Article

Sound Symbolism Facilitates Long-Term Retention of the Semantic Representation of Novel Verbs in Three-Year-Olds

Katerina Kantartzis [1],* **, Mutsumi Imai [2]** **, Danielle Evans [3] and Sotaro Kita [4]**

[1] School of Natural and Social Sciences, University of Gloucestershire, Gloucestershire GL50 4AZ, UK
[2] Faculty of Information and Environment Sciences, Keio University, Tokyo 252-0882, Japan; imai@sfc.keio.ac.jp
[3] School of Psychology, University of Birmingham, Birmingham B15 2TT, UK; danni1122e@gmail.com
[4] Department of Psychology, University of Warwick, Coventry CV4 7AL, UK; S.Kita@warwick.ac.uk
* Correspondence: kkantartzis@glos.ac.uk

check for updates

Abstract: Previous research has shown that sound symbolism facilitates action label learning when the test trial used to assess learning immediately followed the training trial in which the (novel) verb was taught. The current study investigated whether sound symbolism benefits verb learning in the long term. Forty-nine children were taught either sound-symbolically matching or mismatching pairs made up of a novel verb and an action video. The following day, the children were asked whether a verb can be used for a scene shown in a video. They were tested with four videos for each word they had been taught. The four videos differed as to whether they contained the same or different actions and actors as in the training video: (1) same-action, same-actor; (2) same-action, different-actor; (3) different-action, same-actor; and (4) different-action, different-actor. The results showed that sound symbolism significantly improved the childrens' ability to encode the semantic representation of the novel verb and correctly generalise it to a new event the following day. A control experiment ruled out the possibility that children were generalising to the "same-action, different-actor" video because they did not recognize the actor change due to the memory decay. Nineteen children were presented with the stimulus videos that had also been shown to children in the sound symbolic match condition in Experiment 1, but this time the videos were not labeled. In the test session the following day, the experimenter tested the children's recognition memory for the videos. The results indicated that the children could detect the actor change from the original training video a day later. The results of the main experiment and the control experiment support the idea that a motivated (iconic) link between form and meaning facilitates the symbolic development in children. The current study, along with recent related studies, provided further evidence for an iconic advantage in symbol development in the domain of verb learning. A motivated form-meaning relationship can help children learn new words and store them long term in the mental lexicon.

Keywords: sound symbolism; verbs; memory; iconicity

1. Introduction

Word learning presents certain challenges to children. In order to learn new words children must identify the referent of the word in a complex reality. Identifying the correct referent of a word is no small task, as discussed by Quine (1960). For example, when a parent points to a rabbit hopping and says, "hopping", the child must identify the action as the referent of the word, but the word "hopping" could refer to, for example, the name of the animal, the color or the fur. Previous studies demonstrated that sound symbolism helps children overcome the problem of determining the referent of novel

words and hence facilitate verb learning (Imai et al. 2008a, 2008b; Kantartzis et al. 2011). The current study further investigated whether the facilitative effect is long lasting, i.e., whether children are not only better in identifying the meaning of a novel verb but also in retaining it when the word contains sound symbolism.

Sound symbolism is the inherent (or motivated) link between the sound and meaning of a word, and it is recognized and used cross-linguistically. One of the earliest and most celebrated examples is Köhler's (Köhler 1947) sound symbolism for shapes. When asked to match rounded and angular objects to two novel words, *maluma* and *takete*, a large majority of adults from different linguistic backgrounds pick *takete* for the angular object and *maluma* for the rounded object (e.g., Davis 1961; Holland and Wertheimer 1964). Shape sound symbolism can also be recognized by children (Asano et al. 2015; Imai et al. 2015; Maurer et al. 2006; see Fort et al. 2013 and Fort et al. 2018 for an extensive review on shape sound symbolism in infants and toddlers). Sound symbolism has also been shown in size, for example, with words containing /i/ being assigned to smaller objects and those containing /a/ being assigned to larger objects (Sapir 1929). Another classic method used to demonstrate a cross-linguistic sensitivity to sound symbolism is antonym pair matching. Here, participants could match foreign antonym pairs to English antonym pairs above chance (Brown et al. 1955; Brown and Nuttall 1959; Gebels 1969; Klank et al. 1971; Kunihara 1971; Siegel et al. 1965; Tsuru and Fries 1933). Other studies showed that sound symbolic words in one language (Japanese) can be recognized by speakers of another language (English) in the domain of action (Imai et al. 2008a; Kantartzis et al. 2011; Kita et al. forthcoming) and pain (Iwasaki et al. 2007a).

The idea that a motivated link between form and meaning facilitates childrens' symbolic development (Werner and Kaplan 1963; Deloache 1995) has received much empirical scrutiny in recent years (e.g., Imai and Kita 2014). We focus our review on supporting evidence from studies on 2-4-year-olds, which are close to the age group of the current study (three-year-olds), though recent meta-analyses of the maluma-takete effects in infants suggest that this effect is robust even in younger children (Fort et al. 2013). When gestures were taught as a label for an object and children had to select the correct object among a set of objects, children performed better with iconic gestures (with a motivated form-meaning relationship) than with arbitrary gestures (26 month olds, Namy et al. 2004; 26 month olds, Namy 2008; three- and four-year-olds, Marentette and Nicoladis 2011). Furthermore, in an event recognition task, three year olds remembered the actor and/or the action better when an adult experimenter highlighted the actor and/or action with an iconic gesture (depicting the event) when presenting the video recording of an event to the child (Aussems and Kita, in press). When children were shown where an object is hidden in a scaled-down model of a room and had to find the object in a real room, children performed better as the scale model becomes more and more similar to the real room (2.5-year-olds, Deloache and Sharon 2005). In this context, it is interesting to see whether a similar advantage for a motivated form-meaning relationship (i.e., sound-symbolic words) can be observed in the language domain.

The current study focuses on the learning of action labels, which children find more difficult than object labels (e.g., Gentner 1982; Childers and Tomasello 2002; Imai et al. 2005; Imai et al. 2008b). The primary difficulty lies in extracting action as the essential part of a scene for the verb meaning (Golinkoff and Hirsh-Pasek 2008; Maguire et al. 2002). In particular, three-year-olds found it difficult to generalize newly learnt novel verbs to a novel situation on the basis of action identity (Abbot-Smith et al. 2017; Imai et al. 2005; Kersten and Smith 2002). Specifically, having learnt a novel verb while watching action A performed by actor X, children typically did not extend the meaning of the novel verb to action A performed by actor Y. However, both Japanese-speaking and English-speaking children could generalize novel verbs on the basis of the same action if the novel verbs sound-symbolically matched the action (Imai et al. 2008a; Kantartzis et al. 2011; Yoshida 2012; Yoshida and Smith 2003). It was argued that this is because sound symbolism highlighted the action, and this made it easier for children to extract the action apart from the object as the referent of the verb.

Previous research (Imai et al. 2008a; Kantartzis et al. 2011; Yoshida 2012; Yoshida and Smith 2003) has shown that sound symbolism facilitates action label learning when the test trial immediately followed the training trial. However, it is not clear whether sound symbolism brings a long-term benefit. It is possible that in the training trial sound symbolism activates action representation only temporarily, which helps children to correctly focus on the action in the immediately following test trial, but this activation decays quickly. In that case, sound symbolism has limited utility as it does not help create a long-lasting semantic representation in the mental lexicon. Thus, the current study used a delayed test to investigate the stability of lexical representation after fast mapping (e.g., Childers and Tomasello 2002; McGregor et al. 2009). More specifically, it investigated whether sound symbolism can help three-year-old children retain the semantic representation of a novel verb after a one-day delay.

To investigate the role of sound symbolism in the long-term memory of verbs, children were assigned to two groups. Those in the sound-symbolically matching group were taught four novel verbs with sound-symbolically matching referent actions, while those in the sound-symbolically mismatching group were taught four novel words with sound-symbolically mismatching referent actions. On the second day, the children were asked whether a verb can be used for a scene shown in a video. They were tested with four videos for each word they had been taught. The four videos differed as to whether they contained the same or different action and actor as in the training video: (1) same-action, same-actor; (2) same-action, different-actor; (3) different-action, same-actor; and (4) different-action, different-actor. The sound-symbolically mismatching condition was set up in such a way that the different-action, same actor video (i.e., the incorrect choice in the verb generalization task) was sound-symbolically matching to the taught action label. If we obtain the predicted result in the sound-symbolically matching condition, one might argue that this outcome is possible without learning the mapping and generalizing the verb; instead, children could have simply detected a sound-symbolic match between a word and an action just at test. If this alternative interpretation is correct, then children in the sound-symbolically mismatching condition should erroneously say "yes" for the different-action, same-actor test-type above chance level. For example, in the training phase, children in the sound-symbolically mismatching condition may have been taught the word *batobato* for a sound-symbolically mismatching action in which an actor walks slowly, with arms loosely bent and hands touching in the front (see Table 1, Figure 1b). Following this, in the testing phase, the children were shown a sound-symbolically matching action in the different-action, same-actor trial; that is, they were asked whether the word *batobato* can refer to a different but sound-symbolically matching action (see Figure 1b). If children correctly understood the referent of the verb *batobato*, they should have said "no" in this trial because the action was different from the training phase; however, if children were simply saying "yes" to a sound-symbolically matching action in the test phase, then they would have said "yes".

Table 1. Description of the matching and mismatching actions for each sound-symbolic word.

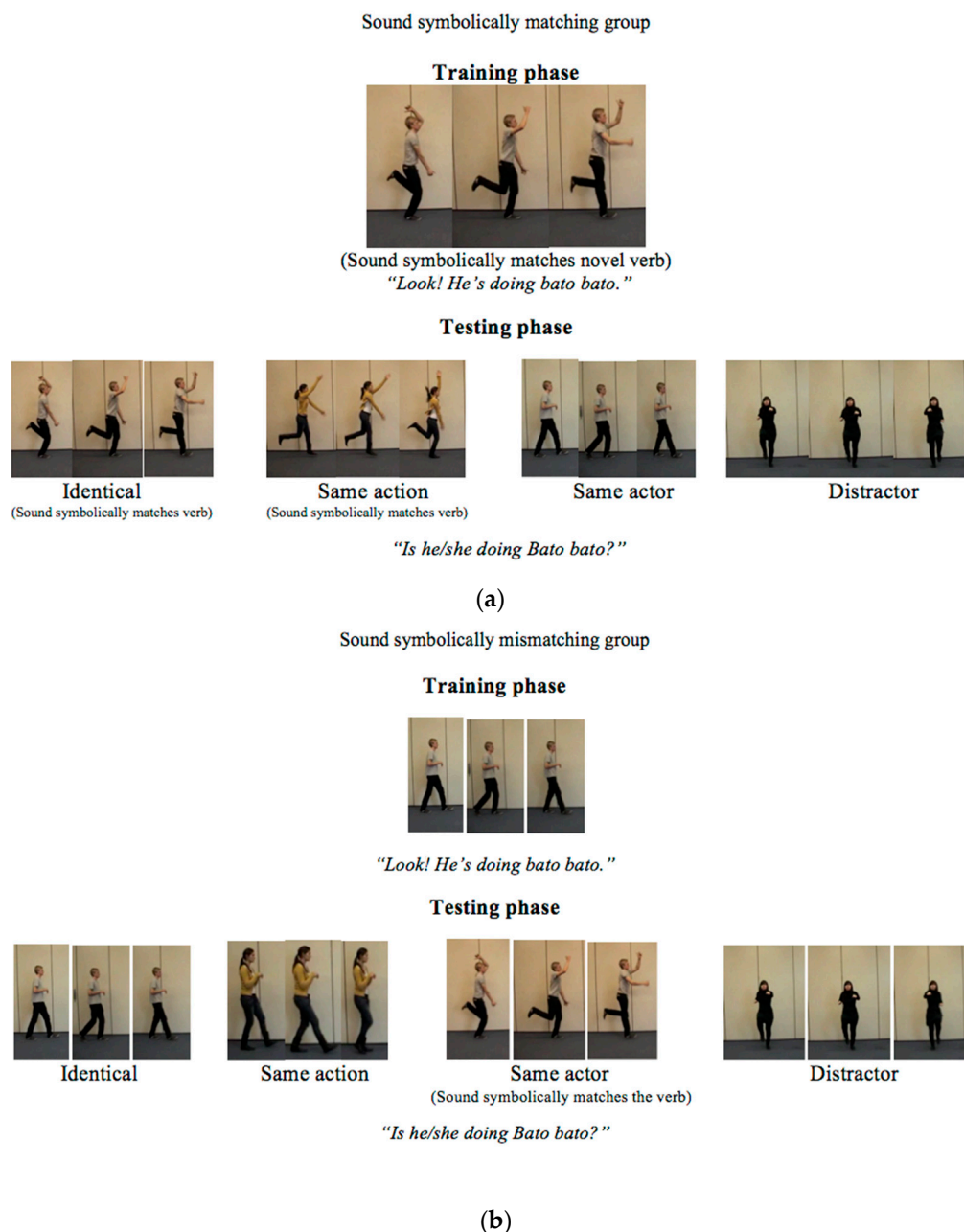| Word	Sound-Symbolic Matching Action	Sound-Symbolic Mismatching Action
Bato bato	A large energetic movement, arms are swinging back and forth outstretched, whereas legs are making large leaping movement	Walking slowly, with arms loosely bent and hands touching in the front
Choka choka	Walking quickly in very small steps with the arms swinging quickly with bent elbows	Legs slightly bent, walking slowly and in a controlled fashion, with arms bent and held out in front of body (as if carrying a tray);
Nosu nosu	Walking slowly in large steps with bent knees and hands on knees	Legs making large steps forward, with a bounce, arms swinging freely from side to side
Toku toku	A small shuffling movement, with straight arms rigidly at the side and legs moving very slightly and rigidly.	Creeping-type walk with medium sized steps, with arms bent and held closely in front of body

Sound symbolically matching group

Training phase

(Sound symbolically matches novel verb)
"Look! He's doing bato bato."

Testing phase

Identical	Same action	Same actor	Distractor
(Sound symbolically matches verb)	(Sound symbolically matches verb)		

"Is he/she doing Bato bato?"

(a)

Sound symbolically mismatching group

Training phase

"Look! He's doing bato bato."

Testing phase

Identical	Same action	Same actor	Distractor
		(Sound symbolically matches the verb)	

"Is he/she doing Bato bato?"

(b)

Figure 1. Examples of the videos children were shown for the novel word "bato bato" in the training and testing sessions (**a**) in the sound-symbolically matching condition, and (**b**) in the sound-symbolically mismatching condition. The correct responses would be "yes" for the Identical and Same action items, and "no" for the Same Actor and Distractor items.

If sound symbolism helps children learn the referent of the verb in such a way that they can generalize what they learned to a new situation after one day, children in the sound-symbolically matching condition should be better at identifying the referent action (accepting (1) and (2), and rejecting (3) and (4)). This pattern of results is compatible with two possibilities, only one of which involves a true generalization. The first possibility is that children can remember who the actor is in the original video shown in the training session and generalize the verb to the same-action-different-actor event even though they are aware of the actor change. This would indicate that children in the sound symbolic match condition form the correct semantic representation of the verb that contains only the action, which enables generalization. The second possibility is that children have a poor memory for actors, and thus they fail to distinguish the versions with the same actor or a different actor shown

in the test session. In this case, children in the sound symbolic match condition associated the action to the verb, due to sound symbolism, but they might also associate the actor to the verb as well. However, a poor memory for actors prevents them from using the actor component of the semantic representation in the test session. To disambiguate the two possibilities, in Experiment 2, we tested the children's recognition memory for the videos shown in the training session of Experiment 1. If the first possibility is correct, children should notice when the actor is different in the test video.

2. Experiment 1

2.1. Method

2.1.1. Participants

Seventy-three three-year-old monolingual English speaking children were recruited from various nurseries around Birmingham, UK. 24 children were excluded, 12 due to a "yes" bias (i.e., a "yes" answer in all trials), one due to a no bias (i.e., a "no" answer in all trials), 11 because they were not available to complete the testing on the second day. Thus, there were 49 children (Mean: 3 years; 7 months, Range: 3;0–3;11) who were included in the final analysis, 25 in the sound-symbolically matching group and 24 in the sound-symbolically mismatching group. All participants had full consent from parents or caregivers before they participated in the study, and only children who were willing to take part were included. The study was conducted according to British Psychological Society Ethical Guidelines as approved by the University of Birmingham Ethics Committee (Identification code: ERN_10-1169).

2.1.2. Stimuli

We used the same word-action pairs as in Kantartzis et al. (2011), except that the videos were re-recorded. There were four novel words (*bato bato, choka choka, nosu nosu* and *toku toku*) borrowed and modified from conventional Japanese mimetics (see Imai et al. 2008a for additional details on the procedure of creating novel words). Each novel word had a sound-symbolically matching action and a sound-symbolically mismatching action (see Table 1).

The stimuli in Kantartzis et al. (2011) had been pretested to verify the sound-symbolically matching and mismatching relationships between the novel words and actions, and so this was not repeated for the current study. For the purpose of the current study, four new distractor actions were also created—one action for each word—in which the actor and the action were different from any of the other videos (see Figure 1 for an example for the distractor action). One female and one male actor performed all the sound-symbolically matching and mismatching actions, while a second female actor performed the distractor actions. The identical test-type was the same video as in the training session. The same-action test-type contained the same-action as in the training session but performed by a different actor. The same-actor test-type contained the same-actor as in the training session performing a different action. Finally, the distractor test-type contained a different actor performing a different action (see Figure 1 for an example of the training video and the four test videos).

2.1.3. Condition

The children were randomly assigned to one of the two groups. In the sound-symbolically matching group (*n* = 25; mean age = 3 years 7months, range: 3;0–3;11) the children were taught novel verbs that sound-symbolically matched the action they represented in the training session (see Figure 1a.). Consequently, in the test session, the verbs sound-symbolically matched the action in the identical test video and the same-action test video (the "yes-items" in the verb generalization task), but not in the same-actor and the distractor test videos (see Figure 1a).

In the sound-symbolically mismatching group, (*n* = 24; mean: 3 years, 6 months, range: 3;0–3;11), the children were taught novel verbs that did not sound-symbolically match the action in the training

session. Consequently, in the test session, the verbs did not sound-symbolically match the actions in the identical and same-action test videos (the "yes-items"). However, the new action in the same-actor test video (a "no-item" in the verb generalization task) sound-symbolically matched the verb (see Figure 1b). This same-actor, sound-symbolically matching test trial was therefore challenging for the children because they had to say "no" (i.e., the verb does not apply to the scene) despite the fact that the actor was the same as in the training video and the action in the test video sound-symbolically matched the verb.

2.1.4. Procedure

The children participated in the task in two sessions over two days: a training session on the first day and a test session on the second day. The test session always took place during the same time period (for example in the morning) as in the training session. During the training session, the children were shown four warm-up slides with a familiar object (cat, apple and biscuits) and a video of a familiar action (hopping). For each slide, the experimenter labeled the object or the action. For each novel word, the experimenter showed one of the training videos and said, "Look! He is doing X" (where X represents the novel word). This was repeated for the four novel words. The test session started with four practice trials. In each practice trial, either a picture of an object (apple, cat) or a video of an action (hopping, walking) was shown, and the experimenter asked a yes-no question (e.g., "Is this an apple?", "Is he clapping?") in such a way that the children were expected to give two yes-answers and two no-answers. In the main trials, the experimenter showed one of the test videos and said, "Ellie thinks he/she is doing XXX. Is he/she doing XXX?" (Ellie is the name of the puppet which was also present at the training session to make the experiment more child-friendly). For each novel word, four trials of the different test-types were carried out: identical, same-action, same-actor and distractor. The order of the four test-types was counterbalanced across the four words for each participant and across participants. The childrens' response was either "yes" or "no", and the researcher recorded this.

2.2. Results and Discussions

If children are successfully learning and generalizing the verbs, they should respond "yes" to the question, "Is he/she doing X?" for the identical and same-action test-types, and "no" to the same-actor (different action) test-type and the distractor test-type both in the sound-symbolically matching and mismatching conditions. In order to test whether sound symbolism improved their performance, the proportions of the trials with correct responses were entered into a 2 (between participant, group: sound-symbolically matching, sound-symbolically mismatching) × 4 (within participant, test-type: identical, same-action, same-actor and distractor) Repeated Measures ANOVA (see Figure 2 for the means). When the sphericity assumption is violated, the Greenhouse-Geisser correction of degrees of freedom is used throughout the paper, and we report the corrected degrees of freedom. There was a significant main effect of the test-type, $F(2.25, 105.71) = 15.99$, $p < 0.001$, $\eta_p^2 = 0.25$. More specifically, the children responded poorly in the same-actor condition compared to the other three conditions (Tukey HSD $p < 0.05$). The main effect of sound symbolism indicates that the children in the sound-symbolically matching group performed better than the children in the sound-symbolically mismatching group $F(47,1) = 5.10$, $p < 0.05$, $\eta_p^2 = 0.10$. The interaction between test-type and sound symbolism was not significant.

In order to investigate whether sound symbolism facilitated the performance in each test-type, a planned comparison of sound-symbolic match versus mismatch was carried out for each test-type (See Figure 2 for the means). The children in the sound-symbolically matching condition performed better than those in the sound-symbolically mismatching condition in the identical ($t(47) = 2.48$, $p < 0.05$) and same-action test-types ($t(47) = 2.43$, $p < 0.05$), but not in the distractor and same-actor test-types.
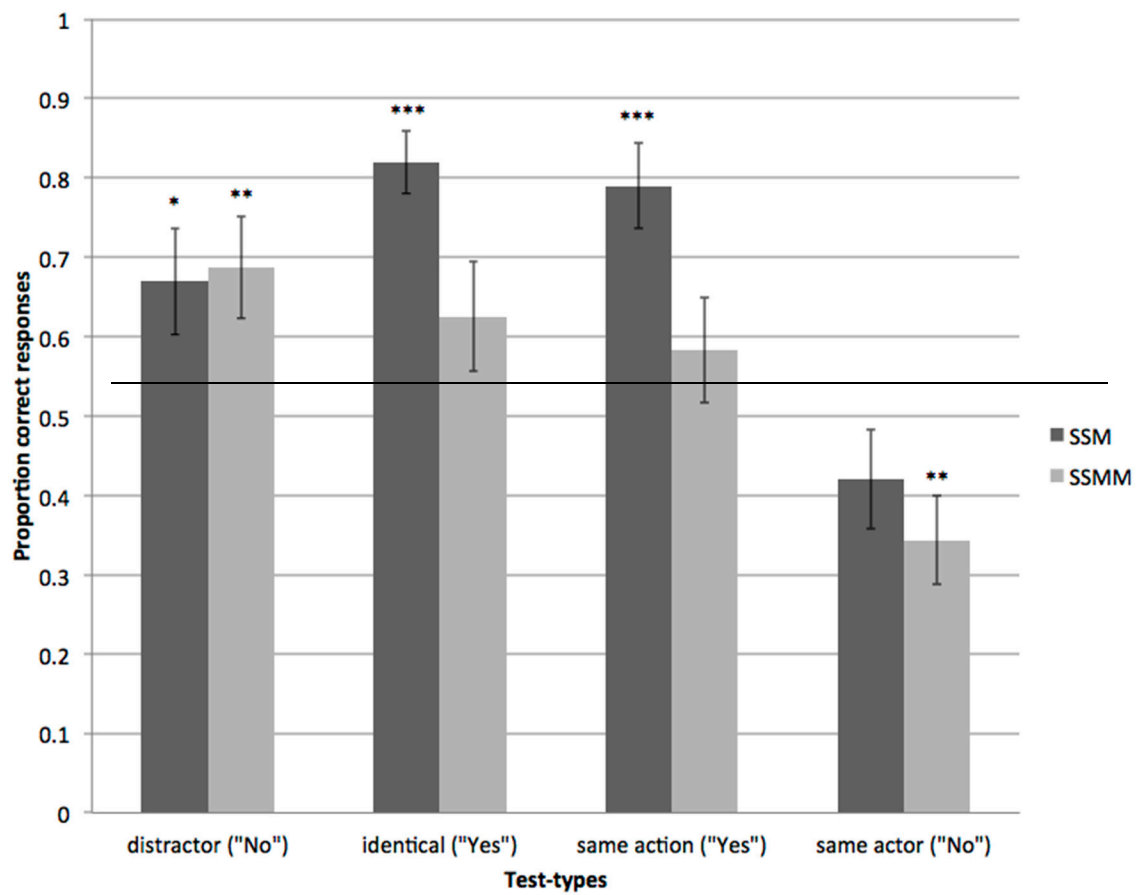
Figure 2. The proportion of correct responses in the sound-symbolically matching (SSM) and sound-symbolically mismatching (SSMM) conditions, in the four test-types: distractor, identical, same-action and same-actor. The correct response ("Yes" or "No") is indicated in brackets. The error bars represent the standard error of the means. Whether or not the mean was significantly different to chance (0.5) is indicated by *** $p < 0.001$, ** $p < 0.01$ and * $p < 0.05$.

The proportion of correct responses for each test-type (distractor, identical, same-action, and same-actor) for both the sound-symbolically matching and sound-symbolically mismatching groups were compared against chance (0.5) (See Figure 2). The children in the sound-symbolically matching condition gave correct responses at an above chance level ($p < 0.05$) for the identical, same-action, and distractor tests. In contrast, the children in the mismatching group responded correctly above chance for the distractor test ($p < 0.05$), but not for the identical or the same action test. Their responses were significantly below chance ($p < 0.05$) for the same-actor test, which means that children falsely gave an incorrect 'yes', possibly because the word sound-symbolically matches the video in the test session.

This raises the possibility that the children may have just been using sound symbolism in the test session to guide their response and not in the training session to help them learn the verb-action combination. As we discussed, however, this possibility could be addressed by comparing the same-action tests (correct answer = "yes") in the sound-symbolically matching group and the same-actor tests (correct answer = "no") in the sound-symbolically mismatching group. In both cases the children were presented with a sound-symbolically matching word-action combination at test. If the children were simply identifying sound symbolism during the test, they should be equally likely to respond with "yes" in these two test-types ("yes" would be a correct answer in the same action condition, but a wrong answer in the same actor condition). The proportion of "yes" responses was higher in the same-action test type in the sound-symbolically matching group than in the same-actor test-type in the sound-symbolically mismatching group, $t(47) = 2.36$, $p < 0.05$, ruling

out the possibility that children were simply responding based on sound symbolism at test without learning and generalizing the taught words (see Figure 2 for the proportion of correct responses. Note that for the same-actor action test-type, "yes" is an incorrect response). Thus, sound symbolism helped the children to map the word to the referent and to encode the mapping in their memory during the training session.

The results indicated that sound symbolism helped the children identify and store the correct semantic representation of the verb, containing only an action, and then generalize the verb to a new event with the same action by a different actor on the following day. However, an alternative interpretation is possible. That is, in the test session, children may have had a poor memory for actors. Sound symbolism helped them to remember what the action was, but they may have mistakenly thought that the same-action video was the same as the video from the training session. If this was the case, the success in the same-action condition cannot be taken as evidence for the children's ability to generalize to a new situation with a different actor. Thus, Experiment 2 was set up to rule out this alternative explanation.

Experiment 2 investigated the children's recognition memory for actions in the events presented in Experiment 1. In the training session, the children were shown the same training videos as in Experiment 1, but words were not presented with them. The test session took place one day later: the children were shown the same four test videos as in Experiment 1 and asked whether they were identical to the videos they saw one day earlier. Their response to the actor-change items in Experiment 2 would tell us whether they had a poor memory for actors.

3. Experiment 2

3.1. Method

3.1.1. Participants

The participants were 34 3-year-old monolingual English-speaking children, from various nurseries around Birmingham, UK. Fifteen children were removed: eight due to a yes-bias and seven because they were not available for the testing session. Therefore, a total of 19 children (Mean: 3 years, 6 months; Range 3;0–3;9) were used in the analysis.

3.1.2. Stimuli

The videos used in this task were identical to those used in the sound-symbolically matching condition of Experiment 1. Thus, there were 4 training videos (i.e., the target videos for recognition), and 16 test videos. The pictures and a video for practice trials were also the same as for Experiment 1. Unlike in Experiment 1, no novel words were involved.

3.1.3. Procedure

The procedure was essentially the same as in Experiment 1 except that no novel words were involved and the children were tested on their recognition memory for the videos. During the training session, the children were shown four warm-up slides with a familiar object and a video of a familiar action, followed by four slides, each with one of the four experimental target videos. For each slide, the experimenter encouraged the children to look at the picture or action by saying, "Look! Look! Look!" During the test session children were again shown four practice slides with a familiar object or action, and 16 experimental test videos of the four test-types: identical, same-action, same-actor and distractor (in the same counterbalanced orders as in Experiment 1). For each video, the experimenter said, "Ellie thinks he/she saw this video yesterday. Did you see this video yesterday?" (Ellie is the name of the puppet also present at the training session to make the experiment more child-friendly.).

3.2. Results and Discussions

The aim of the experiment was to check whether the children could detect the changes from the original event a day after their exposure to the original event. Importantly, this experiment investigated the memory of which parts of a scene were fragile and likely to decay fast. The proportion of "yes" responses was entered into a repeated measures ANOVA, with the test-type (identical, distractor, same-action, and same-actor) as the within subject factor. Note that "yes" was a correct response for the identical test type, but incorrect for other test types. There was a significant effect of test-type $F(3,54) = 21.29$, $p < 0.001$, $\eta_p^2 = 0.54$ (see Figure 3). Posthoc *t*-tests with Bonferroni corrections showed that there were significantly more "yes" responses in the identical than in the distractor ($p < 0.001$) and the same action ($p < 0.001$) test-types. There were significantly more "yes" responses in the same-actor test-types than in the distractor test-type ($p < 0.001$). Finally, the proportion of "yes" responses was significantly different from chance (0.5) for the distractor, identical and same-actor test-types ($ps < 0.05$).
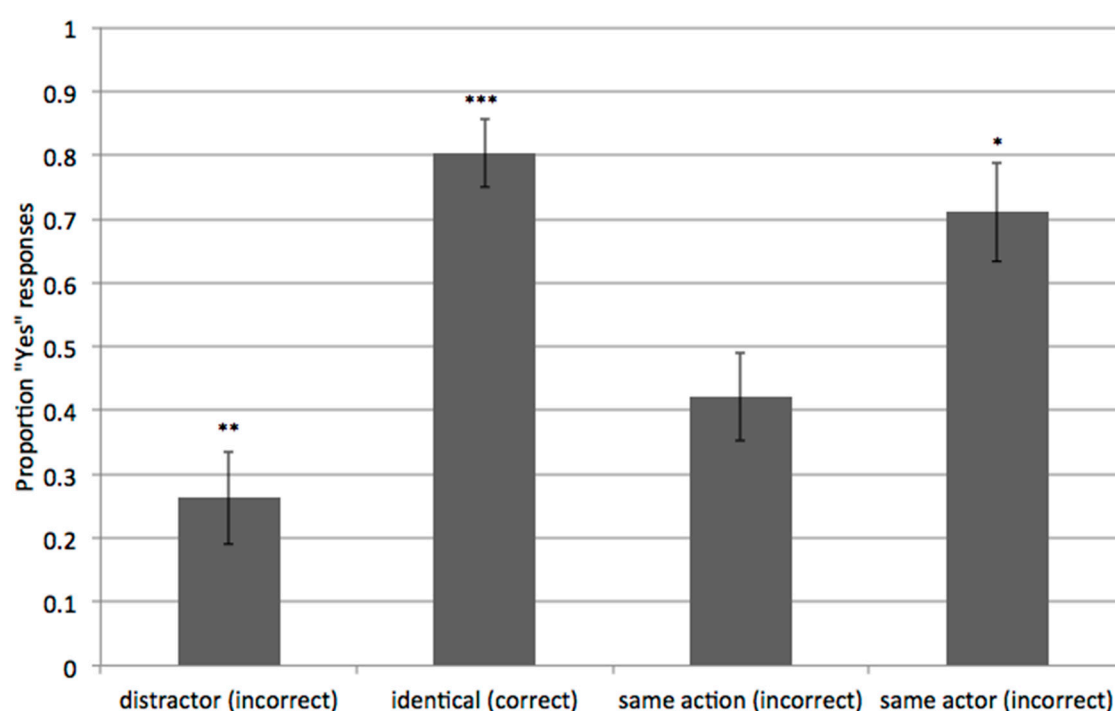

Figure 3. The proportion of "Yes" responses to the identical, same-action, same-actor and distractor test videos. Whether or not a "Yes" response would be correct is indicated in brackets. The error bars indicate the standard error of the means. Whether the mean is significantly different to chance (0.05) is shown by *** indicating $p < 0.001$, ** indicating $p < 0.01$ and * indicating $p < 0.05$.

It is interesting that the children tend to erroneously confuse the event with a different action when the actor was the same (the same actor condition), which suggests that children in general have a more robust memory to objects than to actions. However, the key result is that children responded "yes" more often for the identical test videos than for the same-action test videos (with different actors from the original training videos). This indicates that the children could detect the actor change from the original training video after a day, which is consistent with previous findings (Imai et al. 2005). We thus conclude that in Experiment 1 the children in the matching condition generalized the verb a day after the training to a scene containing the same action, despite the fact that they were aware of the actor change. This in turn indicates that sound symbolism helped the children build a correct semantic representation of the verbs, which contained only actions.

Furthermore, the results from the same-actor condition in Experiment 2 indicate that the children were worse at recognizing the action change than the actor change (see Imai et al. 2005, see also Aussems and Kita 2017, who did not find this difference, but found an at-chance performance for both

action-change and actor-change items): they often mistook the same-actor test videos as being identical to the videos shown in the training. Such a confusion may explain the frequent incorrect acceptance of the same-actor test videos in Experiment 1.

4. General Discussion

This study investigated the role of sound symbolism in a word learning task in which the lexical representation had to be retained until the following day. The children in the sound-symbolically matching condition correctly identified the referent action in the video identical to the training video and in the video showing the same action by a different actor more successfully than those in the sound-symbolically mismatching condition (Experiment 1). Thus, sound symbolism brings long term memory benefits to word learning. An additional analysis and experiment ruled out potential alternative interpretations such as a sound-symbolic benefit at the test session only (Experiment 1) and children's poor memory for actors in the training videos (Experiment 2). Thus, we conclude that sound symbolism helps children create correct semantic representation of verbs, containing only action, and that this representation is robust enough to be retained until the following day.

The current study went beyond the previous studies (Imai et al. 2008a; Kantartzis et al. 2011; Yoshida 2012; Yoshida and Smith 2003) in showing that sound symbolism does not merely boost the activation of the referent temporarily for a short period of time, but sound symbolism helps create a correct semantic representation of novel words that can be maintained over time. Because words need to be stored in the mental lexicon for the long term, it is important to have established that the semantic representation created with the help of sound symbolism is long-lasting.

In the current study, sound symbolism highlighted the action in a scene, which presumably helped children focus on the action as the referent of a novel verb. This dovetails with the findings in the literature that other ways of highlighting action also lead to a better performance in verb generalization tasks. Children performed better when action events were presented as a point-light display, which de-emphasized the actor, thereby emphasizing the action (Maguire et al. 2008), when action events were accompanied by an adult experimenter's iconic gesture that depicted an action (Mumford and Kita 2014), and when multiple scenes that shared the same action were presented to children (e.g., Childers 2011; Snape and Krott 2018). Sound symbolism may therefore be one of many cues children can use to zero-in on the referent in a complex scene, that is, to solve Quine's (1960) problem.

The current study also highlighted a link between event memory and word learning. The children have a poor memory for the action in an event (Experiment 2), which was also shown previously (Imai et al. 2005). This may partly explain why children in the word learning experiment performed poorly in the same-actor test type (Experiment 1), in which the actor was the same as for the training video but where the action was different. The children may have sometimes mistaken this video to be identical to the training video. Thus, a poor memory for action may add to the challenges that children face when learning verbs (Imai et al. 2005).

The current study also contributed to the discussion concerning the role of iconicity, or a motivated form-meaning relationship, in symbolic development. Some theories (Werner and Kaplan 1963; Deloache 1995) proposed that symbols with a motivated form-meaning relationship facilitate children's symbol understanding. This was supported for two-to-four-year-olds in various domains. Children understand a scale model of a room better if the model is more similar to the room (Deloache and Sharon 2005). Children remember gestural labels for objects and associated actions better when the labels are iconic as opposed to arbitrary (Namy et al. 2004; Namy 2008; Marentette and Nicoldis 2011) or as opposed to miniature replicas of the objects (Tomasello et al. 1999). Children performed better in an event recognition task (Aussems and Kita 2017), a verb generalization task (Mumford and Kita 2014) and other word learning tasks (Goodrich and Kam 2009; McGregor et al. 2009) when they saw an adult experiment's iconic gesture (McNeill 1992) depicting the action or spatial relationships in the stimuli. The current study,

along with recent related studies (Imai et al. 2008a; Kantartzis et al. 2011), provided further evidence for the iconic advantage in symbol development in the domain of word learning. A motivated form-meaning relationship can help children learn new words and store them long term in the mental lexicon.

Finally, the current study contributes to the discussion on the role of embodiment in language in two ways. First, sound symbolism grounds words onto our perceptual motor experiences. Ohala (1994) suggested that we associate high vowels such as /i/ with smallness because of our visual and auditory perception. That is, we associate sounds with an energy concentration in higher frequencies (including high vowels) with smallness because smaller animals with shorter vocal tracts tend to produce a vocalization with a higher pitch. In contrast, Sapir (1929) suggested that we associate high vowels with smallness because of our motor experience during articulation; when producing high vowels, the air passes through a narrow (small) oral cavity due to the raised tongue (see also Ramachandran and Hubbard 2001, for a discussion of the motor-based motivation for sound symbolism). Second, sound symbolic words are often produced with co-expressive gestures. Kita (1997, 2001) and Kita et al. (forthcoming) found that when Japanese speakers produce a sound symbolic word ("mimetic") in a narrative, they were highly likely to also produce a gesture that synchronized with the sound symbolic word and depicted the event the sound symbolic words referred to (see also Dingemanse 2013). The current study showed that sound symbolic words with embodied form-meaning relationships are easier to learn for children. This highlights the importance of embodiment in language development in children.

References

Abbot-Smith, Kirsten, Mutsumi Imai, Samantha Durrant, and Erika Nurmsoo. 2017. The role of timing and prototypical causality on how preschoolers fast-map novel verb meanings. *First Language* 37: 186–204. [CrossRef]

Asano, Michiko, Mutsumi Imai, Sotaro Kita, Keiichi Kitajo, Hiroyuki Okada, and Guillaume Thierry. 2015. Sound symbolism scaffolds language development in preverbal infants. *Cortex* 63: 196–205. [CrossRef] [PubMed]

Aussems, Suzanne, and Sotaro Kita. 2017. Seeing iconic gestures while encoding events facilitates children's memory of these events. *Child Development*. [CrossRef] [PubMed]

Brown, Roger, and Ronald Nuttall. 1959. Methods in phonetic symbolism experiments. *Journal of Abnormal and Social Psychology* 59: 441–45. [CrossRef] [PubMed]

Brown, Roger W., Abraham H. Black, and Arnold E. Horowitz. 1955. Phonetic symbolism in natural languages. *The Journal of Abnormal and Social Psychology* 50: 388–93. [CrossRef]

Childers, Jane B. 2011. Attention to multiple events helps two-and-a-half-year-olds extend new verbs. *First Language* 31: 3–22. [CrossRef]

Childers, Jane B., and Michael Tomasello. 2002. Two-year-olds learn novel nouns, verbs, and conventional actions from massed or distributed exposures. *Developmental Psychology* 38: 967–78. [CrossRef] [PubMed]

Davis, Roger. 1961. The fitness of names to drawings: A cross-cultural study in Tanganyika. *British Journal of Psychology* 52: 259–68. [CrossRef]

Deloache, Judy S. 1995. Early Understanding and Use of Symbols—The Model Model. *Current Directions in Psychological Science* 4: 109–13. [CrossRef]

Deloache, Judy S., and Tanya Sharon. 2005. Symbols and Similarity: You Can Get Too Much of a Good Thing. *Journal of Cognition and Development* 6: 33–49. [CrossRef]

Dingemanse, Mark. 2013. Ideophones and gesture in everyday speech. *Gesture* 13: 143–65. [CrossRef]

Fort, Mathilde, Weiß Alexa, Martin Alexander, and Peperkamp Sharon. 2013. Looking for the bouba–kiki effect in prelexical infants. In *Proceedings of the 12th International Conference on Auditory–Visual Speech Processing, Annecy, France, 29 August–1 September 2013*. Edited by Slim Ouni, Frédéric Berthommier and Alexandra Jesse. Le Chesnay Cedex: Inria, pp. 71–76.

Fort, Mathilde, Imme Lammertink, Sharon Peperkamp, Adriana Guevara-Rukoz, Paula Fikkert, and Sho Tsuji. 2018. SymBouki: A meta-analysis on the emergence of sound symbolism in early language acquisition. *Developmental Science*. [CrossRef]

Gebels, Gustav. 1969. An investigation of phonetic symbolism in different cultures. *Journal of Verbal Learning and Verbal Behavior* 8: 310–12. [CrossRef]

Gentner, Dedre. 1982. Why nouns are learned before verbs: Linguistic relativity versus natural partitioning. In *Language development: Vol. 2. Language, thought, and Culture*. Edited by Stan A. Kuczaj. Hillsdale: Erlbaum, pp. 301–34.

Golinkoff, Roberta Michnick, and Kathy Hirsh-Pasek. 2008. How toddlers begin to learn verbs. *Trends in Cognitive Science* 12: 397–403. [CrossRef]

Goodrich, Whitney, and Carla L. Hudson Kam. 2009. Co-speech gesture as input in verb learning. *Developmental Science* 12: 81–87. [CrossRef] [PubMed]

Holland, Morris K., and Michael Wertheimer. 1964. Some physiognomic aspects of naming, or maluma and takete revisited. *Perception of Motor Skills* 19: 111–17. [CrossRef] [PubMed]

Imai, Mutsumi, and Sotaro Kita. 2014. The sound symbolism bootstrapping hypoethiesis afor language acquisition and language evolution. *Philosophical Transactions of the Royal Society B: Biological Sciences* 369: 20130298. [CrossRef]

Imai, Mutsumi, Etsuko Haryu, and Hiroyuki Okada. 2005. Mapping novel nouns and verbs on to dynamicaction events: Are verb meanings easier to learn than noun meanings for Japanese children? *Child Development* 76: 340–55. [CrossRef]

Imai, Mutsumi, Sotaro Kita, Miho Nagumo, and Hiroyuki Okada. 2008a. Sound-symbolism between a word and an action facilitates early verb learning. *Cognition* 109: 54–65. [CrossRef]

Imai, Mutsumi, Lianjing Li, Etsuko Haryu, Hiroyuki Okada, Kathy Hirsh-Pasek, Roberta Michnick Golinkoff, and Jun Shigematsu. 2008b. Novel noun and verb learning in Chinese-, English-, and Japanese-speaking children. *Child Development* 79: 979–1000. [CrossRef]

Imai, Mutsumi, Michiko Miyazaki, H. Henny Yeung, Shohei Hidaka, Katerina Kantartzis, Hiroyuki Okada, and Sotaro Kita. 2015. Sound symbolism facilitates word learning in 14 month olds. *PLoS ONE* 10: e0116494. [CrossRef]

Iwasaki, Noriko, David P. Vinson, and Gabriella Vigliocco. 2007a. How does it hurt, "kiri-kiri" or "siku-siku"? Japanese mimetic words of pain perceived by Japanese speakers and English speakers. In *Applying Theory and Research to Learning Japanese as a Foreign Language*. Edited by Masahiko Minami. Newcastle upon Tyne: Cambridge Scholars Publishing, pp. 2–19.

Kantartzis, Katerina, Mutsumi Imai, and Sotaro Kita. 2011. Japanese sound-symbolism facilitates wordlearning in English-speaking children. *Cognitive Science* 35: 575–86. [CrossRef]

Kersten, Alan W., and Linda B. Smith. 2002. Attention to novel objects during verb learning. *Child Development* 73: 93–109. [CrossRef]

Kita, Sotaro. 1997. Two-dimensional semantic analysis of Japanese mimetics. *Linguistics* 35: 379–415. [CrossRef]

Kita, Sotaro. 2001. Semantic schism and interpretive integration in Japanese sentences with a mimetic: A reply to Tsujimura. *Linguistics* 39: 419–36. [CrossRef]

Kita, Sotaro, Asli Özyürek, Shanley Allen, Tomoko Ishizuka, and Mihoko Fujii. Forthcoming. The role of iconicity in symbolic development: Children's preference for sound symbolic words and their early coupling with iconic gestures. Under review.

Klank, Linda J. K., Yau-Huang Huang, and Ronald C. Johnson. 1971. Determinants of success in matching word pairs in tests of phonetic symbolism. *Journal of Verbal Learning and Verbal Behavior* 10: 140–48. [CrossRef]

Köhler, Wolfgang. 1947. Gestalt Psychology, 2nd ed.Liverlight: New York. First published in 1929.

Kunihara, Shirou. 1971. Effects of the expressive voice on phonetic symbolism. *Journal of Verbal Learning and Verbal Behavior* 10: 427–29. [CrossRef]

Maguire, Mandy J., Elizabeth A. Hennon, Kathy Hirsh-Pasek, Roberta Michnick Golinkoff, Carly B. Slutzky, and Jennifer Sootsman. 2002. Mapping words to actions and events: How do 18-month-olds learn a verb? Paper presented at 26th Annual Boston University Conference on Language Development, Boston, MA, USA, November 2–4; Edited by Paola Escudero, P. Boersma, B. Skarabela, S. Fish and A. H. J. Do. vols. 1 and 2, pp. 369–82.

Maguire, Mandy J., Kathy Hirsh-Pasek, Roberta Michnick Golinkoff, and Amanda C. Brandone. 2008. Focusing on the relation: Fewer exemplars facilitate children's initial verb learning and extension. *Developmental Science* 11: 628–34. [CrossRef]

Marentette, Paula, and Elena Nicoladis. 2011. Preschoolers' interpretations of gesture: Label or action associate? *Cognition* 121: 389–99. [CrossRef] [PubMed]

Maurer, Daphne, Thanujeni Pathman, and Catherine J. Mondloch. 2006. The shape of boubas: Sound-shape correspondences in toddlers and adults. *Developmental Science* 9: 316–22. [CrossRef] [PubMed]

McGregor, Karla K., Katharina J. Rohlfing, Allison Bean, and Ellen Marschner. 2009. Gesture as a support for word learning: The case of under. *Journal of Child Language* 36: 807–28. [CrossRef] [PubMed]

McNeill, David. 1992. *Hand and Mind*. Chicago: University of Chicago Press.

Mumford, Katherine H., and Sotaro Kita. 2014. Children use gesture to interpret novel verb meanings. *Child Development* 85: 1181–89. [CrossRef]

Namy, Laura L. 2008. Recognition of iconicity doesn't come for free. *Developmental Science* 11: 841–46. [CrossRef]

Namy, Laura L., Aimee L. Campbell, and Michael Tomasello. 2004. The changing role of iconicity in non-verbal symbol learning: A U-shaped trajectory in the acquisition of arbitrary gesture. *Journal of Cognition and Development* 5: 37–57. [CrossRef]

Ohala, John J. 1994. The frequency code underlies the sound-symbolic use of voice pitch. In *Sound Symbolism*. Edited by Leanne Hinton, Johanna Nichols and John J. Ohala. Cambridge: Cambridge University Press, pp. 325–47.

Quine, Willard Van Orman. 1960. *Word and Object*. Cambridge: MIT Press.

Ramachandran, Vilayanur S., and Edward M. Hubbard. 2001. Synaesthesia—A window into perception, thought, and language. *Journal of Consciousness Studies* 8: 3–34.

Sapir, Edward. 1929. A study in phonetic symbolism. *Journal of Experimental Psychology* 12: 225–39. [CrossRef]

Siegel, Allen, Irwin Silverman, and Norman N. Markel. 1965. On the effects of mode of presentation on phonetic symbolism. *Journal of Verbal Learning and Verbal Behavior* 6: 171–73. [CrossRef]

Snape, Simon, and Andrea Krott. 2018. The benefit of simultaneously encountered exemplars and of exemplar variability to verb learning. *Journal of Child Language* 45: 1412–22. [CrossRef]

Tomasello, Michael, Tricia Striano, and Phillippe Rochat. 1999. Do young children use objects as symbols? *British Journal of Developmental Psychology* 17: 563–84. [CrossRef]

Tsuru, Shigeto, and H. S. Fries. 1933. Sound and meaning. *Journal of General Psychology* 8: 281–84. [CrossRef]

Werner, Heinz, and Bernard S. Kaplan. 1963. *Symbol Formation: An Organismic-Developmental Approach to Language and the Expression of Thought*. London: Wiley.

Yoshida, Hanako. 2012. A cross-linguistic study of sound symbolism in children's verb learning. *Journal of Cognition and Development* 13: 232–65. [CrossRef] [PubMed]

Yoshida, Hanako, and Linda B. Smith. 2003. Sound symbolism and early word learning in two languages. In *Proceedings of the Twenty-Fifth Annual Conference of the Cognitive Science Society*. Edited by Richard Alterman and David Kirsh. Boston: Cognitive Science Society, pp. 1287–92.