Journal of Memory and Language 76 (2014) 47–60

Contents lists available at ScienceDirect

Journal of Memory and Language

journal homepage: www.elsevier.com/locate/jml

Nomina sunt consequentia rerum – Sound–shape correspondences with every-day objects figures

Andrea Flumini [a,*], Mariagrazia Ranzini [a,c], Anna M. Borghi [a,b]

[a] Department of Psychology, University of Bologna, Viale Berti Pichat 5, 40127 Bologna, Italy
[b] Institute of Cognitive Sciences and Technologies (ISTC), Italian National Research Council, Via San Martino della Battaglia 44, 00185 Rome, Italy
[c] FNRS, CRCN, Université Libre de Bruxelles (ULB), Avenue Franklin Roosevelt 50, 1050 Bruxelles, Belgium

ARTICLE INFO

Article history:
Received 18 July 2013
revision received 10 June 2014
Available online 9 July 2014

Keywords:
Sound-symbolism
Embodied cognition
Shape
Referent
Natural objects
Artifacts

ABSTRACT

Prior research on sound-symbolism has demonstrated the existence of sound–shape correspondences using ad hoc figures in double forced-choice paradigms. This led sound-symbolic skeptics to affirm that the reported effects were due to the properties of the figures shown or to the structure of the task used. In the present study, we hypothesized that the sound–shape correspondence effect would be observed when participants were required to choose which of two invented words would better suit an image representing a common object/entity. In addition, we hypothesized that the effect would be modulated by the object/entity category, and that natural objects would be represented with smoother shapes compared to artifacts. Results confirmed the "classic" takete–maluma effect both when participants chose a name for figures of natural objects (e.g., leaf) and artifacts (e.g., fork), and when they chose a name for figures of natural (e.g., animals) and artificial agents (e.g., robots). Moreover, when participants had to name agents, a modulation of the category (artificial vs. natural) emerged: sound–shape correspondence was not observed with robots, which were associated more often with jagged responses despite their actual shape. Results are discussed in the framework of embodied cognition theories.

Introduction

In his *Institutiones* (VI cent. C.E.) Justinian affirmed that *Nomina sunt consequentia rerum* to indicate that verbal language takes its origin from the things it denotes. The actual beginning of the philosophical debate on the origin of language precedes this Latin laconism by at least some centuries. In Plato's *Cratylus* (IV cent. B.C.E.), for example, there is talk pertaining the possible existence of a *resemblance relation* between the structure of words and what they denote. Even if this is not Plato's position, this dialog suggests the emergence of a naturalistic vision of language as opposed to the alternative sophists' view, according to which the word–referent relation is totally arbitrary.

The principle of arbitrariness of human language, however, still remains widely accepted among linguists, philosophers and psychologists (Kovic, Plunkett, & Westermann, 2009; Nielsen & Rendall, 2011; Nygaard, Herold, & Namy, 2009a, 2009b; for a different position; see Reilly, Westbury, Kean, & Peelle, 2012). Ever since the *Course in General Linguistics* (De Saussure, 1916), contemporary sciences of language have followed the perspective which affirms that a label is always arbitrarily assigned to a referent (e.g., object, event, relation, etc.), each assignation being grounded only on socio-cultural conventions (Nielsen & Rendall, 2011; Pietrandrea, 2002). Even names with a transparent iconic relation with their referents (e.g., sounds) are considered to be accidentally assigned.

---

* Corresponding author.
E-mail address: andrea.flumini2@unibo.it (A. Flumini).

http://dx.doi.org/10.1016/j.jml.2014.06.004

By iconicity we refer to the similarity between certain properties of the words and the sensorimotor characteristics of their referents, as occurs in onomatopoeias, in which words evoke acoustic experiences (e.g., *buzz*, *hiss*).

In spite of the wide acceptance of the principle of arbitrariness of language, some dissonant voices are starting to emerge. Research on language evolution has suggested that the emergence of lexicon conventionality could be a belated stage in the evolution of human language. Indeed, from a phylogenetic point of view, linguistic conventions might be based on originally iconic linguistic forms (Merleau-Ponty, 1945; Steels, 2011), and so considered as the outcome of a process which has recursively and efficiently maximized communication (Burling, 1999; Corballis, 2009; Fay, Garrod, Roberts, & Swoboda, 2010; Garrod, Fay, Lee, Oberlander, & MacLeod, 2007; Zipf, 1949; for modeling work in this direction see Baronchelli, Gong, Puglisi, & Loreto, 2010).

A further important source of inspiration for research aimed at demonstrating that language is not always arbitrary, in particular when we consider face-to-face interactions, comes from the embodied and grounded approach to cognition (reviews in Barsalou, 2008; Borghi & Pecher, 2011). This approach solves the so-called symbol-grounding problem (Harnad, 1990) by proposing that symbols are grounded in the same systems as used by perception, action and emotion. According to this perspective, during the processing of words we would re-activate previous experiences with their referent. For example, the word "cat" would re-enact the experience of seeing a cat, caressing it, feeding it, and so on. In line with this view, recent research has started to emphasize the role and importance of iconicity in language, as a possible mean to connect the linguistic-communicative form with the sensorimotor characteristics of word referents. This research hypothesizes that arbitrariness might not be the rule during face-to-face communication, where gestures and iconic words might be consistently used. For example, Perniss and Vigliocco (in press) have highlighted the importance of iconicity in both spoken and signed language. They suggest that the role of iconicity is crucial in three important aspects of language processing: phylogenesis, ontogenesis and processing. In phylogenesis iconicity facilitates displacement, i.e., the ability to refer to things distant in time and space, while in ontogenesis it provides a mechanism for establishing referentiality, linking linguistic form and meaning and hence facilitating word learning. Finally, in language processing iconicity facilitates the grounding of words in sensorimotor and emotional systems, thereby determining embodiment.

In this paper we investigate the word–referent relationship focusing on the direct bindings between the word sound and certain aspects of the referent's appearance (e.g., shape). By word sound we refer to a multimodal experience, including both the acoustic experience during language comprehension and the phono-articulatory experience during word production. The idea underlying this investigation is that words can entertain a non-arbitrary relation with their referents. This process has been identified in speech as sound-symbolism or phonosemantics (Hinton, Nichols, & Ohala, 1994; Parise & Pavani, 2011), and it is something that works in a way similar to iconicity in sign languages (e.g., Corballis, 2002, 2009; Gentilucci & Corballis, 2006; Perniss, Thompson, & Vigliocco, 2010; Pietrandrea, 2002; Pizutto & Volterra, 2000; Thompson, Vinson, & Vigliocco, 2009, 2010).

The psychological literature on sound-symbolism is longstanding, dating back at least to the first half of the last century. In fact, Edward Sapir, in his most well known essay on language, already suggested that verbal labels are able to catch aspects of what they refer to. As evidence of this, Sapir (1929) reported results of an experiment where almost all the participants – mother tongue English speakers – assigned the invented names *mil* or *mal* to a small or a big table respectively, thus intuitively coupling the sound of the words with the size of the objects. Similarly, Wolfgang Köhler (1929, 1947) revealed the existence of correspondences between word sounds and visual shapes: Spanish participants intuitively assigned an invented word with rounded vowels (*baluma*, or in 1947 *maluma*) to rounded invented figures and an invented word with unrounded vowels (*takete*) to jagged invented figures. Finally, the cross-linguistic ability to guess the meaning of foreign words, examined using samples of various different mother tongues (e.g., Brown, Black, & Horowitz, 1955; Gebels, 1969; Hinton et al., 1994; Koriat & Levy, 1979; Kunihira, 1971), drove some authors to explicitly affirm that speech may have emerged from universal imitative connections between sounds and meanings (Kovic et al., 2009).

In the last ten years, research conducted on speakers of different languages has gathered results which support the idea of sound–shape correspondences (e.g., Arata, Imai, Sotaro, Guillaume, & Okada, 2010; Asano et al., 2011; Iwasaki, Vinson, & Vigliocco, 2007; Kantartzis, Imai, & Kita, 2011; Kovic et al., 2009; Nielsen & Rendall, 2011; Nygaard et al., 2009a, 2009b; Parault, 2006; Ramachandran & Hubbard, 2001; Spector & Maurer, 2008; Westbury, 2005). In particular, Maurer, Pathman, and Mondloch (2006) investigated the takete–maluma phenomenon in 2.5-year-old children and adults to assess whether children reliably match words with rounded vowels to rounded shapes, and words with unrounded vowels to jagged shapes (for a different interpretation, see Nielsen & Rendall, 2011). Participants were simultaneously presented with two shapes, a rounded and a jagged one, and had to name each shape choosing the name from two alternatives, one with rounded vowels and the other with unrounded ones. Results showed that children, like the adults, matched names with rounded vowels to rounded shapes and names with unrounded vowels to the jagged ones, indicating that sound–shape correspondences are already at work at the earliest stages of language acquisition.

As in Maurer et al. (2006), several studies on sound–shape correspondences have adopted naming tasks with a two alternatives forced-choice design. Although results of studies adopting this kind of paradigm support the hypothesis of a non-arbitrary relation between words and their referents, there are some methodological issues which might bring a sound-symbolism skeptic to say that the results may be influenced by confounders. For example, as far as we know, most experiments adopting the naming paradigm proposed forced-choice tasks where two words, one sonorant and one strident, were simultaneously presented together with the pairs of stimuli figures. The risk involved in such a design is twofold, as clearly highlighted by Nielsen and Rendall (2011). First, the purpose of the experiment can become very transparent to subjects, who are required to compare both figures and names in each trial. Second, the simultaneous presentation does not allow disentangling when the results are due to two matches, one for strident sounds/jagged shapes and another for sonorant sounds/rounded shapes, and when there is only one match in one of the two directions. Indeed, in each trial the subject's second-choice is automatically defined by the image and the name they have coupled first. Another problem is that these experiments might be poorly ecological and might not reflect what happens in real life, since both the words and the stimulus figures are created ad hoc for the experiment. Moreover, these ad hoc figures emphasize the features under investigation (e.g., roundness, jaggedness), thus possibly inducing an enhancement of the reported behavioral effects (see Nielsen & Rendall, 2011 for a similar critique).

For these reasons, in the present study we tested the correspondences between sounds and shapes avoiding a double forced-choice paradigm, and using visual stimuli representing every-day objects and entities. Participants were required to choose, from between two words, the most suitable name for an image which represented a well known object/entity, common in every-day life.

Using every-day stimuli allowed us to overcome one further limitation of current studies on sound-symbolic correspondences. To the best of our knowledge, no study so far has taken into account the possible effects of categorical differences on sound-symbolic correspondences. In contrast, the literature on concepts and categorization has highlighted profound differences in the representation of artifacts and natural objects, and of living and not living entities, as indicated by neural studies on brain activation (for a review, see Martin, 2007), by neuropsychological studies on categorical deficits (e.g., Humphreys & Forde, 2001; Gainotti, 2000) and by behavioral studies on categorization in children (e.g., Mandler, 1992; Rakison & Oakes, 2003) and adults (e.g., Borghi et al., 2007; Laurence & Margolis, 2007; Roversi, Borghi, & Tummolini, 2013). Some authors have underlined that categorical distinctions in infants might be based on perceptual cues as well as on motion cues (e.g., Mandler, 1992, 2004). Motion cues differ for animals and artifacts: animals are characterized by self-propelled movements and by nonlinear and smooth motion paths, while artifacts are characterized by induced movements and linear motion paths (Mandler, 1992; Rakison & Poulin-Dubois, 2001, 2002). In a similar vein, recent research has shown that some features of words, such as their grammatical gender, are not arbitrarily related to the characteristics of the referents belonging to different categories. Sera et al. (2002) asked English, Spanish, French and German children and adults to attribute male or female voices to inanimate objects. The results showed that for gendered languages the relation between grammatical gender and the perception of items as being male- or female-like was not arbitrary: speakers of Spanish and French, unlike speakers of German, relied on grammatical gender in their assignment of male and female voices to inanimate objects. The different results obtained with French and Spanish compared to German speakers suggest that the effect is present within grammatical systems characterized by two gender categories (German has three), where grammatical and natural gender are highly correlated. More crucially for us, the gender assignment interacted with the category of the objects, showing that artificial objects were more often perceived as male-like than natural ones, which were instead considered more female-like. Thus, the authors pointed out that the natural/artificial distinction may be correlated to several factors, such as item shape (rounded/jagged), density (light/heavy), or common use (typically used by females/males), which map onto the grammatical gender assignments in a phonesemantic-like manner.

On the basis of the reviewed evidence, we reasoned that manipulating the category would allow us to obtain some hints on how categories are represented with a paradigm never used in this context. As suggested by Sera et al. (2002) it is possible, in fact, that one further feature distinguishing natural objects and artifacts regards their shape: the shape of natural objects could be mentally represented as smoother compared to that of artifacts. In light of these considerations, we designed a paradigm that allowed us to study the development of the sound-symbolic correspondence with every-day objects/entities, which belonged to different categories, that is artifacts and natural objects.

We will now expose our hypotheses, based on the main manipulations we decided to introduce. First, we hypothesized that the sound–shape correspondence effect would be conserved if the figures to be labeled represent every-day entities and, second, if they are presented one by one. If the two hypotheses are confirmed, this will show that the effect is due neither to the structure of the task nor to the properties of the shown figures. Third, we hypothesized that the sound–shape effect is modulated by the category of the visual stimuli. To this aim, in Experiment 1 the stimuli figures represented every-day objects, which could be natural objects or artifacts, and in Experiment 2 the stimuli were figures representing natural (i.e., animals) or artificial agents (i.e., robots), in order to explore whether natural entities are represented with smoother shapes as compared to artificial entities.

Experiment 1

Method

Participants

Twenty-four undergraduate students from the University of Bologna participated in the experiment for course credits (9 males; mean age = 20.79 (2.23); 2 left-handed by self-report). All participants had normal or corrected-to-normal vision and were naive as to the purposes of the experiment.

Materials

The task consisted in choosing one word from a pair as the preferred name for the picture stimulus. The word pairs were simultaneously displayed on a computer screen, with no acoustic presentation. This reduced the possible sources for any sound-symbolic effects to the phonetic and phonotactic, eliminating the prosodic one (Kantartzis et al., 2011). Stimuli consisted of 24 black-and-white line figures chosen from the graphic database by Lotto, Dell'Acqua and Job (2001). Twelve figures referred to natural objects and twelve figures referred to artifacts, and each set was composed of 6 rounded-shaped and 6 angular-shaped figures. The pictures were rated by each participant after the experimental session on a 7-point Likert scale for sharpness/roundness (1 = *very sharp* and 7 = *very rounded*). Participants were presented with each figure individually, and reported autonomously their ratings on a paper form. Ratings were analyzed in a 2×2 ANOVA with the within factors Figure Type (Natural vs. Artificial) and Figure Shape (Rounded vs. Jagged). Across the manuscript, for the participants analysis (indicated by F_1), condition means were obtained by averaging across items, and for the materials analysis (indicated by F_2) they were obtained by averaging across participants.

The ANOVA showed the main effects of Figure Type, $F_1(1, 23) = 68.11$, $MSe = 0.25$, $p < .001$; $F_2(1, 5) = 24.98$, $MSe = 0.17$, $p < .01$ (Natural $M = 4.21$, Artificial $M = 3.37$), and of Figure Shape, $F_1(1, 23) = 718.69$, $MSe = 0.31$, $p < .001$; $F_2(1, 5) = 134.23$, $MSe = 0.41$, $p < .001$ (Rounded $M = 5.31$, Jagged $M = 2.26$). The interaction between Figure Type and Figure Shape was also significant, but only with subjects as random factor, $F_1(2, 46) = 11.26$, $MSe = 0.24$, $p < .01$; $F_2(2, 10) = 1.06$, $MSe = 0.63$, $p = .35$ (Natural: Rounded shape $M = 5.89$, Jagged shape $M = 2.51$; Artificial: Rounded shape $M = 4.72$, Jagged shape $M = 2.01$) (LSD, all *ps* < .001).

To verify if participating in the experiment exerted an influence on participants' attitudes about object shapes, an additional independent group was asked to make an identical evaluation of the pictures. The independent group was composed of twenty-four students from the University of Bologna, participating for course credits (7 males; mean age = 23.42 (3.51); 5 left-handed by self-report). They had normal or corrected-to-normal vision and were naive as to the purposes of the questionnaire. The independent ratings were again analyzed with the 2×2 ANOVA, with the within factors Figure Type (Natural vs. Artificial) and Figure Shape (Rounded vs. Jagged).

The ANOVA showed the main effects of Figure Type, $F_1(1, 23) = 68.54$, $MSe = 0.20$, $p < .001$; $F_2(1, 5) = 11.26$, $MSe = 0.30$, $p < .05$ (Natural $M = 4.37$, Artificial $M = 3.62$), and Figure Shape, $F_1(1, 23) = 718.69$, $MSe = 0.31$, $p < .001$; $F_2(1, 5) = 258.17$, $MSe = 0.27$, $p < .001$ (Rounded $M = 5.69$, Jagged $M = 2.30$). The interaction between Figure Type and Figure Shape was also significant, but only in the by-subjects analysis, $F_1(2, 46) = 11.37$, $MSe = 0.14$, $p < .01$; $F_2(2, 10) = 3.47$, $MSe = 0.12$, $p = .12$ (Natural: Rounded shape $M = 6.19$, Jagged shape $M = 2.55$; Artificial: Rounded shape $M = 5.18$, Jagged shape $M = 2.06$) (LSD, all *ps* < .001). Thus, the ratings of the independent group showed a pattern identical to the ratings of the experimental group. In fact, the ratings predict both the sound-symbolic correspondence of names and shapes, and an effect of the category, which should emerge in the experimental results in a form similar to the standard sound-symbolic effect.

The 8 words, used as names for the 24 pictures, were taken from the study by Maurer et al. (2006) and manipulated in the way they were written to obtain the same sound in Italian as they have in English (e.g., the English *bouba* was transformed into the Italian *boba*, see the Appendix for the complete list of stimuli). The 8 names were coupled in the same four pairs as the experiment by Maurer et al. (2006), always consisting of one sonorant, round-sounding word (e.g., *maluma*) and one strident, sharp-sounding word (e.g., *takete*). Each pair of words was visually presented on a computer screen under the picture to be named (see Fig. 1). Thus, depending on the object's appearance in the figure (Figure Shape: Rounded vs. Jagged) and on the phonological characteristics of the name (Response Type: Rounded vs. Jagged), in each trial it was possible to obtain a sound-symbolic combination (e.g., choosing *maluma* as the name of a rounded-shaped figure) or a non-sound-symbolic combination (e.g., choosing *maluma* as the name of a jagged-shaped figure). For the sake of simplicity we defined the two levels of both Figure Shape and Response Type factors as Rounded vs. Jagged.

Design and procedure

Participants sat 50 cm from the computer screen. Each trial began with a fixation point (+) lasting for 500 ms. The stimulus picture was then displayed centrally, remaining on the screen for 5 s or until a response was made. The two names were simultaneously presented under the picture, one on the left and the other on the right (the order of the names was counterbalanced among subjects). Participants were required to decide which of the two names was more suitable for the picture displayed above them by pressing two keys on an Italian QWERTY keyboard. The keyboard was positioned close to the screen so that each of the two keys was located perfectly below the name to which it corresponded: participants pressed the 5 key to choose the name on the left and the 9 key for the name on the right (see Fig. 1). At the beginning of the experiment they were instructed to respond as quickly as they could. They did not receive any feedback about the accuracy of their responses, as they were told no correct/incorrect values had been established for the trials. Considering that each of the 24 pictures was presented once with each of the 4 word pairs, the experiment consisted overall in 96 experimental trials, preceded by 8 training trials to familiarize participants with the procedure.

Data analysis and results

Missing responses (i.e., responses that required more than 5 s to be given) were removed. The very low rate of missing responses (0.17%) testified that the task was easy to perform. All the remaining responses were transformed into a percentage of choosing a rounded response (the percentages of rounded and jagged responses add up to 100%, so cannot be considered independent) and were entered in a 2×2 ANOVA with the within factors Figure Type (Natural vs. Artificial) and Figure Shape (Rounded vs. Jagged).

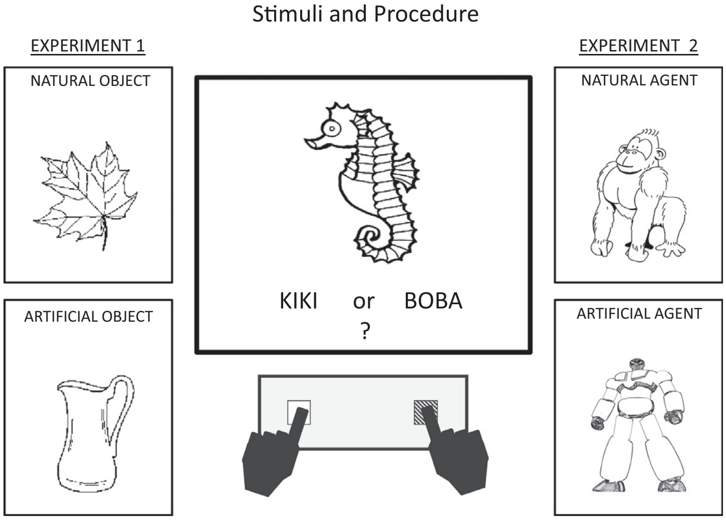

Fig. 1. Experimental stimuli and procedure.

Fisher's LSD *post hoc* tests were conducted on significant interactions.

The ANOVA on the percentage of rounded responses showed only the expected main effect of Figure Shape, $F_1(1, 23) = 20.43$, $MSe = 341.23$, $p < .001$; $F_2(1, 5) = 21.04$, $MSe = 80.29$, $p < .01$, due to rounded sounding names being more frequently assigned as labels to Rounded shapes ($M = 61.86\%$) than to Jagged shapes ($M = 44.82\%$), as reported in Fig. 2. No other effects were observed.

In this task, each picture was repeated 4 times. To verify if picture repetition could have affected the results, we repeated the analysis on the data collected in Experiment 1 considering only the trials of the first presentation of each picture. The percentages of rounded responses to the first presentation were entered in a 2 × 2 ANOVA with the within factors Figure Type (Natural vs. Artificial) and Figure Shape (Rounded vs. Jagged). Fisher's LSD *post hoc* tests were conducted on significant interactions.

The ANOVA on the percentage of rounded responses showed in the by-subjects analysis a marginally significant main effect of Figure Shape, $F_1(1, 23) = 3.95$, $MSe = 525.59$, $p < .06$; $F_2(1, 5) = 2.61$, $MSe = 205.99$, $p = .17$, due to rounded sounding names being more frequently associated to Rounded shapes ($M = 55.90\%$) than to Jagged shapes ($M = 46.60\%$). No other effects were observed.

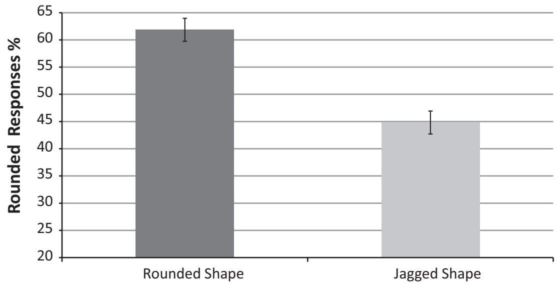

Fig. 2. Experiment 1 – main effect of Figure Shape.

We were also aware that ANOVAs are frowned on for proportional data (see Jaeger, 2008), but we are confident in our results since we were able to confirm the pattern of results observed using ANOVA by mixed effect models[1].

Finally, to control if the experimental results were related to the pictures ratings, we correlated subjects' mean rating to each picture and the mean percentage of rounded response then assigned to that picture. The correlation was positive and highly significant ($r = .24$, $p < .001$), indicating that the more the picture was subjectively perceived as round, the higher the probability for participants to label it with a rounded word.

Discussion of Experiment 1

As predicted, participants more frequently chose rounded words (e.g., *maluma*) as names for rounded shaped object figures (e.g., orange) and jagged words (e.g., *takete*) as names for jagged shaped object figures (e.g., fork), showing a high sensitivity to the correspondence between word sounds and visual shapes even if the figures to be named represented familiar objects. The same effect was marginally observed also when we considered only the first presentation of the visual stimuli.

---

[1] The results of the ANOVA were confirmed by mixed effect models through which we simultaneously took into account the fine-grained effect of perceiving roundness (i.e., rating values) and the variance due to random factors such as Item, or Name Pair. We defined the first model as the most complete one, i.e., with Figure Type and Rating as fixed effects, and Item (i.e., each picture), Name Pair (i.e., the two names to choose between) and Participant in interaction with Figure Type and Rating as random effects. Then, at each consecutive step we tried another model by firstly eliminating the random effects one by one, then the interaction between fixed effects, and finally the fixed effects one by one. Each time we performed a log-likelihood test to determine if one model was worse than the previous one in fitting the data: if no significant difference was observed, or if the last model was significantly better than the previous one (i.e., log-likelihood value closer to 0) then the less complete model was chosen. The final model resulting from this procedure consisted of all random effects, and only Rating as fixed effect (AIC = 2888, BIC = 2945, Log-likelihood = −1434).

Importantly, the results of the correlational analysis between ratings and performance and of mixed effects models confirmed this finding, and indicated that sound–shape correspondence is a continuous, fine-grained effect. This result confirms evidence on the sound–shape correspondence effect. Furthermore, it suggests that attaching labels to external entities, and specifically to every-day objects which already have a common and conventionalized label, is not necessarily an arbitrary activity. Finally, we were able to obtain the sound–shape correspondence effect by presenting the stimuli one by one, thus avoiding potential limits of previous studies, such as the transparency of the experimental aim and the enhancement of the observed effects (see Nielsen & Rendall, 2011).

However, despite the fact that results from all the ratings made on the pictures of Experiment 1 showed that artifacts were perceived as sharper than natural objects, the predicted effect of the object category (natural vs. artificial) was not found in the choice of name.

Experiment 2

In Experiment 1 we asked participants to choose a name for pictures of already known objects. We found the predicted sound-symbolic correspondence between names and shapes, but no effect of the object category. One possible cause of the absence of a category effect is that very different items were compared – for example, the category of artifacts included both very simple tools (e.g., spoon) and more complex ones (e.g., compass). In Experiment 2 we investigated whether the effect would be found using a more compact sub-category within the artificial and natural entities, i.e., the category of *agents*. We define an *agent* as an entity perceivable as having the ability to autonomously *act* or *move*, and endowed with features typically linked to animacy (Backscheider, Gelman, Martinez, & Kowieski, 1999; Landau, Smith, & Jones, 1988, 1992), i.e., eyes. In this sense we consider as agents both animals (i.e., natural agents) and robots (i.e., artificial agents). In contemporary cultures robots have become a quite credible kind of agent, partly on account of common sense beliefs about Turing machines and Artificial Intelligence, and mostly due to the part they play in popular culture (e.g., science fiction books, comics, movies).

The stimuli of Experiment 2 were figures representing animals and anthropomorphous robots. This choice allowed us to compare two categories whose members are more similar than those of the previously used artifacts and natural objects categories. Lastly, considering that an ontogenetic continuity of sound-symbolism has already been shown in the literature with ad hoc stimuli (e.g., Maurer et al., 2006), here we tested a sample composed of adults and children, to investigate the sound-symbolic phenomenon related to every-day categories in function of age. The reasons why we decided to test participants of different ages are multifold. First, studies on sound-symbolism typically also involve children, and considering that Experiment 1 confirmed that sound-symbolic correspondences are able to affect the labeling of every-day objects, we thought it was important to verify whether different processes characterize sound-symbolism with common objects in adults and children. Second, we were interested in whether the relationship between categories and sounds is similarly present in childhood as it is in adulthood, as observed for the relationship between shapes and sounds (e.g., Maurer et al., 2006). Unlike previous studies, which investigated sound–shape correspondences in children mostly below the age of three years (e.g., Maurer et al., 2006), with the sole exception of Davis (1961), here we tested older children (i.e., about 9 years old) in order to investigate the relationship between categories and sounds in populations with a clear knowledge of semantic categories such as artifacts and natural agents. We started from the assumption that language influences categorization (e.g., Lupyan, 2012; Sera et al., 2002), and that a prolonged use of a given label may render the categories less malleable and more stable than they are initially. On this basis, we predicted that children would be more flexible than adults in forming categories such as natural and artificial agents, since the features that appear to be more salient depend on developmental processes in which language plays a fundamental role (Sera et al., 2002). While the formation of categories of natural and artificial agents should be profoundly influenced by language, this should be less so for shape categories. Our interest in the different developmental pattern of the relation between sound and shape compared to that between sound and category derives from these considerations. To the best of our knowledge, no studies up to now have investigated the development of sound-symbolism effects in relation to semantic categories in childhood. On this matter, the question of whether sound-symbolism effects pertaining to categories (e.g., the categories of natural and artificial agents) are influenced by linguistic experiences and changes during development still remains to be investigated. In this way, we would be able to relate our results to other findings in the literature on sound-symbolism, and assess whether the symbolic relationship between categories and sounds develops earlier or later in life than the one between shapes and sounds.

To summarize, we predicted that by using the more specific subcategory of agents, we would find a modulation of the sound-symbolic effect in function of both category and age. In particular, we expected a more marked effect of category on the choice of name in adults, as they may have a more clear distinction between natural and artificial agents due to experience and prolonged use of verbal labels (see also Sera et al., 2002, for a similar conclusion pertaining to gender). In addition to linguistic influences on categorization (e.g., Lupyan, 2012), we expected that children's categories would be more flexible and malleable compared to the categories of adults also because in children the category of animated entities might be broader, and the representations of natural and artificial agents might overlap. For this reason we decided to use artificial agents rather than standard artificial objects. Finally, we expected that adults' categories would be less permeable to learning effects occurring during the experiment (e.g., effects due to figure repetitions).

Method

Participants

Twenty-four children (15 males; mean age = 8.79 (1.06); all right-handed) participated in the experiment as volunteers, and twenty-four students from the University of Bologna (10 males; mean age = 21.04 (2.91); 3 left-handed by self-report) participated for course credits. All participants had normal or corrected-to-normal vision and were naive as to the purposes of the experiment.

Materials

The materials consisted of twenty-four black-and-white pictures of manmade drawings, of which 12 represented animals (6 rounded and 6 jagged-shaped) and 12 robots (6 rounded and 6 jagged-shaped), and the same eight words used in Experiment 1 (see the Appendix for the complete list of stimuli).

As in Experiment 1, after the experimental session, each subject rated the pictures on a 7-point Likert scale for roundness/sharpness. A mixed 2 × 2 × 2 ANOVA with the between factor Group (Children vs. Adults), and the within factors Figure Type (Animal vs. Robot) and Figure Shape (Rounded vs. Jagged) was performed.

The ANOVA showed a significant effect of Group, $F_1(1, 46) = 5.46$, $MSe = 0.67$, $p < .05$; $F_2(1, 10) = 3.72$, $MSe = 0.25$, $p < .09$ (Children $M = 3.86$, Adults $M = 3.58$), Figure Type, $F_1(1, 46) = 98.33$, $MSe = 0.50$, $p < .001$; $F_2(1, 10) = 31.14$, $MSe = 0.40$, $p < .01$ (Animal $M = 4.23$, Robot $M = 3.21$), and Figure Shape, $F_1(1, 46) = 316.51$, $MSe = 1.02$, $p < .001$; $F_2(1, 10) = 331.59$, $MSe = 0.24$, $p < .001$ (Rounded $M = 5.01$, Jagged $M = 2.42$). One interaction was also reliable (marginally significant in F_2), the Group × Figure Type, $F_1(2, 92) = 11.39$, $MSe = 0.50$, $p < .01$; $F_2(2, 20) = 3.61$, $MSe = 0.40$, $p < .09$ (Children: Animal $M = 4.19$, Robot $M = 3.52$ – LSD $p < .05$; Adults: Animal $M = 4.26$, Robot $M = 2.90$ – LSD $p < .001$). No other interaction reached significance.

As in Experiment 1, in order to verify the possible influence of the repeated exposure to the labeling task on the attitudes about visual stimuli shapes, an additional independent group similarly performed the ratings on the pictures of Experiment 2. The independent group was composed of twenty-four children (11 males; mean age = 9.13 (0.45); 2 left-handed by self-report) who participated as volunteers, and twenty-four students from the University of Bologna (12 males; mean age = 23.54 (3.37); 3 left-handed by self-report) who participated for course credits. All participants had normal or corrected-to-normal vision and were naive as to the purposes of the questionnaire. The independent ratings were analyzed using a 2 × 2 × 2 ANOVA with the between factor Group (Children vs. Adults) and the within factors Figure Type (Animal vs. Robot), Figure Shape (Rounded vs. Jagged).

The ANOVA showed the following factors as significant: Group, $F_1(1, 46) = 6.74$, $MSe = 1.06$, $p < .05$; $F_2(1, 10) = 2.61$, $MSe = 0.69$, $p = .14$ (Children $M = 3.88$, Adults $M = 3.49$), Figure Type, $F_1(1, 46) = 67.99$, $MSe = 0.44$, $p < .001$; $F_2(1, 10) = 7.43$, $MSe = 0.99$, $p < .05$ (Animal $M = 4.08$, Robot $M = 3.30$), and Figure Shape, $F_1(1, 46) = 589.67$, $MSe = 0.55$, $p < .001$; $F_2(1, 10) = 404.36$, $MSe = 0.20$, $p < .001$ (Rounded $M = 4.99$, Jagged $M = 2.39$). Three interactions almost reached significance in the by-participant analysis, starting from the Group × Figure Shape, $F_1(2, 92) = 3.42$, $MSe = 0.55$, $p < .08$; $F_2(2, 20) = 2.33$, $MSe = 0.20$, $p = .16$ (Children: Rounded shape $M = 5.08$, Jagged shape $M = 2.68$; Adults: Rounded shape $M = 4.89$, Jagged shape $M = 2.10$). Then, the Figure Type × Figure Shape, $F_1(2, 92) = 3.45$, $MSe = 0.37$, $p < .07$; $F_2 < 1$ (Animal: Rounded shape $M = 5.46$, Jagged shape $M = 2.71$; Robot: Rounded shape $M = 4.51$, Jagged shape $M = 2.08$). Finally, the three-way interaction of the factors Group, Figure Type and Figure Shape was almost significant, $F_1(3, 184) = 4.01$, $MSe = 0.37$, $p < .06$; $F_2 < 1$ (Children – Animal: Rounded shape $M = 5.47$, Jagged shape $M = 3.08$; Robot: Rounded shape $M = 4.69$, Jagged shape $M = 2.29$. Adults – Animal: Rounded shape $M = 5.45$, Jagged shape $M = 2.33$; Robot: Rounded shape $M = 4.33$, Jagged shape $M = 1.88$). No other effects were observed.

Overall, the independent group ratings partially confirmed the pattern of the experimental group, as both ratings predict the "classic" sound-symbolic correspondence of sounds and shapes, and also an effect of the category (by itself, or in interaction with the stimuli shape) for both experimental groups. For more details of the analyses on ratings see the Supplementary Materials section, where we report the results of an additional ANOVA on the ratings made in Experiment 2 by the experimental and independent groups (i.e., considering the participation in the experiment as a factor) in order to highlight any differences. The Supplementary Materials section also reports two additional ANOVAs, one that compared the ratings of the experimental adult groups and the other the ratings of the independent adult groups from both Experiment 1 and 2 (i.e., considering the experiment as a factor, as this analysis was conducted on the data collected for the adult group of both experiments to clarify their relations).

Design and procedure

The design and procedure were exactly the same as in Experiment 1, except that the stimuli used, instead of being pictures of natural objects and artifacts, were pictures of animals and robots.

Data analysis and results

Missing responses were removed (1.28%) and the remaining responses were entered as percentages in a mixed 2 × 2 × 2 ANOVA with the between factor Group (Children vs. Adults) and the within factors Figure Type (Animal vs. Robot), Figure Shape (Rounded vs. Jagged). Fisher's LSD *post hoc* tests were conducted on significant interactions.

The analysis showed, as expected, the significant main effect of the factor Figure Shape, $F_1(1, 46) = 11.54$, $MSe = 132.49$, $p < .01$; $F_2(1, 10) = 28.41$, $MSe = 33.12$, $p < .001$, due to Rounded shapes ($M = 52.45\%$) being more often labeled with a rounded sounding name than Jagged shapes ($M = 46.81\%$). The main effect of Figure Type, $F_1(1, 46) = 3.72$, $MSe = 1079.55$, $p < .06$; $F_2(1, 10) = 4.65$, $MSe = 57.34$, $p < .06$, almost reached significance, due to Animals ($M = 54.21$) being more often labeled with rounded names than Robots ($M = 45.06$).

The interaction Group × Figure Type was significant, $F_1(2, 92) = 5.85$, $MSe = 1079.55$, $p < .05$; $F_2(2, 20) = 39.76$, $MSe = 33.12$, $p < .001$. While an undifferentiated pattern in respect of Figure Type was observed in the Children's responses (Animal $M = 47.92\%$, Robot $M = 50.24\%$), the category of the stimuli exerted a clear effect on the responses of Adults (Animal $M = 60.49\%$, Robot $M = 39.88\%$; LSD $p < .01$) (see Fig. 3).

The interaction Figure Type × Figure Shape was also significant, $F_1(2, 92) = 23.70$, $MSe = 90.55$, $p < .001$; $F_2(2, 20) = 12.89$, $MSe = 32.86$, $p < .01$, showing that Animal Rounded shape ($M = 60.37\%$) received significantly more rounded responses than all other conditions (Animal Jagged shape $M = 48.04\%$; Robot Rounded shape $M = 44.54\%$, Robot Jagged shape $M = 45.58\%$) (LSD, all $ps < .001$), with the category effect suppressing the sound–shape correspondence for natural jagged shape and especially for artificial rounded shape (see Fig. 4). No other effects or interactions were significant.

As in Experiment 1, a further analysis considering only the first presentation of each picture was performed to check for the effect of figure repetition. The rounded responses to the first presentation were entered as percentages in the mixed 2 × 2 × 2 ANOVA with the between factor Group (Children vs. Adults) and the within factors Figure Type (Animal vs. Robot), Figure Shape (Rounded vs. Jagged). Fisher's LSD *post hoc* tests were conducted on significant interactions.

The analysis showed the interaction Figure Type × Figure Shape as marginally significant in the by-subject analysis, $F_1(2, 92) = 3.65$, $MSe = 376.35$, $p < .06$; $F_2(2, 20) = 2.09$, $MSe = 126.62$, $p = .17$, indicating a modulation of Animal Rounded shape ($M = 59.55\%$), which received significantly more rounded responses than all other conditions (Animal Jagged shape $M = 50.42\%$; Robot Rounded shape $M = 48.30\%$, Robot Jagged shape $M = 49.86\%$) (LSD $p < .001$). The interaction Group × Figure Type was almost significant in the analysis with item as random factor, $F_1(2, 92) = 1.51$, $MSe = 1416.03$, $p = .23$; $F_2(2, 20) = 4.27$, $MSe = 130.13$, $p < .07$, confirming the difference between children and adults in respect of category (Children: Animal $M = 51.89\%$, Robot $M = 52.90\%$; Adults: Animal $M = 57.38\%$, Robot $M = 44.77\%$ – LSD $p < .05$). No other effects or interactions were significant.

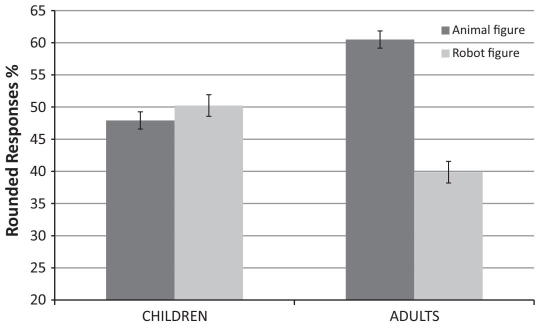

Fig. 3. Experiment 2 – interaction Group × Figure Type.

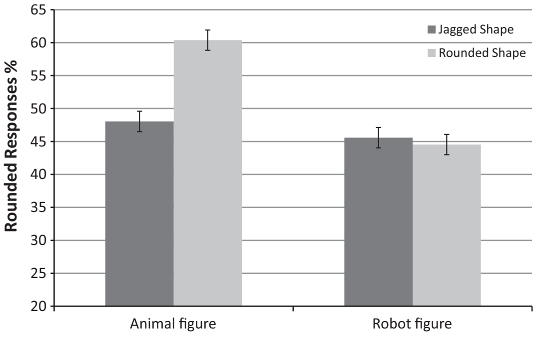

Fig. 4. Experiment 2 – interaction Figure Type × Figure Shape.

Given the different results of the analysis on the first presentation of each picture and on the overall data of Experiment 2, we believed it might be useful to examine the two groups individually. We therefore performed two separate 2 × 2 ANOVA, with the within factors Figure Type (Animal vs. Robot) and Figure Shape (Rounded vs. Jagged), on the data collected at the first appearance of each picture. Fisher's LSD *post hoc* tests were conducted on significant interactions.

The analysis of the Children's responses to the first appearance of each picture did not show any main effects or interactions. The analysis of the Adults, instead, showed a significant main effect of Figure Type with item as random factor, an almost significant main effect of Figure Shape in both analyses, while the interaction was reliable with subject as random factor. The expected main effect of Figure Type was significant only with item as random factor, $F_1(1, 23) = 1.67$, $MSe = 2272.43$, $p = .21$; $F_2(1, 5) = 8.47$, $MSe = 115.21$, $p < .05$ (Animal $M = 57.39\%$, Robot $M = 44.77\%$), whereas Figure Shape was marginally significant in both analyses, $F_1(1, 23) = 3.33$, $MSe = 354.57$, $p < .09$; $F_2(1, 5) = 6.18$, $MSe = 48.18$, $p < .06$, due to Rounded shapes ($M = 54.65\%$) being more often assigned to a rounded sounding name than Jagged shapes ($M = 47.64\%$). The interaction between Figure Type and Figure Shape was reliable in the by-subjects analysis, $F_1(2, 46) = 6.03$, $MSe = 339.62$, $p < .05$; $F_2(2, 20) = 2.07$, $MSe = 240.41$, $p = .21$, showing a modulation of the category especially for Animal Rounded shape ($M = 65.56\%$), which received significantly more rounded responses than all other conditions (Animal Jagged shape $M = 49.31\%$; Robot Rounded shape $M = 43.75\%$, Robot Jagged shape $M = 45.97\%$) (LSD, all $ps < .001$).

As in Experiment 1, mixed effect models were again conducted to confirm and clarify results from the ANOVA[2]. Furthermore, to control if the results were related to the pictures ratings, we correlated subjects' mean rating to each

---

[2] Models starting from the most complete one, with Figure Type, Rating and Group as fixed effects, and Item (i.e., each picture), Name Pair (i.e., the two names to choose between), and Participant in interaction with Figure Type, Rating and Group as random effects. The procedure was the same as described in Experiment 1. The final model resulting from this procedure consisted of all fixed effects in interaction, and Item, Name Pair and Participant in interaction with Figure Type and Group as random effects (AIC = 5840, BIC = 5943, Log-likelihood = −2904).

picture and the mean percentage of rounded responses then assigned to that picture. While the correlation was not significant for children ($r = -.03$, $p > .4$), a positive significant correlation was observed for adults ($r = .20$, $p < .001$), providing evidence that adults' responses were related to shape features, while children responses were not as stable and consistent.

Finally, in order to better understand the relation between the results of Experiment 1 and 2, in the Supplementary Materials section we also report an omnibus ANOVA that compared the results from the two adult groups of Experiment 1 and 2. This analysis was conducted on the data collected for the adult group of both experiments to clarify their relations.

Discussion of Experiment 2

The sound-symbolic correspondence between shapes and sounds was confirmed with figures of natural and artificial agents in Experiment 2. It was also marginally present in the adult group when considering only the first presentation of each visual stimulus. Thus, the results seen in Experiment 1 with natural objects and artifacts were replicated with natural and artificial agents.

Besides this "classic" sound-symbolic effect, in the overall analysis we found that label assignment was modulated by the category, thus confirming our hypothesis (see Fig. 4). As to the developmental pattern, in the adults group we found a clear interaction between sound and category, which was not present in children (see Fig. 3). In fact, a sonorant, rounded-sounding word (e.g., *maluma*) was more frequently assigned to an animal, and a strident, sharp-sounding word (e.g., *takete*) was more frequently assigned to a robot. This result shows, using a paradigm never used in studies on categorization, that natural and artificial agents differ also in some general characteristics related to sounds: natural items are associated with smoother sounds in comparison to artificial ones. The results of Experiment 2 also show that the category effect interacted with the sound–shape correspondence. If the classic effect was present for animals, jagged responses were always preferred for robots, independently of their shape (as reported in Fig. 4). Thus, with artificial agents a sound-category correspondence suppressed the classic sound–shape correspondence. This interaction clearly indicates that the modulation of the takete–maluma effect due to the category appeared both in adults and children, as also confirmed by the marginally significant main effect of Figure Type.

However, it could be argued that the effect of category is due to a confounding influence of the figures stimuli, which might be more rounded for animals than for robots. We do not believe, though, that this is so, for two main reasons. Firstly, neither an interaction between Figure Type and Figure Shape nor a triple interaction between Group, Figure Type and Figure Shape were observed in the experimental groups' ratings. Thus, the results of the ratings by the experimental groups do not predict the interaction between Figure Type and Figure Shape that we observed in the labeling task. Moreover, the difference between jagged and rounded robots in the participants' ratings was significant (Robot: Rounded shape $M = 4.45$, Jagged shape $M = 1.97$; LSD $p < .001$), while the effect on the task performance with robots was not. This pattern of responses was not predicted by the corresponding independent groups' ratings (Robot: Rounded shape $M = 4.51$, Jagged shape $M = 2.08$; LSD $p < .001$). Secondly, and most important, the results for mixed effect models clearly showed that both Figure Type and Rating values (i.e., subjective perception of jaggedness), as well as Group, are crucial factors in explaining our data. While we cannot exclude that perceptual differences between the two kinds of figures we chose might have contributed to the performance, we believe that this cannot be the main factor underlying the resulting effects related to the Figure Type factor.

As regards the developmental pattern, our results suggest that the effect emerges with age, confirming the hypothesis that categories of young children are more malleable, since the effects of language are not yet as marked as in adult categories (Sera et al., 2002). This higher malleability of children's categories is further confirmed by the fact that in children the sound–shape correspondence was not present during the first presentation of the stimuli, while for adults it was. This suggests that children were extracting the sound-symbolic relations and learning how to use them during the labeling task. Furthermore, the results of the correlational analysis between children's ratings and performance indicated that also the link between the perception of external entities features related to shape and the ability to implicitly use them in cognitive tasks is still developing. Thus, even if the literature indicates sensitivity to sound-symbolic relations already in very young children (below the age of three years), our results suggest that sound-symbolic correspondences are modulated during development by aspects of experience that are more abstract compared to physical stimuli features such as shape. These aspects include the semantic knowledge for categories.

General discussion

In Experiment 1 we asked a group of adults to choose a name from between two alternatives for figures of every-day objects, i.e., natural objects (e.g., fruit) and artifacts (e.g., kitchen tools). The results showed the "classic" takete–maluma effect: we found a reliable perceptual correspondence between name sound and shape appearance. This result reveals that the effect can be found also with every-day objects. However, no effect of the objects category (natural vs. artificial) was found, even if it was predicted by the ratings of both experimental and independent groups.

In Experiment 2 we selected a subcategory from among the natural and artificial categories, the category of agents, with the aim to render the two contrasting categories of natural and artificial agents more compact and better comparable. Stimuli were figures depicting animals (natural agents) or robots (artificial agents). In order to investigate the development of the effect, both adults and children were tested. This second experiment confirmed the "classic" takete–maluma effect. In addition, an effect of category (natural vs. artificial) was observed too. First, only adults more frequently assigned rounded sounding names to animals than to robots, while for sharp sounding names the opposite was true (Fig. 3). Second, the category also interacted with the takete–maluma effect, as this effect was present only with animals, while with robots a jagged response was always preferred independently of shape (Fig. 4). Importantly, this interaction indicated that this modulation of the takete–maluma effect due to the category appears both in adults and children.

Our results allow us to address the predictions made. First, we demonstrated that the sound–shape correspondence effect emerges with figures representing every-day objects or agents, that is with more ecological stimuli compared to those typically used in the literature. Second, the pairs of names used (Maurer et al., 2006) showed the predicted sound-to-shape symbolic mapping even when the figures were presented one by one (Nielsen & Rendall, 2011). Thus, the effect was found with a paradigm which minimized both the risk of an enhancement of the results as well as the risk that participants could understand the aim of the study. These two considerations confirm and strengthen the effect observed in previous studies (e.g., Maurer et al., 2006).

Third, modulations of the effect by the stimulus category were found in Experiment 2. Specifically, sound–shape correspondences were not observed with robots, which were associated more often with jagged responses despite their actual shape. Importantly, this effect was not predicted by the ratings, where adults and children both judged jagged robots as sharp, and rounded ones as round. One possible reason why we found the effect in Experiment 2, with the more compact and apparently less differentiated category of "agents", and not in Experiment 1, may depend on the special "naming habit" used by children and adults in their interactions with biological agents (e.g., animals), and with any entity presenting animacy cues (e.g., eyes, mouth) and perceived as able to act autonomously (e.g., robots). Indeed, entities perceived as agents are usually renamed during interaction with them: children and adults typically use a special name for their pets (e.g., their cat is not called just "cat") and for their favorite teddy-bear, or robot toy. In contrast, it might seem more difficult to associate novel names to entities endowed only with generic names and not with proper names, such as the objects of Experiment 1. In support of this explanation, research on the mutual exclusivity or lexical contrast constraint (Clark, 1987; Markman, 1989, 1992) has shown that during language acquisition we experience difficulties in using more names, for example a basic and a superordinate one (e.g., "pineapple" and "fruit", respectively), to indicate the same referent.

Fourth, while the takete–maluma effect and its modulation due to the category were stable across ages, the interaction between the produced sound and the category changed with development. Only adults showed the tendency to associate a jagged name with a robot and a rounded name with an animal *independently of the shape*. Children between the 4th and 6th grade, unlike adults, did not show a sound-symbolic correspondence between word and category independently of shape. Thus, even if semantic aspects of natural and artificial kinds have already been learned at their age, there is clear evidence that from 7 year olds on categorical knowledge changes, as studies on categorization and on categorical induction reveal (e.g., Farrar, Raney, & Boyer, 1992; Kalénine, Bonthoux, & Borghi, 2009; Mounoud, Duscherer, Moy, & Perraudin, 2007). Our results extend this literature suggesting that the symbolic relationship between categories and sounds develops later in life in respect to the relation between shapes and sounds. One could speculate that the emergence of sound-symbolic correspondences at a semantic level requires the acquisition of linguistic and cultural aspects related to categories. This is precisely our position, as we explain the result referring to the greater socio-linguistic experience of adults (see also Sera et al., 2002): it is possible that adults have more experience in hearing or actively associating more nouns with agents such as pets and toys. This experience might have led to associations between sound features (i.e., strident, sonorant) and category properties (i.e., animal more rounded, robot more jagged) which go beyond the "classic" sound-symbolic correspondence between shapes and names based only on perceptual aspects of the stimulus. This result confirms that the children's categories are more perceptually grounded than those of adults (for the importance of shape and perceptual grounding in children's categorization, see the literature on the "shape bias", showing that names are extended on the basis of shape similarity, e.g., Landau et al., 1988, 1992; Smith, 2005). This interpretation is in line with the idea that, once the mapping between perceptual aspects and linguistic aspects is established, no grounding is necessary, but that participants can use a shortcut relying on associative knowledge (Barsalou, Santos, Simmons, & Wilson, 2008; Connell & Lynott, 2013; for a discussion see Borghi, Flumini, Cimatti, Marocco, & Scorolli, 2011).

Importantly, the effects of category we observed cannot be merely attributed to perceptual aspects of the selected stimuli, as they do not mirror the subjective judgments on the figures' roundness observed in the ratings. Nonetheless, interesting effects of category emerged in the ratings both for objects and agents: we suggest that these results might indicate an influence of factors such as category and age on the subjective perception of roundness.

One could argue that the difference between children and adults might be due to the fact that the distinction between animate and inanimate entities becomes better differentiated with age. We tend to exclude this interpretation for at least two reasons. The first is that we found that the sound–shape correspondence was modulated by category in both adults and children. Only the interaction between sound and category independently of shape was present in adults but not in children. Thus, the difference between adults and children does not seem to involve the development of categories, but rather the development of associations between categories and linguistic labels. The second is that the literature has shown the ability to distinguish between artifacts and natural objects as emerging rather early: some studies have demonstrated using habituation, preferential looking or other methods, that even prelinguistic infants are already able to differentiate these two macro-categories (e.g., Behl-Chadha, 1996; Mandler, 2004; Quinn & Johnson, 2000).

Taken as whole, these results extend and strengthen evidence on the takete–maluma effect, indicating that sound-symbolic correspondences may arise at either perceptual or semantic levels. In general, our results bolster the hypothesis of a natural relation between the structure of words and the meanings they convey, extending previous findings of the sound-symbolic literature to entities taken from every-day life. Moreover, they clearly suggest a mutual influence between the naturally biased sound–shape correspondences and the cultural and linguistic learning by which categorization is socially determined. Indeed, not only has a possible iconic relation between the name and its referent been confirmed, but also a sound-symbolic correspondence with semantic aspects of the referent categories too, in interaction with previously acquired knowledge. As our results were collected through presenting real stimuli and invented words, a future direction for our research could be that of conducting experiments in which the presentation is the reverse, that is invented rounded/jagged figures are presented with real words (strident/sonorant, and referring to natural objects/artifact). In such a setting it is very likely that sound–shape correspondences would be seen to be facilitated or interfered with on the basis of word meaning.

Finally, our results may have implications for the classic question about the arbitrariness of verbal language discussed in the introduction. Indeed, they could offer an impetus to speculate about a possible origin of contemporary conventionalized lexicons from more iconic ones, in keeping with the perspectives on cognition which hypothesize a direct, natural line of evolution from gestures to speech (e.g., Corballis, 2002, 2009; Flumini, 2014; Gallese, 2008; Gentilucci & Corballis, 2006; Rizzolatti & Craighero, 2004).

Acknowledgments

This research was supported by the EU funded Projects "ROSSI – Emergence of communication in RObots through Sensorimotor and Social Interaction", contract No. FP7-STREP-216125.

Appendix

See Tables 1–3 and Figs. A1 and A2.

Table 1
Word pairs.

Words pairs	
Jagged names	Rounded names
KIKI	BOBA
KUTI	BAMA
TITI	GOGA
TUKITI	MABUMA

Table 2
Experimental groups ratings.

			Experimental groups					
			Exp. 1 adults		Exp. 2 adults		Exp. 2 children	
Item nr.	Type	Shape	Mean rating	Stand. dev.	Mean rating	Stand. dev.	Mean rating	Stand. dev.
1	Artificial	Rounded	6.04	1.16	5.00	1.64	4.25	1.15
2			4.25	1.11	4.33	1.86	3.54	1.79
3			5.25	0.79	5.17	1.52	4.88	1.33
4			4.21	0.88	4.75	1.80	4.13	0.95
5			4.00	0.88	4.17	1.46	4.04	0.91
6			4.58	1.10	5.17	1.20	4.04	1.20
7		Jagged	2.75	0.94	3.63	1.76	2.71	1.33
8			2.50	1.18	1.58	1.02	1.88	1.30
9			1.83	0.92	2.83	1.61	1.88	1.26
10			1.88	1.15	2.17	1.40	1.29	0.55
11			1.21	0.51	1.42	0.83	1.08	0.41
12			1.88	1.03	2.04	1.16	1.08	0.41
1	Natural	Rounded	6.04	0.62	5.21	1.38	5.42	0.88
2			5.21	0.93	6.29	1.04	6.08	1.02
3			5.79	0.83	4.96	1.33	6.00	0.93
4			6.75	0.44	5.25	1.45	5.54	0.93
5			5.21	1.14	5.46	1.38	5.63	1.13
6			6.38	0.77	5.29	1.60	5.71	1.08
7		Jagged	3.92	1.10	3.21	1.96	2.92	0.88
8			2.38	0.92	3.04	1.40	3.00	0.93
9			2.42	1.14	2.75	1.67	2.29	1.04
10			1.42	0.72	3.17	1.52	3.42	0.78
11			3.08	1.25	2.54	1.59	2.25	0.94
12			1.88	0.90	3.13	1.85	2.88	1.08

Table 3
Independent groups ratings.

			Independent groups					
			Exp. 1 adults		Exp. 2 adults		Exp. 2 children	
Item nr.	Type	Shape	Mean rating	Stand. dev.	Mean rating	Stand. dev.	Mean rating	Stand. dev.
1	Artificial	Rounded	5.83	1.05	4.42	1.18	5.38	1.64
2			5.38	0.71	3.17	1.37	3.38	1.91
3			5.96	0.91	5.29	1.12	5.83	1.47
4			4.71	1.37	4.71	1.12	5.29	1.37
5			3.96	1.20	3.75	1.26	3.38	1.35
6			5.25	0.94	4.63	1.79	4.92	1.35
7		Jagged	2.67	0.96	3.58	1.28	3.83	1.71
8			2.54	0.93	1.54	0.66	1.13	0.45
9			2.17	0.87	1.83	0.82	3.33	1.79
10			1.63	0.65	1.79	0.78	2.00	1.32
11			1.67	0.76	1.25	0.53	1.54	1.38
12			1.67	1.01	1.25	0.53	1.88	1.62
1	Natural	Rounded	6.25	0.74	5.50	0.78	4.88	1.39
2			5.79	0.83	5.88	0.85	5.63	1.56
3			6.25	0.74	5.46	0.93	6.08	1.18
4			6.71	0.55	5.17	1.52	5.58	1.74
5			5.88	0.90	5.46	1.50	5.21	1.64
6			6.29	1.12	5.25	1.36	5.42	1.28
7		Jagged	3.21	1.25	2.04	0.86	2.42	1.50
8			2.33	0.92	2.46	0.98	3.29	1.81
9			2.83	1.05	2.04	1.20	2.25	1.67
10			2.08	1.18	3.00	0.98	3.79	1.56
11			2.92	0.93	1.92	1.06	2.38	1.93
12			1.92	0.83	2.50	1.10	4.38	1.64

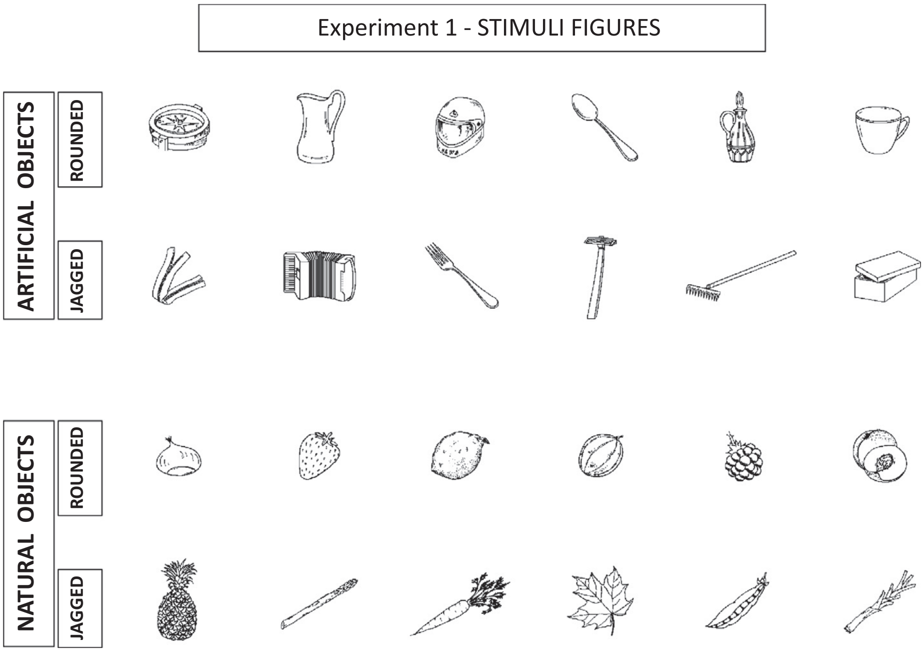

Fig. A1. Experiment 1 – stimuli figures.

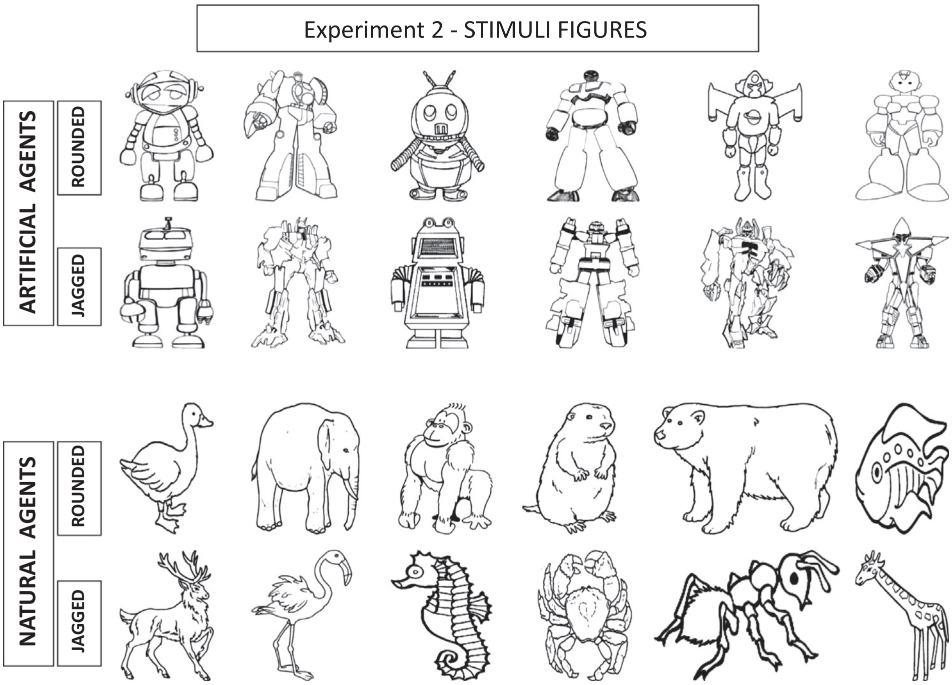

Fig. A2. Experiment 2 – stimuli figures.

A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.jml.2014.06.004.

References

Arata, M., Imai, M., Sotaro, K., Guillaume, T., & Okada, H. (2010). Perception of sound symbolism in 12 month-old infants: An ERP study. *Neuroscience Research, 68S*, 223–334.

Asano, M., Kitajo, K., Arata, M., Guillaume, T., Sotaro, K., Okada, H., et al. (2011). Preverbal infants' sensitivity to sound symbolism: An EEG study. *Neuroscience Research, 71S*, 408–416.

Backscheider, A. G., Gelman, S. A., Martinez, I., & Kowieski, J. (1999). Children's use of different information types when learning homophones and nonce words. *Cognitive Development, 14*(4), 515–530. http://dx.doi.org/10.1016/S0885-2014(99)00017-9.

Baronchelli, A., Gong, T., Puglisi, A., & Loreto, V. (2010). Modeling the emergence of universality in color naming patterns. *Proceedings of the National Academy of Science, 107*(6), 2403–2407.

Barsalou, L. W. (2008). Grounded cognition. *Annual Review of Psychology, 59*, 617–645.

Barsalou, L. W, Santos, A., Simmons, W. K, & Wilson, C. D (2008). Language and simulation in conceptual processing. In M. De Vega, A. M. Glenberg, & A. C. Graesser (Eds.). *Symbols, embodiment, and meaning* (pp. 245–283). Oxford: Oxford University Press.

Behl-Chadha, G. (1996). Basic-level and superordinate-like categorical representations in early infancy. *Cognition, 60*, 105–141.

Borghi, A. M., Bonfiglioli, C., Lugli, L., Ricciardelli, P., Rubichi, S., & Nicoletti, R. (2007). Are visual stimuli sufficient to evoke motor information? Studies with hand primes. *Neuroscience Letters, 411*, 17–21.

Borghi, A. M., Flumini, A., Cimatti, F., Marocco, D., & Scorolli, C. (2011). Manipulating objects and telling words: A study on concrete and abstract words acquisition. *Frontiers in Cognition, 2*, 15. http://dx.doi.org/10.3389/fpsyg.2011.00015.

Borghi, A. M., & Pecher, D. (2011). Introduction to the special topic embodied and grounded cognition. *Frontiers in Psychology, 2*, 187. http://dx.doi.org/10.3389/fpsyg.2011.00187.

Brown, R. W., Black, A. H., & Horowitz, A. E. (1955). Phonetic symbolism in natural languages. *Journal of Abnormal and Social Psychology, 50*, 388–393.

Burling, R. (1999). Motivation, conventionalization, and arbitrariness in the origin of language. In B. J. King (Ed.), *The origins of language: What nonhuman primates can tell us* (pp. 307–350). Santa Fe, NM: School of American Research Press.

Clark, E. V. (1987). The principle of contrast: A constraint on language acquisition. In B. MacWhinney (Ed.), *Mechanisms of language acquisition* (pp. 1–33). Hillsdale, NJ: Lawrence Erlbaum Associates.

Connell, L., & Lynott, D. (2013). Flexible and fast: Linguistic shortcut affects both shallow and deep conceptual processing. *Psychonomic Bulletin and Review*, 1–9.

Corballis, M. C. (2002). *From hand to mouth: The origins of language*. Princeton, NJ: Princeton University Press.

Corballis, M. C. (2009). The evolution of language. *Annals of the New York Academy of Sciences, 1156*, 19–43.

Davis, R. (1961). The fitness of names to drawings: A cross-cultural study in Tanganyika. *British Journal of Psychology, 52*, 259–268.

Farrar, M. J., Raney, G. E., & Boyer, M. E. (1992). Knowledge, concepts, and inferences in childhood. *Child Development, 63*(3), 673–691.

Fay, N., Garrod, S., Roberts, L., & Swoboda, N. (2010). The interactive evolution of human communication systems. *Cognitive Science, 34*, 351–386.

Flumini, A. (2014). *Cognition in context: Evidence on affordances and verbal language*. Doctoral thesis retrievable from <http://amsdottorato.unibo.it/6449/>.

Gainotti, G. (2000). What the locus of brain lesion tells us about the nature of the cognitive defect underlying category-specific disorders: A review. *Cortex, 36*(4), 539–559.

Gallese, V. (2008). Mirror neurons and the social nature of language: The neural exploitation hypothesis. *Social Neuroscience, 3*, 317–333.

Garrod, S., Fay, N., Lee, J., Oberlander, J., & MacLeod, T. (2007). Foundations of representation: Where might graphical symbol systems come from? *Cognitive Science, 31*(6), 961–987.

Gebels, G. (1969). An investigation of phonetic symbolism in different cultures. *Journal of Verbal Learning and Verbal Behavior, 8*, 310–312.

Gentilucci, M., & Corballis, M. C. (2006). From manual gesture to speech: A gradual transition. *Neuroscience Biobehavioural Review, 30*, 949–960.

Harnad, S. (1990). The symbol grounding problem. *Physica D, 42*, 335–346.

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

Humphreys, G. W., & Forde, E. M. (2001). Hierarchies, similarity, and interactivity in object recognition: "Category-specific" neuropsychological deficits. *Behavioral and Brain Sciences, 24*(3), 453–476. discussion 476–509.

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 Education in the World, 17*, 53–78.

Jaeger, T. F. (2008). Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of memory and language, 59*(4), 434–446.

Kalénine, S., Bonthoux, F., & Borghi, A. M. (2009). How action and context priming influence categorization: A developmental study. *British Journal of Developmental Psychology, 27*, 717–730.

Kantartzis, K., Imai, M., & Kita, S. (2011). Japanese sound-symbolism facilitates word learning in English-speaking children. *Cognitive Science, 35*(3), 575–586.

Köhler, W. (1929). *Gestalt psychology*. New York: Liveright.

Köhler, W. (1947). *Gestalt psychology* (2nd ed.). New York: Liveright.

Koriat, A., & Levy, I. (1979). Figural symbolism in Chinese ideographs. *Journal of Psycholinguistic Research, 8*(4), 353–365.

Kovic, V., Plunkett, K., & Westermann, G. (2009). The shape of words in the brain. *Cognition, 114*, 19–28.

Kunihira, S. (1971). Effects of expressive voice on phonetic symbolism. *Journal of Verbal Learning and Verbal Behavior, 10*, 427–429.

Landau, B., Smith, L. B., & Jones, S. (1988). The importance of shape in early lexical learning. *Cognitive Development, 3*(3), 299–321.

Landau, B., Smith, L. B., & Jones, S. (1992). Syntactic context and the shape bias in children's and adults' lexical learning. *Journal of Memory and Language, 31*(6), 807–825.

Laurence, E., & Margolis, S. (Eds.). (2007). *Creations of the minds. Theories of artifacts and their representation*. Oxford: Oxford University Press.

Lotto, L., Dell'Acqua, R., & Job, R. (2001). Le figure PD/DPSS. Norme di accordo sul nome, tipicità, familiarità, età di acquisizione e tempi di denominazione per 266 figure. *Giornale Italiano di Psicologia, 28*, 193–207.

Lupyan, G. (2012). Linguistically modulated perception and cognition: the label-feedback hypothesis. *Frontiers in Psychology, 3*.

Mandler, J. M. (1992). How to build a baby: II. Conceptual primitives. *Psychological Review, 99*, 587–604.

Mandler, J. M. (2004). *The foundations of mind: The origins of conceptual thought*. New York: Oxford University Press.

Markman, E. M. (1989). *Categorization and naming in children: Problems of induction*. Cambridge, MA: MIT Press.

Markman, E. M. (1992). Constraints on word learning: Speculations about their nature, origins and domain specificity. In M. R. Gunnar & M. P. Maratsos (Eds.), *Modularity and constraints in language and cognition: The Minnesota symposium on child psychology* (pp. 59–101). Hillsdale, NJ: Lawrence Erlbaum Associates.

Martin, A. (2007). The representation of object concepts in the brain. *Annual Review of Psychology, 58*, 25–45.

Maurer, D., Pathman, T., & Mondloch, C. J. (2006). The shape of boubas: Sound–shape correspondences in toddlers and adults. *Developmental Science, 9*(3), 316–322.

Merleau-Ponty, M. (1945/1962). *Phenomenology of perception*. London: Routledge.

Mounoud, P., Duscherer, K., Moy, G., & Perraudin, S. (2007). The influence of action perception on object recognition: A developmental study. *Developmental Science, 10*(6), 836–852.

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*(2), 115–124.

Nygaard, L. C., Herold, D. S., & Namy, L. L. (2009a). The semantics of prosody: Acoustic and perceptual evidence of prosodic correlates to word meaning. *Cognitive Science, 33*, 127–146.

Nygaard, L. C., Herold, D. S., & Namy, L. L. (2009b). Sound to meaning correspondences facilitate word learning. *Cognition, 112*, 181–186.

Parault, S. J. (2006). Sound symbolic word learning in context. *Contemporary Educational Psychology, 31*, 228–252.

Parise, C., & Pavani, F. (2011). Evidence of sound symbolism in simple vocalizations. *Experimental Brain Research, 214*(3), 373–380.

Perniss, P., & Vigliocco, G. (in press). The bridge of iconicity: from a world of experience to the experience of language. *Philosophical Transactions of the Royal Society B.*

Perniss, P., Thompson, R. L., & Vigliocco, G. (2010). Iconicity as a general property of language: Evidence from spoken and signed languages. *Frontiers in Psychology, 1*.

Pietrandrea, P. (2002). Iconicity and arbitrariness in Italian Sign Language. *Sign Language Studies, 2*, 296–321.

Pizutto, E., & Volterra, V. (2000). Iconicity and transparency in sign languages: A cross-linguistic crosscultural view. In K. Emmorey & H. Lane (Eds.), *The signs of language revisited: An anthology to Honor Ursula Bellugi and Edward Klima* (pp. 261–286). Mahwah, NJ: Lawrence Erlbaum Associates.

Quinn, P. C., & Johnson, M. H. (2000). Global-before-basic object categorization in connectionist networks and 2-month-old infants. *Infancy, 1*, 31–46.

Rakison, D. H., & Oakes, L. M. (Eds.). (2003). *Early category and concept development: Making sense of the blooming, buzzing confusion*. New York: Oxford University Press.

Rakison, D. H., & Poulin-Dubois, D. (2001). Developmental origin of the animate–inanimate distinction. *Psychological Bulletin, 127*(2), 209.

Rakison, D. H., & Poulin-Dubois, D. (2002). You go this way and I'll go that way: Developmental changes in infants' detection of correlations among static and dynamic features in motion events. *Child Development, 73*, 682–699.

Ramachandran, V. S., & Hubbard, E. M. (2001). Synaesthesia – A window into perception, thought, and language. *Journal of Consciousness Studies, 8*, 3–34.

Reilly, J., Westbury, C., Kean, J., & Peelle, J. E. (2012). Arbitrary symbolism in natural language revisited: When word forms carry meaning. *PLoS ONE, 7*(8), e42286. http://dx.doi.org/10.1371/journal.pone.0042286.

Rizzolatti, G., & Craighero, L. (2004). The mirror-neuron system. *Annual Review of Neuroscience, 27*, 169–192.

Roversi, C., Borghi, A. M., & Tummolini, L. (2013). A marriage is an artefact and not a walk that we take together: An experimental study on the categorization of artefacts. *Review of Philosophy and Psychology*, 1–16.

Sapir, E. (1929). A study in phonetic symbolism. *Journal of Experimental Psychology, 12*, 225–239.

Saussure, F. (1916/1966). *Course in general linguistics*. New York: McGraw Hill.

Sera, M. D., Elieff, C., Forbes, J., Burch, M. C., Rodríguez, W., & Dubois, D. P. (2002). When language affects cognition and when it does not: An analysis of grammatical gender and classification. *Journal of Experimental Psychology: General, 131*(3), 377.

Smith, L. B. (2005). Action alters shape categories. *Cognitive Science, 29*, 665–679.

Spector, F., & Maurer, D. (2008). The color of Os: Naturally biased associations between shape and color. *Perception, 37*, 841–847.

Steels, L. (2011). Modeling the cultural evolution of language. *Physical Life Review, 8*(4), 339–356.

Thompson, R. L., Vinson, D. P., & Vigliocco, G. (2009). The link between form and meaning in American Sign Language: Lexical processing effects. *Journal of Experimental Psychology: Language, Memory, and Cognition, 35*, 550–557.

Thompson, R. L., Vinson, D. P., & Vigliocco, G. (2010). The link between form and meaning in British Sign Language: Effects of iconicity for phonological decisions. *Journal of Experimental Psychology: Language, Memory, and Cognition, 36*, 1017–1027.

Westbury, C. (2005). Implicit sound symbolism in lexical access: Evidence from an interference task. *Brain and Language, 93*, 10–19.

Zipf, G. K. (1949). *Human behavior and the principle of least effort: An introduction to human ecology*. Cambridge, MA: Addison-Wesley.