcog.2019.14

Nielsen, A. K., & Dingemanse, M. (2021). Iconicity in word learning and beyond. A critical review. *Language and Speech*, *64*(1), 52–72. https://doi.org/10.1177/0023830920914339

Oswalt, R. L. (1994). “Inanimate imitatives in English”. In L. Hinton, J. Nichols, & J. J. Ohala (Eds.), *Sound symbolism* (pp. 293-308). New York: Cambridge University Press. https://doi.org/10.1017/CBO9780511751806.020

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, Retrieved from https://CRAN.R-project.org/package=VWPre.

R Core Team (2020). *R: A language and environment for statistical computing.* R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.

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

Sapir, E. (1970). *Language: An introduction to the study of speech*. London: Rupert Hart-Davis.

Šljakhova, S.S. (2004). *Drebezgi jazyka: Slovar' russkikh fonosemantičeskikh anomalij [Shatters of language: Dictionary of Russian phono-semantically anomalous words]*. Perm′, Russia: Permskij gosudarstvennyj pedagogogičeskij universitet.

Thompson, A. L., & Do, Y. (2019). Unconventional spoken iconicity follows a conventional structure: Evidence from demonstrations. *Speech Communication*, *113*, 36–46. https://doi.org/10.1016/j.specom.2019.08.002

Tsuru, S., & Fries, H. S. (1933). A problem in meaning. *The Journal of General Psychology*, *8*(1), 281–284. https://doi.org/10.1080/00221309.1933.9713186

van Rij, J., Wieling, M., Baayen, R., & van Rijn, H. (2017). *itsadug: Interpreting Time Series and Autocorrelated Data Using GAMMs.* R package version 2.3. Retrieved from https://cran.r-project.org/package=itsadug

Viimaranta, J., Aleksandrova, A. A., Bogomolov, A. S., Pasmor, E. S., & Tikkanen, L. (2016). "Vodnye" vukopodražanija v finskom i russkom jazykakh [Interjections pertaining to sounds of water in Finnish and Russian]. *Jazyk i kul'tura, 1*(33), 6–24. https://doi.org/10.17223/19996195/33/1

Vinson, D. P., & Vigliocco, G. (2008). Semantic feature production norms for a large set of objects and events. *Behavior Research Methods, 40*, 183–190. https://doi.org/10.3758/BRM.40.1.183

Voeltz, F. K. E., & Kilian-Hatz, C. (2001). *Ideophones.* Amsterdam: John Benjamins Publishing.

Weiss, J. H. (1963). Role of "meaningfulness" versus meaning dimensions in guessing the meanings of foreign words. *Journal of Abnormal and Social Psychology*, *66*(6), 541–546. https://doi.org/10.1037/h0044085

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

Table 1.

The means and standard deviations (in the parentheses) of the types of responses for the 23 participants for the words in each cluster.

Response type	Cluster			
	Facilitating	Counteracting	Mixed	Undefined
Matching	17.93 (3.00)	2.61 (3.01)	8.78 (1.72)	2.14 (1.29)
Mismatching	0.00 (0.00)	14.69 (3.40)	3.78 (2.54)	3.50 (3.20)
Random	5.07 (3.00)	5.69 (2.43)	10.44 (2.88)	17.36 (2.82)

Table 2.

Generalized additive mixed model for the empirical logit of looks to the target picture as a function of time by cluster. The upper part reports the parametric coefficients and the lower part the smooths and random effects. For the parametric coefficients, the Facilitating cluster is the baseline.

Parametric coefficients	Estimate	SE	t value	p value
Intercept	1.5447	0.2291	6.743	< .0001
ClusterMixed	-0.3300	0.3357	-0.983	.3256
ClusterCounteracting	-1.0412	0.3021	-3.447	.0006
ClusterUndefined	-1.4895	0.2935	-5.075	< .0001
Smooth terms	**Edf**	**Ref. df**	**F value**	**p value**
Smooth: Time, ClusterFacilitating	5.163	5.623	16.316	< .0001
Smooth: Time, ClusterMixed	6.106	6.605	10.688	< .0001
Smooth: Time, ClusterCounteracting	5.479	5.957	6.619	< .0001
Smooth: Time, ClusterUndefined	3.575	3.883	5.629	.0002
Random effect: Time, subject	189.108	242.000	307.207	< .0001
Random effect: Time, item	357.338	446.000	386.737	< .0001

Random effect: Event	1263.749	1346.000	37.891	< .0001

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

Figure captions

Figure 1. The dendrogram for Facilitating/Counteracting/Mixed/Undefined clusters. The numbers correspond with the item numbers of Appendix A.

Figure 2. An example of a stimulus screen. Upper left (A) is the target picture for the sound *ščelk* that depicts mechanical clicking of a lock, for instance, upper right (B) is the visually similar competitor, lower left (C) is the semantic competitor, and lower right (D) is the control distractor. Pictures are taken from the Papunet picture bank, papunet.net, Sergio Palao/ARASAAC (A, B, and D), Paxtoncrafts Charitable Trust (C); B and D are edited versions of the original pictures. Pictures are shared with the Creative Commons BY-NC-SA licence.

Figure 3. The fixation proportions to the target for the four clusters as a function of time. The vertical line denotes the chance level.

Figure 4. The estimated differences of looks to target between the four clusters as a function of time. The statistically significant areas are within the two dotted lines.

Appendix A

The meanings of onomatopoeias, the elicitation test results, and the clusters emerging in the cluster analysis are reported in Table A1. The tokens are given in the same order as in the elicitation test.

[Insert Table A1 about here.]

Table A1.

Meanings of onomatopoeias, the elicitation test results (answers), and the clusters (Cluster 1 = Facilitating, Cluster 2 = Counteracting, Cluster 3 = Mixed, Cluster 4 = Undefined).

Token	Source of sound/nature of event	Answers			Cluster
		Correct	Incorrect	Missing	
1. apčkhi	sneezing	21	1	1	1
2. bultykh	hard or heavy object falling into water	2	9	12	4
3. svis'	whistling, also an object flying through the air or items rubbing against one another	14	9	0	1
4. ljap	hitting, slapping, also something muddy, pasty or wet falling down.	16	4	3	1
5. šarakh	unexpected strong hit, fast throw, shot	4	11	8	4
6. bac	sharp and short sound of hitting or shooting	8	13	2	2
7. grokh	falling or hitting the ground with noise	0	21	2	2
8. bren'	string musical instruments	20	3	0	1
9. trakh	strong and sharp cracking or rumbling	1	15	7	4

10. čukh-čukh	train	16	7	0	1
11. tres'	hitting, bursting or splitting into parts, for example, of something wooden	3	6	14	4
12. brjak	hitting metallic or glass objects against one another	8	6	9	3
13. f'ju- f'ju	objects flying past, also whistling	20	3	0	1
14. čeburakh	strong hitting, falling clumsily	1	7	15	4
15. din'-din'	small bell, glass clinking	19	3	1	1
16. čmok	kissing	3	15	5	4
17. tuk-tuk	sharp hits or knocking on wood	3	18	2	2
18. khlobys'	unexpected hitting, falling, also a massive volume of liquid being poured	8	5	10	3
19. khlop	unexpected falling, hitting, bursting, shooting or throwing	7	13	3	2
20. pljukh	falling or throwing something flat, soft or liquid	7	9	7	3
21. skrip	friction or compression	11	9	3	3
22. bakh	low frequency, dull and prompt sound of hitting, shooting or exploding	6	12	5	3
23. don	big bell other metallic objects	18	3	2	1
24. šmjak	hitting, falling with a flopping noise	10	8	5	3
25. čikh	sneezing	18	5	0	1
26. zvjak	metallic or glass objects hitting against one another	2	11	10	4
27. ukh	dull hitting, falling, exploding or cannon fire	0	20	3	2

28. šlep	dull unclear sound of falling, slapping or walking clumsily, in a lazy manner or on mud	18	4	1	1
29. bul'-bul'	liquid pouring from a vessel, an object falling into water	23	0	0	1
30. bukh	dull abrupt sound of hitting, exploding or falling	1	20	2	2
31. tik-tak	clock ticking	21	2	0	1
32. khlest'	whipping	1	13	9	4
33. om-om	eating or gulping	11	10	2	3
34. kap	water dripping	0	19	4	2
35. cvirk	water trickling, also a thin sheet of water pouring	0	23	0	2
36. bom	big bell ringing	7	14	2	2
37. khrust'	something fragile cracking or breaking	13	7	3	1
38. ščelk	mechanical clicking, also cracking of a nutshell, snapping with fingers	2	8	13	4
39. cok-cok	clatter of hoofs or metallic heel tips	0	20	3	2
40. porkh	wings flapping	1	7	15	4
41. khljup	splashing, a munching sound of something liquid or muddy	14	8	1	1
42. vžik	buzzing or other sharp high pitch sound	5	5	13	4
43. pykh	steam or smoke being let out, also something burning down	9	6	8	3
44. am	eating or devouring	4	19	0	2
45. švark	unexpected, sharp, strong hitting or throwing	3	14	6	4
46. tu-tu	train's horn	9	13	1	3

47. čirk	a match strike or rubbing against something dry or wooden, knife cut	1	19	3	2
48. top-top	stomping, walking	3	17	3	2
49. pif-paf	gun shoot	1	11	11	4
50. šu-šu	human whisper	1	13	9	4

Appendix B

The generalized additive mixed models for the visual and semantic competitor effects are presented in Tables B1 and B2, respectively.

[Insert Table B1 about here.]

Table B1.

Generalized additive mixed model for the difference of empirical logit of looks to the visually similar picture minus the control distractor as a function of time by cluster. The upper part reports the parametric coefficients and the lower part the smooths and random effects. For the parametric coefficients, the Facilitating cluster is the baseline.

Parametric coefficients	Estimate	SE	t value	p value
Intercept	0.2287	0.1557	1.469	0.142
ClusterMixed	-0.1961	0.2420	-0.810	0.418
ClusterCounteracting	-0.1115	0.2183	-0.511	0.609
ClusterUndefined	0.3036	0.2603	1.166	0.244
Smooth terms	**Edf**	**Ref. df**	**F value**	**p value**
Smooth: Time, ClusterFacilitating	1.001	1.001	0.236	.6276
Smooth: Time, ClusterMixed	1.001	1.002	0.438	.5084
Smooth: Time, ClusterCounteracting	1.001	1.002	0.018	.8955
Smooth: Time, ClusterUndefined	7.334	7.767	1.019	.2766
Random effect: Time, subject	183.018	242.000	35.286	.0047
Random effect: Time, item	356.270	446.000	76.414	< .0001
Random effect: Event	1241.555	1346.000	20.107	< .0001

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

[Insert Table B2 about here.]

Table B2.

Generalized additive mixed model for the difference of empirical logit of looks to the semantic object minus the control distractor as a function of time by cluster. The upper part reports the parametric coefficients and the lower part the smooths and random effects. For the parametric coefficients, the Facilitating cluster is the baseline.

Parametric coefficients	Estimate	SE	t value	p value
Intercept	0.30953	0.16575	1.867	.0618
ClusterMixed	-0.09805	0.24443	-0.401	.6883
ClusterCounteracting	0.08333	0.22048	0.378	.7055
ClusterUndefined	-0.16815	0.21628	-0.777	.4369
Smooth terms	Edf	Ref. df	F value	p value
Smooth: Time, ClusterFacilitating	1.001	1.001	0.013	.9110
Smooth: Time, ClusterMixed	1.001	1.001	0.098	.7542
Smooth: Time, ClusterCounteracting	1.002	1.003	0.194	.6607
Smooth: Time, ClusterUndefined	1.002	1.002	0.397	.5294
Random effect: Time, subject	194.985	242.000	68.683	.0271
Random effect: Time, item	345.025	446.000	124.863	< .0001
Random effect: Event	1250.115	1346.000	24.559	< .0001

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

Table A1.

Meanings of onomatopoeias, the elicitation test results (answers), and the clusters (Cluster 1 = Facilitating, Cluster 2 = Counteracting, Cluster 3 = Mixed, Cluster 4 = Undefined).

Token	Source of sound/nature of event	Answers			Cluster
		Correct	Incorrect	Missing	
1. apčkhi	sneezing	21	1	1	1
2. bultykh	hard or heavy object falling into water	2	9	12	4
3. svis'	whistling, also an object flying through the air or items rubbing against one another	14	9	0	1
4. ljap	hitting, slapping, also something muddy, pasty or wet falling down.	16	4	3	1
5. šarakh	unexpected strong hit, fast throw, shot	4	11	8	4
6. bac	sharp and short sound of hitting or shooting	8	13	2	2
7. grokh	falling or hitting the ground with noise	0	21	2	2
8. bren'	string musical instruments	20	3	0	1
9. trakh	strong and sharp cracking or rumbling	1	15	7	4

10. čukh-čukh	train	16	7	0	1
11. tres'	hitting, bursting or splitting into parts, for example, of something wooden	3	6	14	4
12. brjak	hitting metallic or glass objects against one another	8	6	9	3
13. f'ju- f'ju	objects flying past, also whistling	20	3	0	1
14. čeburakh	strong hitting, falling clumsily	1	7	15	4
15. din'-din'	small bell, glass clinking	19	3	1	1
16. čmok	kissing	3	15	5	4
17. tuk-tuk	sharp hits or knocking on wood	3	18	2	2
18. khlobys'	unexpected hitting, falling, also a massive volume of liquid being poured	8	5	10	3
19. khlop	unexpected falling, hitting, bursting, shooting or throwing	7	13	3	2
20. pljukh	falling or throwing something flat, soft or liquid	7	9	7	3
21. skrip	friction or compression	11	9	3	3
22. bakh	low frequency, dull and prompt sound of hitting, shooting or exploding	6	12	5	3
23. don	big bell other metallic objects	18	3	2	1

24. šmjak	hitting, falling with a flopping noise	10	8	5	3
25. čikh	sneezing	18	5	0	1
26. zvjak	metallic or glass objects hitting against one another	2	11	10	4
27. ukh	dull hitting, falling, exploding or cannon fire	0	20	3	2
28. šlep	dull unclear sound of falling, slapping or walking clumsily, in a lazy manner or on mud	18	4	1	1
29. bul'-bul'	liquid pouring from a vessel, an object falling into water	23	0	0	1
30. bukh	dull abrupt sound of hitting, exploding or falling	1	20	2	2
31. tik-tak	clock ticking	21	2	0	1
32. khlest'	whipping	1	13	9	4
33. om-om	eating or gulping	11	10	2	3
34. kap	water dripping	0	19	4	2
35. cvirk	water trickling, also a thin sheet of water pouring	0	23	0	2
36. bom	big bell ringing	7	14	2	2
37. khrust'	something fragile cracking or breaking	13	7	3	1

38. ščelk	mechanical clicking, also cracking of a nutshell, snapping with fingers	2	8	13	4
39. cok-cok	clatter of hoofs or metallic heel tips	0	20	3	2
40. porkh	wings flapping	1	7	15	4
41. khljup	splashing, a munching sound of something liquid or muddy	14	8	1	1
42. vžik	buzzing or other sharp high pitch sound	5	5	13	4
43. pykh	steam or smoke being let out, also something burning down	9	6	8	3
44. am	eating or devouring	4	19	0	2
45. švark	unexpected, sharp, strong hitting or throwing	3	14	6	4
46. tu-tu	train's horn	9	13	1	3
47. čirk	a match strike or rubbing against something dry or wooden, knife cut	1	19	3	2
48. top-top	stomping, walking	3	17	3	2
49. pif-paf	gun shoot	1	11	11	4
50. šu-šu	human whisper	1	13	9	4

Table B1.

Generalized additive mixed model for the difference of empirical logit of looks to the visually similar picture minus the control distractor as a function of time by cluster. The upper part reports the parametric coefficients and the lower part the smooths and random effects. For the parametric coefficients, the Facilitating cluster is the baseline.

Parametric coefficients	Estimate	SE	t value	p value
Intercept	0.2287	0.1557	1.469	0.142
ClusterMixed	-0.1961	0.2420	-0.810	0.418
ClusterCounteracting	-0.1115	0.2183	-0.511	0.609
ClusterUndefined	0.3036	0.2603	1.166	0.244
Smooth terms	**Edf**	**Ref. df**	**F value**	**p value**
Smooth: Time, ClusterFacilitating	1.001	1.001	0.236	.6276
Smooth: Time, ClusterMixed	1.001	1.002	0.438	.5084
Smooth: Time, ClusterCounteracting	1.001	1.002	0.018	.8955
Smooth: Time, ClusterUndefined	7.334	7.767	1.019	.2766
Random effect: Time, subject	183.018	242.000	35.286	.0047
Random effect: Time, item	356.270	446.000	76.414	< .0001

Random effect: Event	1241.555	1346.000	20.107	< .0001

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

Table B2.

Generalized additive mixed model for the difference of empirical logit of looks to the semantic object minus the control distractor as a function of time by cluster. The upper part reports the parametric coefficients and the lower part the smooths and random effects. For the parametric coefficients, the Facilitating cluster is the baseline.

Parametric coefficients	Estimate	SE	t value	p value
Intercept	0.30953	0.16575	1.867	.0618
ClusterMixed	-0.09805	0.24443	-0.401	.6883
ClusterCounteracting	0.08333	0.22048	0.378	.7055
ClusterUndefined	-0.16815	0.21628	-0.777	.4369
Smooth terms	Edf	Ref. df	F value	p value
Smooth: Time, ClusterFacilitating	1.001	1.001	0.013	.9110
Smooth: Time, ClusterMixed	1.001	1.001	0.098	.7542
Smooth: Time, ClusterCounteracting	1.002	1.003	0.194	.6607
Smooth: Time, ClusterUndefined	1.002	1.002	0.397	.5294
Random effect: Time, subject	194.985	242.000	68.683	.0271
Random effect: Time, item	345.025	446.000	124.863	< .0001

Random effect: Event	1250.115	1346.000	24.559	< .0001

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