Japanese sound symbolism facilitates word learning in English speaking children

Katerina Kantartzis (kfk495@bham.ac.uk)
School of Psychology, University of Birmingham, Edgbaston, Birmingham, B15 2TT UK

Sotaro Kita (s.kita@bham.ac.uk)
School of Psychology, University of Birmingham, Edgbaston, Birmingham, B15 2TT UK

Mutsumi Imai (imai@sfc.keio.ac.jp)
Department of Environmental Information, Keio University, Shonan-Fujisawa

Abstract

Sound symbolism is the non-arbitrary link between the sound of a word and its meaning. Imai et al. (2008) showed that Japanese speaking children benefited from the presence of sound symbolism when learning novel verbs. However, Japanese is a language rich in sound symbolic elements, but English on the other hand is not. The present study investigated whether English-speaking three-year olds can benefit from a cross-linguistically recognisable sound symbolic link between a novel word and its referent in word learning. The children were taught a novel verb and asked to generalise this to a new situation. It was found that English-speaking 3-year-olds peformed better when the taught verb and the target action matched sound symbolically. This suggests that sound symbolism can facilitate word learning regardless of the language the children are learning.

Keywords: language acquisition; sound symbolism; verb; semantics; motion.

Introduction

Learning words presents a challenge to children because it is not trivial to map the phonological form of a word to its meaning (Quine, 1960). For instance, the parent may point to a rabbit and says to a child, "Look, a rabbit!", from the viewpoint of the child who hears the word "rabbit" for the first time, the word could be referring indeed to a rabbit or alternatively to attributes of the rabbit (e.g., colour), to parts of the rabbit (e.g., the ears), to the action the rabbit is performing, and etc. There are various proposals in the literature as to how children narrow down possible meanings of a novel word. For example, children may assume that a word tends to refer to the whole object (e.g., a rabbit) or the whole action (e.g., hopping) rather than its respective parts (e.g., the ear, leg extension in hopping) (e.g., Markman & Wachtel, 1988; Clark, 1993). In the case of a verb, children may use the syntactic frame in which it is embedded (e.g., transitive vs. intransitive) to infer its meaning (e.g., Naigles & Kako, 1993). This study investigated another word learning mechanism based on intrinsic sound-meaning links, known as sound symbolism. This mechanism can complement other mechanisms because sound symbolism may help children to find the relevant aspect of reality in a fine-grained way. Sound symbolism, for example, can differentiate subtly different manners of locomotion (Hamano,1998; Matsumoto, 1997). Due to the intrinsic sound-meaning link, sound symbolism allows children to focus on the particular manner of locomotion the novel verb is referring to. An example of this is, if the parent is teaching the child the verb “hopping” while pointing to a rabbit which is hopping very fast. In this situation the novel verb “hopping” could be referring a number of things: for example the fast movement or the movement by a rabbit or the manner of movement (which would be the correct interpretation). However, if the word “hopping” sound symbolically matches the referent, namely hopping manner of movement, then the child is able to focus on the relevant part of the world the adult is referring to, and therefore potentially learn the verb more successfully. Other proposed bootstrapping mechanisms for word learning in the literature cannot help in making such fine differentiations among words belonging to the same syntactic class

It has been shown that sound symbolism can help Japanese speaking children to map a novel verb to its referent action (Imai, Kita, Nagumo, & Okada, 2008). Japanese speaking three-year-olds were taught a novel verb with a video clip of a person walking in a particular manner. Consistent with preceding research (Imai, Haryu, Okada, 2005; McGuire, 2002), three-year-olds failed to generalize the verb to a new instance of the same action performed by a different actor when the novel word did not have a sound symbolic relation to the referent. However, they succeed in the task when the novel word had a sound symbolic relation to the referent. This finding lead the authors to propose the "sound symbolism bootstrapping hypothesis", which states that sound symbolism can help children single out the referent of a novel word in the complex reality.

However, Imai et al.’s results leave an important question unanswered. Japanese is a language with a very rich inventory of sound symbolic words (Hamano, 1998; Kita, 1997, 2001) along with many other non-Indo-European languages (see Hinton, Nichols & Ohala, 1994; Voeltz & Killian-Hatz, 2001). For example, a midsize dictionary of Japanese sound-symbolic words (Atoda & Hoshino , 1995) lists more than 1700 entries. However, Indo-European languages, including English, do not have a

large inventory of sound symbolic words (Newmeyer, 1992). Japanese children thus may have an advantage in that they have much experience with sound symbolism, as they are learning a language with a large inventory of sound symbolic words that are frequently used by adults and by three-year-olds (e.g., Allen et al. 2007). For children to take advantage of sound symbolism in word learning, is extensive general experience with sound symbolic words necessary? Alternatively, is sound symbolism beneficial for all children regardless of whether their language has rich sound symbolic lexicon? Does sound symbolism in novel words scaffold English children as well? If this is the case, we have strong evidence that sound symbolism functions as a universal constraint for word learning.

The present study was conducted to address this question by investigating whether English-speaking three-year olds perform better in a word learning task when novel verbs have a cross-linguistically recognisable sound symbolic link to the referent action. Cross-linguistically recognisable sound symbolism is well-documented, even for English speakers. Kohler (1929) found that the novel word "baluma" was considered to be an appropriate label for a round object, and the word "takete" for a pointed object. This type of sound symbolism was recognized both by English speaking children (Davis, 1961; Maurer, Pathman, & Mondloch, 2006) and by Kitongwe speaking children who lived in an isolated part of Tanzania (Davis, 1961). Similarly, adult English speakers with no knowledge of Japanese could correctly guess some aspects of the meaning of Japanese sound symbolic words (Imai, Kita, Nagumo, & Okada, 2008; Iwasaki, et al., 2008). Thus, it is possible that sound symbolism inherent in Japanese sound symbolic words that is recognisable by English adult speakers may facilitate word learning in English speaking children. In contrast, if rich experience with sound symbolic words would be necessary to take advantage of sound symbolism in novel verb learning, the sound symbolism bootstrapping effect found in Japanese children would not be found in English-speaking children.

Current Study

English-speaking children were taught a novel word and then asked to generalize it to a new situation with the same action but a different actor. There were three conditions: the sound symbolic match condition, in which the novel verbs were sound symbolically related to the referent action, and two control conditions in which the novel verbs were not sound symbolically related to the referent action (see below). We used sound symbolic words that were created on the basis of Japanese sound symbolic words and were cross linguistically recognisable (see below for verification). If all children, regardless of the language they are learning, have the capacity to use sound symbolism to scaffold word learning, then English-speaking three-year-olds should perform better in the sound symbolic match condition than in the two control conditions.

Method

Participants and Design

Forty-five monolingual English-speaking 3-year-olds (*M*=41.57 months, range 36-48 months, 20 boys, 25 girls) were recruited from nurseries around Birmingham, UK, with a prior parental consent. The participants were randomly assigned to one of the following three conditions, all of which followed the same structure of a training phase followed by a test phase.

Sound symbolic match condition. Fifteen children in this condition (Mean age= 41.73, range=33-48 M, 9 girls), were presented with a novel verb and a training video showing an actor carrying out the target action (training phase). The verb sound-symbolically matched the target action. Then, they were presented with two videos simultaneously, one in which the action was the same as the training video but the actor was different (target video), the other in which the actor was the same but the action was different (distractor video), and asked in which video the actor was carrying out the action referred to by the newly taught verb (test phase). The newly taught verb did not sound symbolically match the action in the distractor video (see the Stimuli section for verification). The action used in the target or distractor video did not re-appear as the action in the target or distractor video for another word.

Neutral baseline condition. Fifteen children (Mean age = 42.47M, range 35-48, 8 girls) were tested in this condition. This condition provided a baseline for 3-year-olds' performance in this verb generalization task when the newly taught verb did not sound symbolically match the target action or the action in the distractor video. The verbs were presented in a form which resembled typical English verbs. The training videos, the target videos and the distractor videos were all identical to those in the sound symbolic match condition.

Sound symbolic mismatch condition. Fifteen children (mean age = 40.53M, range 33-47M, 8 girls) were taught the same set of words as in the s*ound symbolic match* condition. However, the newly taught verb did not sound symbolically match the target action. Instead, the newly taught verb sound symbolically matched the action in the distractor video. That is, the pairs of videos used in the test phase were identical to those in the sound symbolic match condition, but what counted as the target in the sound symbolic match condition now became the distractor and vice versa, for this condition. Accordingly, the training videos differed from those in the sound symbolic match condition because the target actions were different. This condition would allow us to eliminate alternative explanations for the predicted finding that children would

perform better in the sound symbolic match condition than in the neutral baseline condition.

Stimuli

The material used for the experiment consisted of word-action combinations. The actions in the word-action combinations were various manners of walking. Eight novel words were prepared to be associated with the four target actions. Four of them were altered versions of Japanese mimetics (batobato, nosunosu, chokachoka and tokutoku) used in the sound symbolic match and mismatch conditions. The other four were non-words that have the structure of typical monosyllabic English verbs (bretting, blegging, blicking and truffing).

The four novel sound symbolic words were paired with sound symbolically matching actions: batobato = a large energetic movement, arms are swinging back and forward outstretched, while legs are making large leaping movement, chokachoka = walking quickly in very small steps with the arms with bent elbows swinging quickly; nosunosu (see below) = walking slowly in large steps with bent knees and the hands on knees (see the upper pane of Figure 1); tokutoku = a small shuffling movement, with straight arms rigidly at the side and legs moving very slightly and rigidly. The four words were also paired with an action that does not sound symbolically match the words, which was used in the distractor video in the sound symbolic match condition (see the lower pane of Figure 1 for an example).


Target


Distractor

Figure 1. An example of a target video and distractor videos used in the test phase of the sound symbolic match condition for the novel word "nosu-nosu".

A pre-test was conducted to check if these verbs have the assumed sound symbolic relationships with the actions in the four target videos and four distractor videos in the three conditions described above. The novel sound symbolic words (as an audio recording) were presented along with the eight videos to 21 English speaking adults (without any knowledge of Japanese) and 15 Japanese speaking adults. Each novel word and video were presented in pairs and the participants were asked to rate how well they thought each word-action combination matched on a scale from 1 (did not match) to 7 (matches very well). The mean rating was significantly higher for the target videos than for the distractor videos for English speakers (targets, M = 4.4, SD = 1.02; distractors, M = 3.50, SD = 1.04), $t(20) = -3.8$, $p < .001$, $d = 7.67$ and Japanese speakers (targets, $M = 5.71$, $SD = .66$; distractors, $M = 2.06$, $SD = .78$). $t(14) = -14.7$, $p < .001$, $d = 10.79$.

For the neutral baseline condition, four novel, English-sounding verbs were prepared. The degree of the sound-action match was also tested by 20 English speaking adults. Again words and actions were presented in pairs, one at a time. The results ensured that the novel verbs did, indeed, not match any of the actions in the target or distractor videos: The degree of match was judged to be poor for both the target videos and the distractor videos (targets, M=3.81, SD=.57, distractors, M= 3.63, SD=.79) $t(19)$=-1.0, and there was no difference between the two.

Procedure

Each child was tested individually in a quiet room of the nursery. Two warm-up trials took place first, using familiar nouns to establish the procedure, of indicating the referent of a word by pointing. Then, a practice trial with a familiar verb preceded the main experiment to ensure that the children understood the training-test procedure. The practice trials followed the same procedure as the experiment proper. During the training phase in the *sound symbolic match* condition and the *sound symbolic mismatch* condition, children saw an action video on a laptop computer, and heard a target verb embedded in the sentence "Look! He is doing X" ("doing X" represents the novel sound symbolic verb). In the test phase, the experimenter asked them to indicate their response by pointing to one of the two action videos on the screen, saying "which one is doing X?" The procedure for the *neutral baseline* condition was identical to the other two conditions, except that the novel verb was introduced without the light verb "do" (i.e., "He is Xing").

Results and Discussion

When a child correctly extended the novel verb on the basis of the same action, the response was coded as correct. For each child, the proportion of correct responses out of the four trials was calculated and served as the dependent variable. As we expected, the children performed differently across the three groups, $F(2, 42)$= .388 p<.05, $\eta 2$= .161 (See Figure 2). More importantly, children more successfully extended novel verbs to new situations based on the identity of the action when the word sound symbolically matched the action than when the word did not sound symbolically match the action, as the children in the sound symbolic match condition performed better

than those in the sound symbolic mismatch condition or in the neutral baseline condition (Fisher's LSD, both p's <.05).

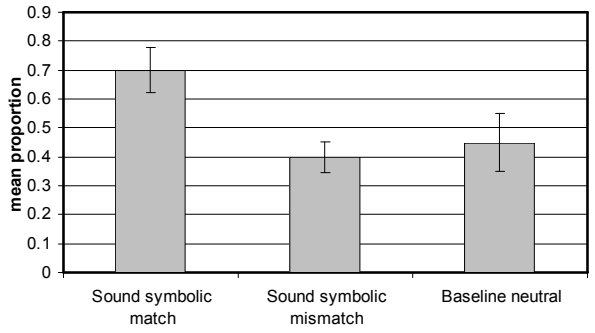

Figure 2. Mean proportion of correct responses given in the sound symbolic match, sound symbolic mismatch and neutral baseline neutral conditions in the verb generalisation task. The error bars represent the standard error of the mean for each condition.

Consistent with the previous findings (Imai, Haryu, & Okada, 2005; Imai et al. 2008; Maguire et al. 2002), the children in the two control conditions failed to generalize novel verbs to the same action in the two control conditions, $t(14)$=-1.87 (the sound symbolic mismatch condition), and $t(14)$=-.49 (the neutral baseline condition). In sharp contrast, the children in the sound symbolic match condition successfully generalized the novel verb, $t(14)$=2.57, p<.05. d=.663.

The performance in the sound symbolic mismatch condition ruled out two possible alternative interpretations. Firstly, one might suggest that children in the sound symbolic match condition were merely detecting sound symbolism at the test phase. In other words, the results may not have reflected the success of verb generalization but reflected success in detecting sound symbolism between the word and the action at test. However, the fact that the children in the sound symbolic mismatch condition did not select the sound symbolically matching distractor suggests that this alternative is unlikely. Secondly, one might also question whether the sentence structure ("doing X") or the reduplication of the sound symbolic verbs caused good performance in the sound symbolic match condition. These possibilities can also be ruled out, because the children were presented with the same set of novel sound symbolic verbs in both sound symbolic match and mismatch conditions, but only the latter group performed at chance and the difference between the two groups was significant.

General discussion

The present research demonstrated that English-speaking children performed better in a verb generalisation task when the novel verb sound-symbolically matched the referent action than when it did not. The sound symbolism in the stimuli was derived from Japanese sound symbolic words, and the pretest of the stimuli showed that adult native speakers of Japanese and English recognized the sound symbolism. As the English adults did not have any knowledge of Japanese, the sound symbolism in the stimulus is potentially universal. We found that this cross-linguistically recognisable (potentially universal) sound symbolism helped the English-speaking children to focus on the relevant part of the world as the intended referent of a novel word, and therefore fostered novel verb learning. In other words, the sound symbolism can help constrain the meaning of the word which can be highly ambiguous in the environment (Quine, 1960) even for children who are learning a language which is not rich in a sound symbolic lexicon.

It has been known that children can detect universally recognisable sound-symbolic links between word form and meaning (Davis, 1961; Maurer, et al., 2006), but this study shows that they can utilize sound symbolism in word learning. "Sound symbolic bootstrapping" may be a mechanism that children can use when learning new words, in addition to other strategies discussed in the literature such as the whole-object bias (e.g., Markman & Wachtel, 1988; Clark, 1993) and syntactic bootstrapping (Naigles & Kako 1990). By incorporating the meaning in the sound of the word, the relevant part of the world which the novel sound symbolic word refers to is made obvious to the child.

Sound symbolic bootstrapping has an advantage of over other word learning constraints proposed in the literature. Sound symbolism can differentiate events and states in a very fine grained way. For example, Japanese sound symbolic words differentiate types of motion such as "chokochoko" (walking in a short, rapid movement), "noshinoshi" (walking heavily), "pyonpyon" (hopping repeatedly) and "gorogoro" (a heavy object rolling repeatedly). Thus, sound symbolic bootstrapping may be able to facilitate fine-grained distinctions of events and states that other word learning constraints cannot.

The finding of the present study, along with the analogous finding for the Japanese children (Imai et al., 2008), suggests that sound symbolism can serve as a language-universal constraint for young children in word learning. The universality of this phenomenon would point to the possibility that children are innately disposed to develop an ability to detect universal sound symbolism and use it for word learning. The fact that a large number of geographically dispersed and historically unrelated languages having a large number of sound symbolic words (Hinton, Nichols & Ohala, 1994; Voeltz & Killian-Hatz, 2001) and some aspects of sound symbolism are crosslinguistically recognisable (Davis 1961; Imai, Kita, Nagumo, & Okada, 2008; Iwasaki, et al., 2008; Maurer, Pathman, & Mondloch, 2006) also provide further support for the biological basis for sound symbolism, which might, in turn, shed light on the evolution of language.

Some researchers have suggested that sound symbolic words played an important role in the evolution of

human language (Kita, 2008; Ramachandran & Hubbard, 2001). The very idea of a word as an association between a speech sound and a referent might have evolved on the basis of the ability to directly link representations in auditory and other sensory modalities and motor processes among each other (Ramachandran & Hubbard, 2001). For example, small objects are associated with a high closed vowel /i/ and large objects are associated with a low open vowel /a/ (e.g., Sapir, 1929). This may be because we consider the size of objects to be analogous to the size of the oral cavity or the "size" of acoustic image for different vowels (Sapir, 1929) and because we draw analogy between the size of the oral cavity and the aperture of a grasping hand (Ramachandran & Hubbard, 2001). The ability to recognize such analogies between representations in different modalities might lead to sound symbolism and the emergence of the first word in evolution. Once our ancestors evolved the innate disposition to develop universal sound symbolic intuitions and their use in word learning, they could rapidly agree on what words refer to what. This would have facilitated the growth of a shared lexicon (Ramachandran & Hubbard, 2001). Given that sound symbolic words in modern languages can refer to information in various domains such as vision, touch, smell, taste, manners of movement, emotion and attitude (e.g., Kita, 1997; Voeltz & Killian-Hatz, 2001), sound symbolic proto words of our ancestors may have had a considerable expressive power (Kita, 2008). This in turn would have given a considerable adaptive advantage to our ancestors. Universal sound symbolism in modern languages may be "fossils" reflecting this evolutionary process.

We suggest that all humans have the innate disposition to develop sound symbolic abilities and use them for word learning and that the emergence of this disposition was a crucial step in language evolution. It is possible that the present study tapped into the vestige of this evolutionary process still present in all children.

References

Atoda, T. & Hoshino, K. (1995). *Giongo gitaigo tsukaikata jiten [Usage dictionary of sound/manner mimetics]*. Sotakusha.

Allen, S., Ozyurek, A., Kita, S.S Brown, A., Furman, R., Ishizuka, T., & Fujii, M. (2007). How language specific is early syntactic packaging of Manner and Path? A comparison or English, Turkish, and Japanese. *Cognition*, 102, 16-48.

Childs, G. T. (1996). Where have all the ideophones gone?: The death of a word category in Zulu. Toronto Working paper in Linguistics, 15(2), 81-104.

Davis, R. (1961). The fitness of names to drawings. *British Journal of Psychology*, 52, 259-268.

Dediu, D., & Ladd, D. R. (2007). Linguistic tone is related to the population frequency of the adaptive haplogroups of two brain size genes, ASPM and Microcephalin. *Proceedings of National Academy of Sciences, 104*, 10944-10949

Hamano, S. (1998). *The sound symbolic system of Japanese.* Stanford, CA: CSLI Publications & Tokyo: Kuroshio.

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

Imai, m., Haryu, E., & Okada, H. (2005). Mapping novel nouns and verbs on to dynamic action events: Are verb meanings easier to learn than noun meanings for Japanese children? *Child Development*, 76, 340-355.

Imai, M. Kita, S., Nagumo, M., & Okada, H. (2008). Sounds Symbolism between a word and an action facilitates early verb learning. *Cognition*, 109, 54-65

Kita, S. (2001). Semantic schism and interpretive integration in Japanese sentences with a mimetic: A reply to Tsujimura. *Linguistics*, 39, 419-439.

Kita, S. (2008). *World-view of protolanguage speakers as inferred from semantics of sound symbolic words: A case of Japanese mimetics*. In N. Masataka (Ed.), Origins of language. Tokyo: Springer.

Kita, S. (1997). Two-dimensional semantic analysis of Japanese mimetics. *Linguistics*, 35, 379-415.

Kohler, W. (1947). *Gestalt Psychology*. (2nd ed.) NY: Liverlight.

Maguire, M., Hennon, E., Hirsh-Pasek, K., Golinkoff, R., Slutzky, C. & Sootsman, J. (2002). Mapping words to actions and events: How do 18-month-olds learn a verb? In B. Skarabela, S. Fish, & A. H. Do (Eds.), *Proceedings of the 27th Annual Boston University Conference on Language.* Somerville, MA: Cascadilla Press.

Markman, E. M. & Watchtel, G.F. (1988). Childrens use of mutual exclusivity to constrain the meaning of words. *Cognitive Psychology*, 20, 121-157.

Matsimoto, Yo (1997). Kuukan idoo no gengohyoogen to sono kakuchoo [Linguistic expressions of motion in space, and their extensions]. In Shigenoti Tanaka & Yo Matsumoto, *Kuukan to idoo no hyoogen [Expressions of space and motion]*. Tokyo: Kenkyuusha.

Maurer, D., Pathman, T., 7 Mondloch, C. J. (2006). The shape of boubas: Sound-shape correspondences in toddlers and adults. *Developmental Science*, 9, 316-322.

Naigles, L. G. & Kako, E. T. (1993). First contact in verb acquisition: Defining a role for syntax. *Society for Research in Child Development*, 64, 1665-1687.

Newmeyer, F. J. (1992). Iconicity and generative grammar. Language, 68, 756-796.

Nuckolls, J. (2004). To be or not to be ideophonically impoverished. *Texas Linguistics Forum, 47*, 131-142.

Quine, W.V.O.(1960). *Word and object*. Cambridge, MA: MIT Press.

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

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

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