

When the Sound-Symbolism Effect Disappears: The Differential Role of Order and Timing in Presenting Visual and Auditory Stimuli

Jelena Sučević, Dragan Janković, Vanja Ković

Department of Psychology, Faculty of Philosophy, University of Belgrade, Belgrade, Yugoslavia
Email: jelena.sucevic@gmail.com

Received April 24th, 2013; revised May 26th, 2013; accepted June 23rd, 2013

Copyright © 2013 Jelena Sučević et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Köhler's observation that most people match pseudoword "maluma" to curvy objects and "takete" to spiky objects represented the well-known example of sound symbolism—the idea that link between sound and meaning of words was not entirely arbitrary. This study was aimed to examine the existence of sound symbolism in natural language and to consider the potential role of some aspects of experimental design and stimuli features which had not been considered in experimental studies so far. Three experiments were done in order to explore the influence of visual information on language processing. Visual lexical decision task with the sharp-sounding and soft-sounding verbal stimuli presented within the spiky and curvy frames was used. Reaction time analysis in these three experiments highlighted additional aspects of visual and language processing which influence the potential interplay of these two processes. As results revealed, when visual information preceded presentation of verbal material for approximately 1000 ms or when visual and verbal material were presented simultaneously, the processing was being delayed and the interactions of these two processes occurred. The pattern of obtained results gave further support to the idea of sound symbolism as pre-semantic phenomenon and the hypothesis that the effect emerged from very early stages of language processing.

Keywords: Sound Symbolism; Semantics; Words; Natural Language

Introduction

Whether sound of a word is arbitrary or non-arbitrary related to its meaning has been debated at least since Plato's Cratylus dialog in fifth century BC (Plato, 1998). This sound-meaning relation has since been much discussed in the philosophy, linguistic and psychology.

De Saussure's view on language as arbitrary system has often been considered to be the core idea of modern linguistic research approach (Saussure, 1959). According to this approach, there is no systematic relation between characteristics of particular word and object referred by it. On the contrary, certain correspondences between phonological features of words and their meanings have been claimed to exist. This idea, as proposed by linguist Eduard Sapir, became known as "phonetic symbolism" (Sapir, 1929). He claimed that the relation between sound and meaning cannot be considered as entirely arbitrary, and that these correspondences are instances of sound-symbolism, universal feature of language system.

With the well-known Köhler's observation from 1929, debate over word-object relations was set as a matter of interest not only in philosophy and linguistic, but in psychology as well. Using the forced-choice word-picture matching task Köhler determined existence of the systematic tendency to match nonsense word *takete* to the spiky object, and the nonsense word

baluma (in later research *maluma*) to curvy object (Köhler, 1929). Later on, a number of studies confirmed this finding as a robust culturally independent effect (e.g. Davis, 1961; Bremner et al., 2013).

Experimental investigations of this phenomenon have been focused primarily on the vowel content of the words or pseudowords associated with visual objects with specific characteristics like sharpness and roundness (e.g. Davis, 1961; Ramachandran & Hubbard, 2001; Maurer et al., 2006). For example, pseudowords containing rounded vowels /o/ and /u/ were more often associated with rounded shapes and pseudowords containing unrounded vowels /i/ and /e/ were more often associated with spiky shapes (Ramachandran & Hubbard, 2001; Maurer et al., 2006). However, further research showed that beside vowels, consonants and consonant-vowel patterns in the words also play a role in sound symbolism. Using picture-naming task Janković et al. found that pseudowords produced for sharp, spiky objects included significantly more plosives /k/, /t/, /g/, /d/, affricates /ts/, /dz/, and trill /r/, while pseudowords produced for rounded objects included more laterals /l/, /L/ and nasals /m/, /n/ (Janković & Marković, 2001; Janković, Vučković, & Radaković, 2005). In addition, same study showed that pseudonyms produced for spiky objects included more CC (consonant-consonant) syllables while pseudonyms produced for rounded objects included more CV (consonant-vowel) syl-

lables. Similarly, Westbury (2005) focused on consonants and found that strings containing plosive consonants were identified more quickly and accurately within spiky frames while strings containing continuant consonants were identified more quickly and accurately within curved frames.

Interesting account of the role of sound symbolism in language evolution has been provided by Ramachandran and Hubbard (2001). Within their synesthetic theory of language origins and consciousness, they interpret sound-symbolic correspondence as consequence of coactivation of the motor or somatosensory areas involved in sound articulation with the perception of differently shaped objects. According to Ramachandran and Hubbard, nature of this activation is similar as one present in synesthesia. For instance, such cross-modal correspondences are present in linkage between perception of rounded object and motor representation which is activated when person in saying vowel /o/ (Ramachandran & Hubbard, 2001).

Alternative view on the origin of sound symbolism negates neural basis of this phenomenon and assumes that learning of language occurs prior to sound-symbolism-like effects. In other words, these correspondences come from generalization of knowledge about already acquired word-object mappings to nonsense words stimulus (Rogers & Ross, 1975). Study which confronted these two theoretical accounts on sound-symbolism origin is Maurer, Pathman and Mondloch's study (2006). In this study, sound-shape correspondences are found to be present even in two and a half months old infants. According to these authors, vocabulary size at this age is not big enough to make word-object mapping generalizations possible. Furthermore, this age is considered to be period when influences among contiguous brain areas are stronger than in adults (Spector & Maurer, 2009). In line with Ramachandran's view on language evolution, Maurer suggested that these sound-shape correspondences influence individual language development, but may have influenced the evolution of language as well (Maurer et al., 2006).

Majority of research dealing with the sound-symbolism have been based on artificial material. Based on those insights, certain generalizations concerning natural language properties were made. On the other side, there has been far less research based on natural language data and their findings often were quite inconsistent (Newman, as cited in Westbury, 2005; Diefloth, 1994). However, one of those studies, in which data from 229 languages were analyzed, found certain patterns of language symbolism in majority of those languages (Ciccotosto, 1991). Furthermore, some studies dealing with the structure of words in natural language suggest that words denoting sharp and rounded objects show quite similar patterns of phoneme and consonant-vowel distributions as those found in pseudowords produced for sharp and rounded visual stimuli (Ilić, Ković, & Janković, 2012).

According to Westbury, transparency of the experimental manipulations, small number of stimuli and their artificial nature represent key features of previous studies which lead to absence of any direct sound-symbolism effect and its restriction to post-hoc analysis of phoneme-meaning regularities (Westbury, 2005). To try to surpass these problems, in his study Westbury adapted implicit interference task in which participants undertook a lexical or letter decision task with the word and pseudoword (in second experiment letters and numbers

were used) presented inside spiky or curvy frames. His main idea was that if the hypothesis of sound-symbolism is plausible, sharp-sounding words will be processed more efficiently when they are presented within the spiky frames compared to the situation when presented within curvy frames, and vice versa for soft-sounding words. The results showed that curvy shapes are facilitating the identification of all-continuant strings while interfering with the identification of all-stop strings, and vice versa, but only in case of letters, not in case of words. Based on this, Westbury assumed that the effect of sound-symbolism "happens" on the level of lexical access and claims that it has a pre-semantic nature (Westbury, 2005).

In spite of the growing body of research exploring the nature of sound-symbolism, it is still unclear whether this phenomenon should be interpreted as a natural language feature. Although Ramachandran in his theory addresses this issue, the lack of evidence in experimental studies left the issue still unsolved. In other words, as Westbury formulated this, the question of the extent to which sound symbolism may be constructed, rather than discovered, by experimenters is still opened. To answer this question, this author redefines sound-symbolism as pre-semantic phenomenon and positions it on a lower level of cognitive processing. However, key feature of word is its referring function and the idea of sound-symbolism came from this line of searching for a connection between sounding of a word and characteristics of a particular object referred to. For that reason, it seems important to consider potential role of factors such as meaning of word and its level of abstraction within the experimental paradigm. For example, in afore mentioned study of Westbury (2005) these two aspects of word haven't been systematically controlled, so that words used as stimuli refer to object of different level of abstraction (e.g. noon and nail). Given those facts, it seems necessary to reconsider Westbury's claim that sound-symbolism is pre-semantic effect and to examine a potential role of these factors in natural language processing, as well as the relation between sound and meaning of words, and its relation to the frame within which word is presented.

According to one recent study, sound symbolism effect is not only influenced by the stimuli properties, but also by the characteristics of experimental procedure (Ković & Pejović, 2012). Namely, these authors have found that sound symbolism effect occurs only when mapping from auditory to visual stimuli and not vice versa. This design-dependent aspect of sound symbolism raises the question whether certain characteristics of experimental procedure which are usually not in focus of sound symbolism studies also play important role in discovering or even diminishing potential sound symbolism effect in language processing.

The aim of present study is to investigate the sound-symbolism effect in natural language processing and the influence of the order and timing in presenting visual and auditory stimuli on this effect. Three experiments were designed in order to examine the relation between properties of label, properties of referred object and visual context in which label processing occurs. More precisely, we intend to test whether the verbal stimuli processing differs when the stimulus is presented within sound-symbolic and non-sound-symbolic visual context. The differences in processing of verbal stimuli in these two conditions can provide important insight in the role of sound sym-

bolism in natural language processing.

Experiment 1

Participants

Twenty five participants, second-year undergraduate students of psychology at the Faculty of Philosophy, University of Belgrade (all females) took part in present experiment and received course credit for their participation. All participants reported normal or corrected-to-normal vision.

Method

To examine whether there is a sound-symbolic correspondence effect in natural language, five factors were manipulated in lexical decision task: the frame shape (spiky vs. curvy), the frame typicality (typical vs. atypical), lexical category (word vs. pseudoword) and the phonological structure of the word/pseudoword (sounding spiky vs. sounding soft). Beside these factors, a potential effect of frame exposure time was analyzed as well (1000, 2000, 3000 and 4000 ms).

Stimuli

Frames

Six spiky and six curvy frames were created in order to select typical and atypical stimuli within each of these categories. In order to more closely resemble stimuli used Westbury's study (2005) white figure was placed in the center of black background (each figure fitted in 432×288 pixels rectangle, as in Westbury, 2005). Spiky frames were constructed to systematically vary in "sharpness" (the number and size of spikes were varied) and 6 curvy frames systematically varied in 'curviness' (as shown in **Figure 1**). Considering that sharpness and curviness may not be entirely objective dimensions, but under subjective influence as well, subjective experience of these dimensions was examined. In order to test whether objective criteria used to create frames and subjective experience of these dimensions were congruent, 15 participants (who did not take part in the main experiment) judged these twelve frames on 7 point scale (1 indicating curvy, 7 indicating spiky). Results indicated that objective criteria and subjective judgments were congruent and two spiky and two curvy frames were selected. One frame within each category (spiky and curvy) was selected as typical (the one judged as most spiky) and the other as atypical (judged as least spiky). Identical procedure was done within category of

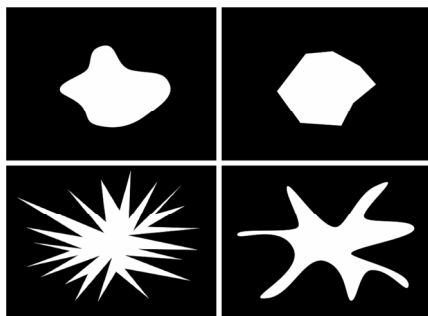


Figure 1.
The frames used in Experiments 1, 2 and 3. Typical and atypical curvy frames are presented above and typical and atypical spiky frames below.

curvy frames (frame judged as most curvy was selected as typical while least curvy frame as atypical).

Words

Stimuli were selected from the corpus of words used in Ilić, Ković and Janković (2012) and which referred to round or spiky real-objects. Only high-frequent words containing consonant-vowel-consonant-vowel-consonant (C-V-C-V-C) structure and referring to concrete objects were recruited from the corpus. In order to obtain two categories (sharp sounding and soft sounding) of the factor named *Phonological Structure*, criteria based on findings of several studies previously mentioned was used (Janković & Marković, 2001; Westbury, 2005; Ilić, Ković, & Janković, 2012). Based on those criteria, 30 words were selected (15 within each category).

Pseudowords

Total of 30 pseudowords were created so that they have the same phonological characteristics as previously selected words. Pseudowords sounding sharp were created so that in each word from sharp sounding category one consonant has been replaced with one "sharp" phoneme. Position of consonant which was replaced was balanced (5 pseudowords were created by replacing first consonant, 5 by replacing second and 5 by replacing third consonant in word) and the inserted phoneme as well ("sharp" consonants /k/, /z/, /t/, /ʃ/, /s/ and "soft" consonants /m/, /l/, /b/, /v/, /n/ were used).

Procedure

The participants were instructed to answer to the presented stimuli as quickly and accurately as possible, by pressing one of two keys on keyboard. They were instructed to place their index fingers on key "V" and "N" and to press key "V" if presented string was a word or to press key 'N' if presented string was a pseudoword. There was no explicit mention of the frames to the participants.

Visual lexical decision task presented to the participants was as follows: each trial began with the presentation of frame for a randomized interval of 1000 to 4000 ms. Then, string of letters was presented within the same frame and it disappeared immediately after participants gave answer. After the removal of stimuli, an inter-stimulus interval of 500 ms followed (as shown in **Figure 2**). Sixty letter strings (30 words and 30 pseudowords) were presented within each of four frames in random order. Thus, the task consisted of 240 trials altogether and it took participants approximately 20 minutes to complete it. Reaction time and accuracy of responses were collected.

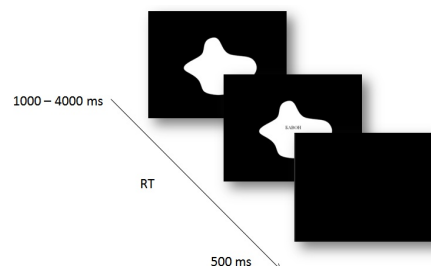


Figure 2.
Experimental procedure in Experiment 1.

Results

All participants made less than 20% errors in the task, thus none subject was excluded from further analysis (Criteria of exclusion as in Westbury, 2005). An average correct decision rate was 97% (SD = 1.84%). No significant differences in the error rate were found for Frame Shape, Frame Typicality, Phonological Structure and Frame Exposition Time, but chi-square test showed a significant difference for Lexical Category ($\chi^2(1) = 6.88$; $p < .01$), whereby more incorrect answers was given for words (124) than for pseudowords (86). Incorrect responses were excluded from the further analysis.

A $2 \times 2 \times 2 \times 2 \times 4$ Repeated Measures ANOVA of RT by subjects with factors Frame Shape (spiky vs. curvy), Frame Typicality (typical vs. atypical), Lexical Category (word vs. nonword), Phonological structure (sharp vs. curvy/soft sounding) and Frame Exposure Time (1000, 2000, 3000 and 4000 ms) revealed a significant main effect of Lexical Category ($F(1,8) = 14.52$; $p < 0.01$), whereby the words were found to be more quickly recognized than the pseudowords ($t(15) = 8.81$; $p < .01$). There was no significant main effect of the Frame Shape, Frame Typicality, Phonological Structure nor Frame Exposure Time ($p > .05$). None of the one-way or higher-order interaction effects were significant ($p > .05$).

Analyzing the response times by items, a $2 \times 2 \times 2 \times 2 \times 4$ Mixed Measures ANOVA of RT with between-subjects factors Lexical Category and Phonological structure and repeated-measures factors Frame Shape, Frame Typicality and Frame Exposure Time was done. Results revealed a significant main effect of Frame Exposure Time ($F(3, 162) = 6.69$; $p < .01$) and Lexical Category ($F(1, 54) = 12.35$; $p < .01$). As shown in **Figure 3**, the words were faster processed than the pseudowords. Different exposition time lead to differences in processing speed, whereby participants were slower in condition when frame was presented for 1000 ms prior to letter string compared to conditions where the frame was presented 2000 or more ms prior to letter string. No significant effects were found for factors Frame Shape, Frame Typicality nor Phonological Structure ($p > .05$). None of the higher-order interactions were significant ($p > .05$).

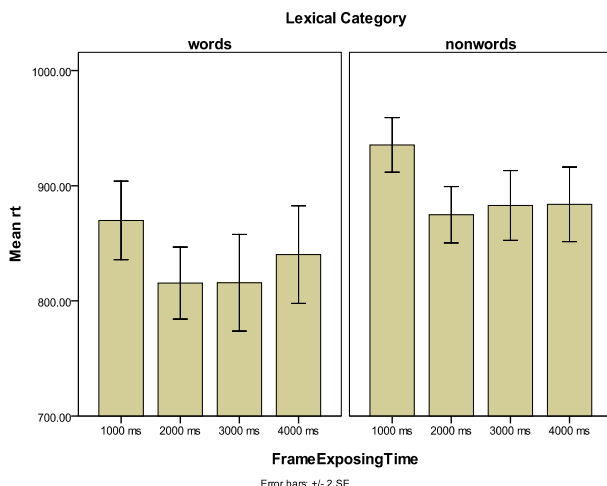


Figure 3. Reaction times for correct decisions to words and pseudowords depending on frame exposition time in Experiment 1.

Discussion

According to the results of the first experiment, only the factor *Lexical Category* showed significant effect. Faster processing of words compared to pseudowords processing is in accordance with the classical psycholinguistic studies as well as the results from the study which used the same task used in this experiment (Westbury, 2005). Beside the lexical category effect, analysis revealed a significant effect of the frame exposure time on the processing time of letter string inside the frame. Experimental design of the experiment followed the one present in the Westbury's study, so the frames were presented for 1000 to 4000 ms prior to presenting letter string which participant needs to process. To our knowledge, there are no explicit theoretical or empirical assumptions that form basis of this manipulation and expectations of its possible effects, if any exist. For that reason, factor Frame Exposure Time was included in the analysis to determine whether it may have influence on a certain characteristics of word processing. The results of this experiment indicate that this aspect of task design had a significant effect on processing time, both for words and pseudowords, whereby in situation when visual information is presented 1000 ms prior to presentation of letter stimuli, processing of this stimuli is slower compared to situation when frame is presented 2000, 3000 or 4000 ms prior to stimuli.

Although effects speaking in favor of sound symbolism correspondence were not found, the lack of potential effects may be due to inadequate timing of frame presentation. It could have lead to sequential processing of the frame and the letter strings, where effect of frame processing faded before letter string was presented. Although string was presented within the frame, it is possible that some sort of habituation on frame happened when letter string was presented, especially if we have in mind that the effects of priming (and the task is quite similar to those within the priming paradigm) are very sensitive to the variations in timing.

Experiment 2

Experiment 2 was conducted in order to examine whether interactions of the frame and the word/pseudoword processing exist in case when presentation of frame shortly precedes string presentation, as indicated in experiment 1. Experimental design and the procedure were identical as in experiment 1, except for the timing of frame exposition which was 1000 milliseconds.

Participants

Twenty three participants (5 males) participated in this experiment. All participants were second-year undergraduate students of psychology at the Faculty of Philosophy, University of Belgrade and received course credit for their participation. All participants reported normal or corrected-to-normal vision.

Stimuli

All stimuli were identical as in experiment 1. There were 30 words (15 sounding sharp and 15 sounding soft) and 30 corresponding pseudowords, always presented within four frames (typical and atypical spiky frame and typical and atypical curvy frame).

Procedure

Experimental design of experiment 2 differed from the one in previous experiment only regarding the duration of frame presentation. In the experiment 1 prior to presenting string of letter within the frame, frame was exposed 1000 to 4000 ms. In this experiment exposure time of the frames was 1000 ms with ± 200 ms of jitter. Then, string of letters appeared within the frame and the participant gave an answer whether it was word or pseudoword by pressing one of two keys on keyboard. The rest of procedure was identical as in the experiment 1. It took approximately 10 to 15 minutes to complete the task. Reaction time and accuracy of responses were collected.

Results

Participants made less than 20% errors in the task, thus none subject was excluded from further analysis. An average correct decision rate was 95% (SD = 2.27%). No significant differences in the error rate were found for Frame Shape, Frame Typicality, Frame Exposition Time and Lexical Category. There was a significant difference in number of errors for factor Phonological Structure. Incorrect answers were more frequent for sharp sounding (167) than for soft sounding words and pseudowords (133), ($\chi^2(1)=3.85$; $p = .05$). Incorrect responses were excluded from the further analysis.

Analysis by subjects of the $2 \times 2 \times 2 \times 2$ Repeated Measures ANOVA of reaction times with factors Frame Shape (spiky and curvy), Frame Typicality (typical and atypical), Lexical Category (word and pseudoword) and Phonological Structure (sharp and curvy/soft sounding = for words also meaning) revealed a significant main effects of the Phonological Structure ($F(1,22) = 15.54$; $p < .01$) and Lexical Category ($F(1,22) = 83.35$; $p < .01$). According to these results, verbal stimuli which sound softly are being processed faster in comparison to strings which sound sharply ($t(183) = 3.75$; $p < .01$) and words are being processed more quickly compared to pseudowords ($t(183) = 13.01$; $p < .01$). There were no significant main effects of the Frame Shape and the Frame Typicality ($p > .05$).

The following interaction effects were significant: 3-way interaction Frame Shape \times Frame Typicality \times Lexical Category ($F(1, 22) = 5.79$; $p < .05$) and four-way interaction Frame Shape \times Frame Typicality \times Phonological Structure \times Lexical Category ($F(1, 22) = 9.76$; $p < .01$) (as shown in **Figure 4**). Follow-up tests showed that sharp words are being processed faster when presented within typical curvy frame compared to when presented in atypical curvy frame ($t(22) = 2.57$; $p < .05$) and that sharp-sounding pseudowords are being faster processed when presented inside typical spiky frame compared to when being presented in atypical spiky frame ($t(22) = 2.79$; $p < .05$).

Analyzing the response times by items, a $2 \times 2 \times 2 \times 2$ Mixed Measures ANOVA of reaction times with between-subjects factors Lexical Category and Phonological structure and repeated-measures factors Frame Shape and Frame Typicality was done. Results revealed a significant main effect of Lexical Category ($F(1, 56) = 39.75$; $p < .01$), whereas neither Phonological Structure, Frame Shape nor Frame Typicality factor showed significant main effect ($p > .05$). The 2-way interaction Frame Typicality \times Lexical Category was significant ($F(1, 56) = 7.42$; $p < .01$) and the 3-way interaction Frame Typicality \times Lexical Category \times Phonological Structure was

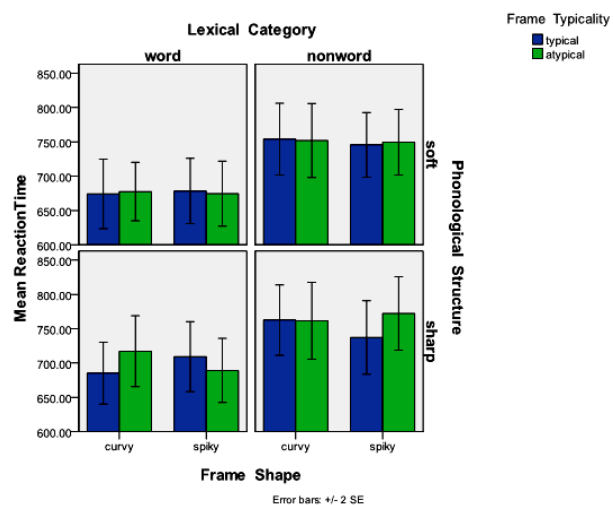


Figure 4. Reaction time for correct decisions in Experiment 2.

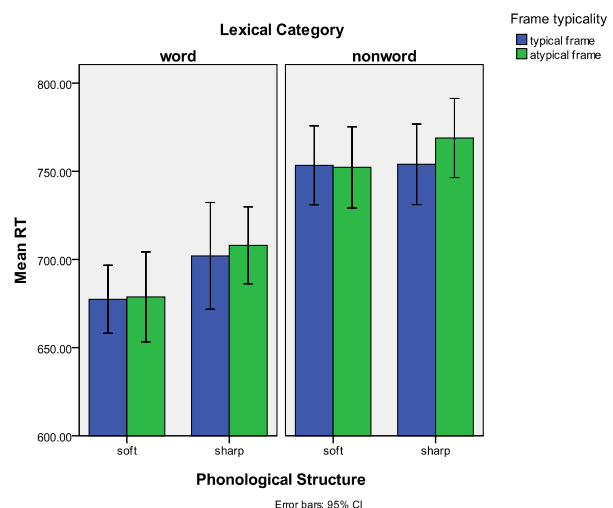


Figure 5. Reaction time for correct decisions to words and pseudowords in typical and atypical frames in Experiment 2.

near significant ($F(1, 56) = 3.35$; $p = .073$). Follow-up tests showed that sharp-sounding pseudowords are being faster processed when presented inside typical frames compared to when presented inside atypical frames ($t(14) = 2.39$; $p < .05$), as shown in **Figure 5**.

Discussion

As shown in the experiment 2, when processing of the visual information only slightly precedes presentations of word/pseudoword, beside the effect of lexical category (i.e. faster processing of words than pseudowords), some additional effects, which were not found in experiment 1, arose. In other words, when frame presentation precedes string presentation for 1000 ms (in comparison to 1000 - 4000 ms used in the first study), words and pseudowords which have “soft” phonological structure are being processed faster than those having “sharp” phonological structure. More importantly, higher-order interactions

which appeared significant indicated that there are certain differences in processing words and pseudowords depending on whether they are presented in curvy or spiky frame, and whether the frame was typical or atypical. The pattern of obtained interactions still does not give us a clear picture of influence of visual information processing on word and pseudoword processing, and whether these interactions can be interpreted as products of mechanisms functioning on the principle of sound symbolism. However, it gives us insight in important aspects of stimuli which also have their role in the potential interplay of visual and lexical information processing, which have not been considered so far.

Experiment 3

In the first and second experiment the frame was always presented to the participants prior to the presentation of verbal stimuli, thus the early stages of visual information processing were already done prior to the presentation of the verbal stimuli, especially in the first experiment.

Recent studies of sound symbolism indicate that one of the important factors which influence sound-symbolic effects is temporal sequence of the stimuli (Ković & Pejović, 2012). Having this and the results of previous two experiments in mind, this experiment was conducted in order to examine the influence of visual information on processing of verbal stimuli when visual and verbal stimuli are simultaneously presented.

Participants

Twenty participants (4 males) participated in this experiment. All participants were students at the University of Belgrade.

Stimuli

All stimuli were identical as in previous two experiments. There were 30 words (15 sounding sharp and 15 sounding soft) and 30 corresponding pseudowords. Verbal stimuli were presented within four frames (typical and atypical spiky frame and typical and atypical curvy frame).

Procedure

Experimental design of experiment 3 differed from those in previous experiments regarding the timing of frame presentation. In this experiment, presentation of the frame did not precede verbal stimuli presentation. The frame and the verbal stimuli were presented to the participant at the same time. String of letters appeared within the frame and the participant answered whether it was word or pseudoword by pressing one of two keys on keyboard. It took approximately 10 minutes to complete the task. Reaction time and accuracy of responses were collected.

Results

None subject was excluded from further analysis, since all subjects made less than 20% errors in the task. An average correct decision rate was 94% (SD = 2.23%). No significant differences in the error rate were found for Frame Shape, Frame Typicality, Frame Exposition Time and Lexical Category. There was a significant difference in number of errors for factor Lexical category ($\chi^2(1) = 5.39$; $p < .05$). Incorrect answers were

more frequent for words (153) than for pseudowords (115). Incorrect responses were excluded from the further analysis.

The $2 \times 2 \times 2 \times 2$ Repeated Measures ANOVA of reaction times by subjects with factors Frame Shape (spiky and curvy), Frame Typicality (typical and atypical), Lexical Category (word and pseudoword) and Phonological Structure (sharp and soft) revealed a significant main effects of the Lexical Category ($F(1, 19) = 36.43$; $p < .01$) and Frame Typicality ($F(1, 19) = 7.25$; $p < .05$), whereas no significant main effects of the Frame Shape nor Phonological Structure ($p > .05$). As results revealed, Frame Shape X Phonological Structure interaction was significant ($F(1, 19) = 5.47$; $p < .05$), (as shown in **Figure 6**). Follow-up comparisons showed that soft sounding verbal stimuli (both words and pseudowords) are being processed faster when presented within spiky frame compared to when presented within curvy frame ($t(22) = 2.12$; $p < .05$).

Analyzing the response times by items, a $2 \times 2 \times 2 \times 2$ Mixed Measures ANOVA of reaction times with between-subjects factors Lexical Category and Phonological structure and repeated-measures factors Frame Shape and Frame Typicality was done. Results revealed a significant main effect of Lexical Category ($F(1, 56) = 15.70$; $p < .01$), while Frame Typicality effect was near significant ($F(1, 19) = 3.84$; $p = .055$). There were no significant effects of Frame Shape, Phonological Structure nor Frame Typicality ($p > .05$).

Discussion

This experiment was conducted in order to further examine whether characteristics of visual stimuli influence verbal processing as the possible result of sound-symbolic correspondences. The frame and the letter string were presented to the participant at the same time, which lead to simultaneous visual and verbal processing. The results revealed a significant interaction of frame shape and phonological structure. However, the pattern of obtained interaction was reversed compared to the one expected if sound-symbolic hypothesis is relevant for both words and pseudowords. Verbal stimuli which sound soft are being processed more efficiently when presented within spiky frames compared to when presented within curvy frames. These

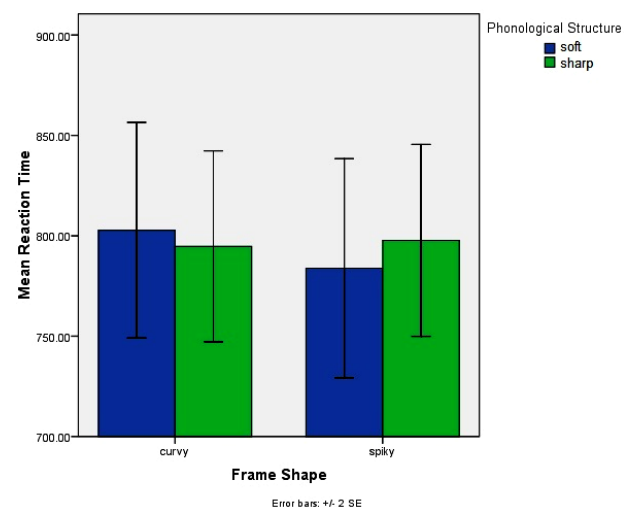


Figure 6. Reaction time for correct decisions to soft and sharp sounding verbal stimuli in curvy and spiky frames in Experiment 3.

results contradict to those presented in the Westbury's study, where soft sounding letters are being more efficiently processed when presented within curvy frames and sharp sounding letters when presented within spiky frames (Westbury, 2005). The reversed pattern of interaction obtained in case of letter string processing reveals a question whether there are some additional factors which influence verbal processing when more than letter is presented, having no influence on isolated letters processing.

General Discussion

Lack of experimental studies dealing with sound symbolism and especially those kinds of "on-line" behavioural measures intended to capture language processing while it happens, motivated this study to try to consider, within the experimental design, some important features of words that were usually being neglected within this research approach.

Another novelty of the study and important aspect of task design process was creation of the frames and selection of those which will be used in the study. Beside the criteria of spikiness/roundness, we included the criteria of typicality in frame selection as well. As results revealed, this was important factor which also influenced processing of words.

The results of the first experiment indicate that the timing of the frame exposition (prior to the presentation of the word or pseudoword inside the frame) is also an important factor which influences the reaction time measures. Although there were no clear indications of the nature of this factor's influence, the results showed that in the case of larger exposition time, the potential effect of frame diminishes or disappears, while in case of shorter frame exposition the processing time was delayed. In this situation, the processing of words and pseudowords is influenced by visual information features—frame shape and frame typicality, whereby the effect of typicality differed for spiky and for curvy frames, as well as for the phonological characteristics within each lexical category. The observed interactions do not provide a clear insight in the way that visual information processing influences processing of verbal material and whether these interactions may be due to sound-symbolic correspondences inherent to nature of language processing. Having in mind that speed of processing differs for words and pseudo-words, and that words are processed faster, it is possible that certain set of effects on word processing cannot be captured by behavioural measures. However, the pattern of obtained results clearly indicates that certain interplay of visual and phonological processing exists.

Main idea of this research was that if the sound-symbolism hypothesis was plausible, the following pattern of interactions would be expected: soft-sounding words would be more efficiently processed when presented within curvy frames and sharp-sounding words when presented within spiky frames compared to the incongruent situations: Soft-sounding words presented within spiky frames and sharp-sounding words presented within curvy frames. Given that this hypothesis was not confirmed, it seems that sound-symbolic mechanisms do not influence natural language processing. However, several recent studies dealing with sound-symbolic correspondences in artificial material have found that sound-symbolism effect "happens" on very early stages of language processing (Ković et al., 2010, Parise & Spence, 2012). It might be possible that some sound-symbolic correspondences which occur during the natural language processing are also positioned on early stages of

language processing but are overridden by higher-order processes, i.e., semantics processing. Pattern of interactions obtained in second and third experiment indicates that visual context influences "soft" verbal stimuli processing when visual and verbal information are presented simultaneously, while effects on "sharp" verbal stimuli are present only if visual context precedes verbal stimuli. However, the pattern of these interactions is reversed then the expected one. One possible explanation for this "reversed" sound symbolism effect could be that the presentation of words in the incongruent context leads to novelty effect. On the other hand, recently proposed language model by Monaghan, Christiansen and Fitneva (2011) could perhaps provide more plausible explanation of this unexpected result. According to Monaghan and his colleagues, certain systematic mappings in language do exist. However, the mappings between the word and general category are systematic, while mappings between the word and its particular meaning are arbitrary (Monaghan et al., 2011). The authors further suggest that this model of language structure provides optimal mode of functioning, since identification of precise meaning of the word is not necessary for determining lexical category—it can be done by identifying the general region of semantic space that the word inhabits, i.e. general category to which a word belongs and not the precise meaning of a word (Monaghan & Christiansen, 2006). On the other side, the existence of systematic mappings between the word and its meaning would strongly constrain the size of vocabulary. For that reason, arbitrary mappings present in this domain of language are optimal since they impose fewer constraints for the number of encoding words. This claim is further supported by the notion that contextual information is also important factor which provides additional information for identification of particular meaning of word (Monaghan et al., 2011). In the light of Monaghan's theory, it is possible that experimental design led to the reversed pattern of interaction obtained in third experiment. In this study, participants' task was to judge whether string of letters has the meaning or not. In order to solve this task, identification of the particular meaning of word was necessary. According to Monaghan's model, words with arbitrary mapping should be more efficiently processed in this kind of task so it might be possible that context assumed as congruent (curvy frame for soft sounding words and spiky for sharp sounding words) actually made processing of word in these experiments more difficult.

Furthermore, these findings directly contradict contemporary theories of language which assume that language is a function independent of other cognitive or sensory functions. Even more, those theories assume that sub-functions of language processing, phonology, orthography and semantics are being processed distinctly (as mentioned in Westbury, 2005). Baring this in mind, a question emerges whether orthography as well could influence the results of the experiments and lead to "orthographical contamination". Although Westbury indicated that this was not the case in his study, and that the letter shape of words written did not influence the effects, this issue can be important since the visual presentation of a word leads to indirect activation of representation of a word. In future studies, it is considered important to examine whether the obtained pattern of results exists when mental representation of word is being directly activated. For that reason, it is necessary to develop audio-visual version of lexical decision task with auditory presentation of verbal stimuli.

According to one recent study, sound symbolism effect is not

only influenced by the stimuli properties, but also by the characteristics of experimental procedure (Ković & Pejović, 2012). Namely, these authors have found that sound symbolism effect occurs only when mapping from auditory to visual stimuli and not vice versa. This design-dependent aspect of sound symbolism raises the question whether certain characteristics of experimental procedure which are usually not in focus of sound symbolism studies also play important role in discovering or even diminishing potential sound symbolism effect in language processing.

Beside the influence of temporal sequence of auditory and verbal stimuli, i.e. mapping from audio to visual or vice versa on appearance of sound symbolism effect (Ković & Pejović, 2012), the present study points out the important role of timing in presenting visual and auditory stimuli within the visual to verbal mapping. Having in mind the specific nature of the task, it was not possible to test the opposite direction of influence—from verbal to visual domain, since participants were judging lexical category of presented verbal stimuli. In this case, any kind of delaying response would probably influence reaction times as well.

Aiming to experimentally examine the potential role of sounding and meaning of word in sound-symbolism effect, this study revealed some additional aspects of stimuli which influenced the observed effect. Characteristics of visual information and frame exposition time revealed as important factors which have not been examined so far. Taken together, these results speak in favor of the claim about pre-semantic nature of sound symbolism. The overall pattern of the results in this study points out the importance of the temporal dimension in sound symbolism effect. For that reason, future studies should include event related potentials technique which can provide us a better insight in the dynamic of this effect. Analysis of wave deflections would allow us to further examine the nature of sound symbolism and potentially determine key moments in early phases of processing when this effect emerges. Furthermore, another line of future research on sound symbolism should consider the influence of the stimuli presentation order on these effects. There is a number of studies which have shown that certain sound-symbolic effects are design dependent and that effects are more likely to arise if the order of stimuli presentation is audio-visual compared to visual-auditory condition. Having in mind these findings, the question emerges whether the sound symbolism effect would be more profound in this study if the reversed order of stimuli presentation was used, i.e. could phonological and semantic properties of words also provide a context for visual information processing.

Acknowledgements

We would like to thank Rastko Pajković for his help with creating visual stimuli. This research was supported by the Ministry of Science and Technological Development of Serbia, grant numbers 179006 and 179033.

REFERENCES

- Bremner, A., Caparos, S., Davidoff, J., Fockert, J., Linnell, K., & Spence, C. (2013). "Bouba" and "Kiki" in Namibia: A remote culture makes similar shape-sound matches, but different shape-taste matches to Westerners. *Cognition*, 126, 165-172. doi:10.1016/j.cognition.2012.09.007
- Ciccotosto, N. (1991). Sound symbolism in natural language. *Dissertation Abstracts International*, 53, 541.
- Davis, R. (1961). The fitness of names to drawings: A cross-cultural study in Tanganyika. *British Journal of Psychology*, 52, 259-268. doi:10.1111/j.2044-8295.1961.tb00788.x
- Diffloth, G. (1994). I:big, a:small. In L. Hinton, J. Nichols, & J. Ohala (Eds.), *Sound symbolism* (pp. 107-114). Cambridge: Cambridge University Press.
- Ilić, O., Ković, V., & Janković, D. (2012). Crossmodal correspondences in natural language: Distribution of phonemes and consonant-vowel patterns in Serbian words denoting round and angular objects. *13th International Multisensory Forum*, Oxford.
- Janković, D., & Marković, S. (2001). Takete-Maluma phenomenon. *Perception*, 30, *ECVP 2001 Abstracts Supplement*. doi:10.1068/v010131
- Janković, D., Vučković, V., & Radaković, N. (2005). Consonants in the Takete-Maluma phenomenon: Manner and place of articulation. *Perception*, 34, *ECVP 2005 Abstracts Supplement*. doi:10.1068/v050611
- Köhler, W. (1929). *Gestalt psychology, an introduction to new concepts in modern psychology*. New York: Liveright.
- Ković, V., Plunkett, K., & Westermann, G. (2010). The shape of words in the brain. *Cognition*, 114, 19-28. doi:10.1016/j.cognition.2009.08.016
- Ković, V., & Pejović, J. (2012). Now you see it, now you don't: Design dependant sound symbolism effect in categorization studies. *13th International Multisensory Forum*, Oxford.
- Maurer, D., Pathman, T., & Mondloch, C. J. (2006). The shape of bou-bas: Sound-shape correspondences in toddlers and adults. *Developmental Science*, 9, 316-322. doi:10.1111/j.1467-7687.2006.00495.x
- Nielsen, A., & Rendall, D. (2011). The sound of round: Evaluating the sound-symbolic role of consonants in the classic Takete-Maluma phenomenon. *Canadian Journal of Experimental Psychology*, 65, 115-124. doi:10.1037/a0022268
- Padraic, M., & Christiansen, M. H. (2006). Why form-meaning mappings are not entirely arbitrary in language. *The 28th Annual Conference of the Cognitive Science Society*, Vancouver.
- Padraic, M., Christiansen, M. H., & Fitneva, S. A. (2011). The arbitrariness of the sign: Learning advantages from the structure of the vocabulary. *Journal of Experimental Psychology: General*, 140, 325-347. doi:10.1037/a0022924
- Parault, S. J., & Schwanenflugel, P. J. (2006). Sound-symbolism: A piece in the puzzle of word learning. *Journal of Psycholinguistic Research*, 35, 329-351. doi:10.1007/s10936-006-9018-7
- Parise, C. V., & Spence, C. (2012). Audiovisual crossmodal correspondences and sound symbolism: A study using the implicit association test. *Experimental Brain Research*, 220, 319-333. doi:10.1007/s00221-012-3140-6
- Plato (1998). *Cratylus*. Cambridge: Hackett Publishing Company.
- Ramachandran, V. S., & Hubbard, E. M. (2001). Synaesthesia—A window into perception, thought and Language. *Journal of Consciousness Studies*, 8, 3-34.
- Rogers, S. K., & Ross, A. S. (1975). A cross cultural test of the Maluma-Takete phenomenon. *Perception*, 4, 105. doi:10.1068/p040105
- Sapir, E. (1929). A study in phonetic symbolism. *Journal of Experimental Psychology*, 12, 225-239. doi:10.1037/h0070931
- Saussure, F. D. (1959). *Course in general linguistics*. New York: Philosophical Library.
- Spector, F., & Maurer, D. (2009). Synesthesia: A new approach to understanding the development of perception. *Developmental Psychology*, 45, 175-189. doi:10.1037/a0014171
- Westbury, C. (2005). Implicit sound symbolism in lexical access: Evidence from an interference task. *Brain & Language*, 93, 10-19. doi:10.1016/j.bandl.2004.07.006