Symbolic play connects to language through visual object recognition
Linda B. Smith and Susan S. Jones
Department of Psychological and Brain Sciences, Indiana University-Bloomington, USA

Abstract
Object substitutions in play (e.g. using a box as a car) are strongly linked to language learning and their absence is a diagnostic marker of language delay. Classic accounts posit a symbolic function that underlies both words and object substitutions. Here we show that object substitutions depend on developmental changes in visual object recognition: 18- to 30-month old children (n = 63) substitute objects in play after they have developed the adult-like ability to recognize common objects from sparse models of their geometric structure. These developmental changes in object recognition are a better predictor of object substitutions than language or age. A developmental pathway connecting visual object recognition, object name learning, and symbolic play is proposed in which object substitutions are like the canary in the coal mine: they are not causally related to language delay, but their absence is an easily detected signal of a problem in language acquisition.

Introduction
Two-year-old children often play with objects in a way that has been of special interest to researchers of early language. In this play, children substitute one object for another – for example, using a pot as a hat, a stick as a sword, or a cardboard box as a boat (Bergen, 2002; Bretherton, O’Connell, Shore & Bates, 1984; McCune, 1995; Piaget, 1962). These object substitutions are linked to early language development, with their absence being a diagnostic marker of significant language delay (e.g. Bergen, 2002; McCune, 1995; Rescorla & Goossens, 1992; Rutherford, Young, Hepburn & Rogers, 2006). Recent findings suggest that developments in visual object recognition – a seemingly unrelated domain – are also related to early language, and the absence of those developments also predict language delay. Visual object recognition and object substitutions in play have not been considered in relation to each other. In the present study, we examine the possibility that developmental changes in visual object recognition play a role in connecting object substitutions in play to language development.

‘Symbolic play’ is an umbrella term used to refer to a range of pretend play behaviors including dress-up and role-playing as well as object substitutions (Lilliard, 2001; Lewis, Boucher, Lupton & Watson, 2000; McCune-Nicolich, 1981). Many of these early forms of representational play are related to early language learning (McCune, 2008). However, object substitutions are the form of symbolic play that has been most systematically related to future language development (e.g. Lewis et al., 2000; Shore, O’Connell & Bates, 1984; Rescorla & Goossens, 1992), and it is the play most widely used in clinical assessments of language and other developmental disorders (e.g. Johnson, DesJardin, Quittner & Winter, 2008; Lewis et al., 2000; O’Toole & Chiat, 2006). This form of play emerges in typically developing children between 18 and 30 months – the same age range in which children’s object name vocabularies are rapidly expanding (Bergen, 2002; Lewis et al., 2000; McCune-Nicolich, 1981; McCune, 1995; Shore et al., 1984). The tie between object substitutions and language development is classically attributed to a shared ‘symbolic function’: for example, the pot on the child’s head and the word hat both stand for a real hat (e.g. Baron-Cohen, 1987; Lilliard, 1993; Nichols & Stich, 2000; Piaget, 1962; Winner, 1979). Consistent with this idea, a number of researchers (McCune-Nicolich, 1981; Shore et al., 1984; Striano, Tomasello & Rochat, 2001) have noted constraints on the types of objects that children substitute for others. The substituted objects tend to be simple in shape and to have minimal surface details, and thus perhaps are symbol-like. Thus, a banana might be substituted for a phone, but a richly detailed toy truck would not be.

Critically, the shape of the substituted object is also geometrically similar to the shape of the replaced object. This observation suggests a possible relation between the emergence of object substitutions in play and recent...
findings about developmental changes in visual object recognition that occur between 18 and 30 months (Smith, 2003, 2009). Early in this period, children appear to recognize instances of common categories primarily by their surface properties and surface features (Rakison, 2003; Pereira & Smith, 2009); at the end of this period, children also use geometric properties of 3-dimensional shape (Smith, 2003; Pereira & Smith, 2009). The ability to recognize common objects from sparse representations of their 3-dimensional geometric structures is well established in adults (Biederman, 1987; Hummel, 2000). Biederman (1987) has shown that adults readily recognize instances of basic-level categories given representations composed of just 2–4 simple geometric components (volumes such as cylinders, cones, spheres, cubes) in the proper spatial structure. In theories of adult visual object recognition, this ability is linked to the generalized recognition of multiple views of an object and of varied instances of the same category (cites). Using both linguistic and non-linguistic tasks, Smith (2003) tested young children’s ability to recognize common categories from such Biederman-like caricatures of their 3-dimensional shapes. Children with vocabularies containing more than 100 object names recognized these sparse geometric models as well as they did richly detailed instances. Children with fewer than 100 object names recognized only the detailed examples, not the sparse models. In brief, these findings indicate that the ability to recognize objects from sparse geometric models of 3-dimensional object shapes develops early in young children and is linked to individual children’s object name vocabulary size.

Smith’s (2003) findings have been replicated (Augustine, Jones & Smith, in press; Jones & Smith, 2005; Pereira & Smith, 2009; Son, Smith & Goldstone, 2008) and extended in two ways relevant to the present study. First, Son et al. (2008) asked children to match 3-dimensional objects by shape and found that they were more successful if given a sparse geometric representation of a few major parts than if given richly detailed and realistic objects. Thus, children are better able to match 3-dimensional shapes when surface details are minimized, just as they are more likely to substitute objects in play that have simple shapes without many surface features. Second, Jones and Smith (2005) showed that children with delayed vocabulary development were less able than typically developing children to recognize common objects from sparse models of geometric shape. Thus, there is a link between delays in the development of object recognition and in language learning, just as there is a link between delays in object substitutions in play and in language learning.

Therefore, the question arises as to whether the emergence of object substitutions in play may be a manifestation of a fundamental change in children’s visual object representations. Answering this question is important for understanding how advances in visual object representations may support cognitive development in other domains. In addition, since the recognition of objects from their geometric structure and object substitutions in play have both been separately linked to language development, evidence on their relationship to each other may provide deeper insights into the cascading consequences of early word learning beyond the realm of language itself. Accordingly, the following study measures the relation between individual children’s recognition of common objects from sparse Biederman-like models of their geometric shapes and children’s use of object substitutions in play. The study also examines the relation of each of these developments to object name vocabulary size and age.

Method

Participants

Because our primary analytic approach to assessing the developmental relationship between visual object recognition and object substitutions was a correlational design, we recruited children representing a broad range of vocabulary sizes within the target developmental period of 18 to 30 months. There were no restrictions on participation other than no known developmental or neurological disorders. Sixty-three children (32 males) aged 18 to 27 months \( (M = 21.3 \text{ mos}, SD = 1.84) \) participated. Parents reported children’s productive word vocabularies on the MacArthur-Bates Communicative Development Inventory (Fenson, Dale, Reznick, Bates, Thal & Pethick, 1994). Because nouns in particular have been shown to be related to object recognition, we also report the numbers of nouns in children’s productive vocabularies. The range was 5 to 301 nouns with a mean of 112.8 nouns \( (SD = 83.9) \). This broad sample of vocabulary size within the target age range enables a generalizable assessment of the correlation between visual object recognition, object substitutions, and noun vocabulary size.

Design

Children were tested in an object recognition task and a thematic play task with both rich and sparse objects. The object recognition task, like one task used by Smith (2003), was a name comprehension task: children were asked to indicate which of three objects was the named item. Performance with the richly detailed objects provided a measure of the children’s receptive understanding of the object names and performance with the sparse objects provided a measure of their use of geometric structure to recognize instances. The object substitution task involved no explicit naming of any of the objects, and little verbal instruction. Thematic play with the rich objects provided a measure of their understanding of the task and grasp of the intended theme, and inclusion of the sparse object in this play was the measure of object
substitution. For all children, the object recognition task was first and sparse recognitions preceded rich ones.

Procedures

Object recognition

Sparse and rich representations of the following six common categories were used: pizza, ice cream, cat, butterfly, camera and toothbrush. The rich representations were detailed and realistic toy instances. The sparse caricatures were constructed from 2–4 (Styrofoam) geometric components and painted grey, as illustrated in Figure 1. The rich and sparse objects ranged in volume from 10 cm³ to 18 cm³. For the sparse test trials, three objects were placed 10 cm apart on a tray. While holding the tray away from the child, the experimenter asked the child to indicate one named object, e.g. ‘Where is the camera? Show me the camera.’ After the question was asked, the tray was moved forward so the child could respond. The first object selected was recorded as the child’s choice. Each category was tested once for a total of six trials. These were randomly ordered for each child as were the spatial locations of the requested objects. For the rich object test trials, which followed the sparse test trials, the procedure was identical.

Symbolic play

The play task in the present experiment is one commonly used to elicit object substitutions in experimental studies and clinical assessments (e.g. Shore et al., 1984; Striano et al., 2001; Johnson et al., 2008; Lewis et al., 2000). In the present variant, the child was presented with a set of thematically related objects with one typical component in the theme missing, and with a simple object of similar geometric shape to that of the missing object available for substitution. We refer to this simple object as the ‘target’ object in each set. Figure 1 shows the ‘sleeping’ theme set used in the present experiment: a doll, a pillow and a blanket, and a block (the target) that could be used as a bed. The other two sets were an ‘eating’ theme set consisting of a doll, a spoon, a plate, a cup, and small spheres (the target) that could be used to represent food; and a ‘person action’ set consisting of an open-sided building, a chair, an airplane, and a peg (the target) that could be used as a person.

Before testing began, children were introduced to the play task. The experimenter gave the child a set of four warm-up toys that were not thematically related (a teddy bear, a ball, a car and a flower), said ‘You can play with these’, and encouraged the child to engage with each of the objects. After this warm-up, each play set was presented for 90 seconds. The experimenter said ‘Play with these’: the only added encouragement provided was to direct the children’s attention in a general way (e.g. with a sweeping gesture) to the objects.

The order of presentation of the three play sets was counterbalanced across children. Each act on an object – both sparse and rich – was coded (‘yes’ or ‘no’) as fitting the specified thematic role (e.g. putting the blanket on the doll; using the block for a bed) or not (e.g. spinning the doll on the tabletop). Agreement in two judges’ independent coding of a randomly sampled 25% of the acts was 97%.

Results

The main empirical question is whether there is a relation between the recognition of sparse models of objects in the Object Recognition task, and the use of simple geometric forms to substitute for missing objects in thematic play. Before considering the correlational analyses that address this question, we first describe children’s performances as a group in the recognition and play tasks with both the rich and the sparse objects.

In the Object Recognition task, which is also a receptive language task, the children as a group readily recognized the rich instances of these common categories (Mean = .72 correct, SD = .23) showing receptive knowledge of the category names. However, given these same names, they were much less likely to recognize the sparse geometric caricatures of the same things (Mean = .48 correct, SD = .24; t(62) = 7.05, p < .001). Importantly, this general group difference does not characterize all children in the sample. Indeed, 40% of the children correctly identified four or more of the six sparse objects.

In the Symbolic Play task, we scored thematic actions with both the rich and the sparse objects. To compare the thematic use of rich and sparse objects, we categorized play with each of the three sets as to whether the child exhibited thematic play that involved the rich objects and thematic play that involved the sparse target. The categorization criterion for rich object thematic play with a set was two theme-related actions, and the categorization criterion for sparse object thematic play with a set was also two theme-related actions on that object. This criterion was designed to minimize the over-interpretation of accidental actions that resembled thematic play. By
this measure, children played thematically with the rich objects in .78 of the three sets on average (SD = .28) showing that they understood the task. Children as a group included the target object in thematic play in only .58 of the three sets on average (SD = .31). The difference between thematic play with rich and sparse objects by this categorical measure is reliable (t(62) = 7.31, p < .001). For the correlational analyses, we also counted the total numbers of thematic acts involving the target object across all three play sets. Children performed an average of 4.49 thematically related acts involving the target objects across all three play sets. There was considerable variation among children in this measure (SD = 4.43, range = 0–11 actions).

Figure 2 provides evidence on the main question: is the recognition of sparse geometric models of common object categories predictive of object substitutions in play? As shown in the figure, there is a strong relation between children’s sparse object recognition scores and the numbers of target object substitutions they produce in play (r = 0.52; t(61) = 4.76, p < .001). Both tasks involve simplified geometric objects but the use of such objects in the object recognition and symbolic play literatures has been thought about in different ways. In the object recognition literature, these abstract geometric representations of known objects are thought to be physical instantiations of the internal representations that enable view-independent and generalizable recognition of common objects. In the symbolic play literature, the simple forms have principally been thought of as simple signs upon which symbols will be built (Bates et al., 1979). The correlation between children’s performances in the two tasks supports the idea that the advances in visual object recognition due to the development of abstract geometric representations may make object substitution in play possible.

The object recognition task is also a name comprehension task, whereas the play task is a non-linguistic task. Prior studies of both object recognition and object substitution indicate that both are correlated with language development. The object recognition task with rich objects is a receptive language task, albeit with only six categories. Performance in this rich object recognition task was not related to sparse object recognition (r = 0.05, ns), nor to object substitutions (r = −0.037, ns). However, productive noun vocabulary was related to both sparse object recognition (r = 0.39, t(61) = 3.31, p < .01) and to object substitutions (r = 0.29; t(61) = 2.28, p < .03). This suggests that the critical aspect of language development may not be knowing specific names for specific object categories, but instead acquiring a sufficient number of object categories to support the abstraction of the geometric properties of the objects.

We carried out a step-wise multiple regression analysis, with number of object substitutions the predicted variable, to examine the relations among productive noun vocabulary, sparse object recognition, and object substitutions. We did not consider Age because it was not related to either sparse object recognition (r = −0.06, ns) or object substitutions (r = 0.08, ns); nor did we consider rich object recognition scores because, as reported above, these were also unrelated to sparse object recognition or object substitutions. In the regression, only sparse object recognition was a significant predictor of object substitutions (Multiple R² = 0.27; beta = 0.52; F(1, 61) = 22.63, p < .001). The language measure made no additional independent contribution (partial correlation = 0.10, F(1, 61) = 0.65, ns). This result – that attention to the sparse geometric shapes of common objects is a better predictor of symbolic play than language – suggests a new understanding of the processes that support object substitutions in play – that is, that object substitutions depend upon the ability to perceive the geometric structure of 3-dimensional shape.

Because previous research (Smith, 2003; Pereira & Smith, 2009) has shown differences in visual object recognition between children grouped by object name vocabulary size, we looked for similar differences between children in this study by dividing them into five vocabulary groups: children with 0 to 25 nouns (n = 14; M = 12.5 nouns; SD = 6.8), 26 to 50 nouns (n = 12; M = 36.5 nouns; SD = 5.5), 51 to 100 nouns (n = 14; M = 80.2 nouns; SD = 13.5), 100 to 200 nouns (n = 13; M = 128.5 nouns; SD = 24.6), and more than 200 nouns (n = 10; M = 257.2 nouns; SD = 41.0). As noun vocabulary size increases across groups, so too does mean age (F(4, 58) = 3.00, p < .03). However, the differences in age among the groups are small (mean ages are 20.15, 20.73, 21.93, 21.2, and 22.6) and only the comparison of the first with the last is significant (t(22) = 2.89, p < .01). Figure 3 shows the mean proportions of sparse and rich objects correctly identified by each vocabulary group in

Figure 2 Scatterplot of performances in the two tasks. Number of object substitutions (use of the ambiguous object in the specified thematic role) as a function of number of sparse representations that were recognized in the object recognition task. (Locations of individual scores have been jittered by less than 2% to reveal density of points in the plot.)

© 2011 Blackwell Publishing Ltd.
the object recognition task (Panel A) and the mean proportions of sparse and rich objects that each group included in thematic play (Panel B). As is apparent in Panel A of the figure, recognition and play with both sparse and rich objects increased across the five vocabulary groups. The important result is that the only significant differences in sparse object recognition (pairwise t-tests, Bonferroni correction) are between the two lowest vocabulary groups and the three highest vocabulary groups. Children in the two lowest vocabulary groups did not recognize sparse objects at above chance levels (= .33), whereas children in the three highest vocabulary groups did. This finding is similar to the relation between vocabulary size and sparse object recognition reported by Pereira and Smith (2009). There is also a similar pattern shown in Panel B of Figure 3, where the largest difference in thematic play with sparse objects is between the second and third vocabulary groups \((t(24) = 1.67, p = .054\) one-tailed; in post-hoc t-tests with Bonferroni correction, the only reliable between-group comparison is between the second vocabulary group and the fifth vocabulary group). The biggest increase in object substitutions (thematic play with the sparse object) also occurred in the 51–100 noun group. The relevance of these vocabulary group comparisons is that they align with previous analyses in studies of children’s object recognition (Pereira & Smith, 2009). The major finding in this study is that these developments in visual object recognition strongly predict object substitutions in play as shown in Figure 2.

**General discussion**

Children’s object substitutions in play have been studied in relation to language because the behavior is a strong predictor of healthy language development, and also researchers have hoped that this form of play might provide insights into the development of symbolic representation. Changes in children’s visual object recognition have been studied from the perspective of how experience with categories and multiple instances may yield more abstract and more robust visual object recognition (Pereira & Smith, 2009; Smith, 2009). The present results suggest a previously unsuspected connection between developments in these two domains and a developmental pathway that may connect early language to symbolic play via the visual object recognition system. The present data show that it is when children can recognize instances of common categories from sparse geometric models of 3-dimensional shape that they begin to substitute similarly shaped objects for one another in play.

Past research (Smith, 2003; Pereira & Smith, 2009) as well as the present results indicate that changes in visual object recognition are related to noun vocabulary size. As yet, there is no clear evidence as to the direction of this link. However, past research shows that learning object names and categories trains children’s attention to shape (e.g. Perry, Samuelson, Malloy & Schiffer, 2010; Smith, Jones, Landau, Gershkoff-Stowe & Samuelson, 2002). Further, computational models of object recognition succeed on the assumption that sparse representations of 3-dimensional object shape are a consequence of category learning (Edelman & Intrator, 2003; Doumas & Hummel, 2010). Thus, abstracting the underlying geometric structure of an object’s shape – an ability important to mature object recognition and, by the present results, to object substitutions in play – may depend on learning names for a range of basic-level category instances (see Perry et al., 2010). The formation of these more abstract representations of object shape may in turn support the more rapid and efficient learning of additional object names (Son et al., 2008, Smith et al., 2002; Perry et al., 2010). By this account, learning object names builds geometric object representations and speeds up new category learning. The more abstract object representations also support object substitutions in play. Thus, the relation between language and object

![Figure 3](image-url) **Figure 3** Panel A: Mean proportion correct recognition of rich and sparse objects for children at five different productive noun vocabulary levels. Panel B: Mean proportion of sets meeting the criterion for thematic play with rich and sparse objects (object substitutions) for the same five groups of children.
substitutions may be mediated (Baron & Kenny, 1986) by changes in visual object recognition.

Figure 4 provides a proposed developmental pathway that illustrates these relations and also explains why failure to develop object substitution is diagnostic of future language delay. The idea is that object substitutions are like the canary in the coal mine: they are not causally related to language delay, but their absence is an easily detected signal of a problem in language acquisition. As shown in the figure, early learning of object names promotes (and is supported by) the formation of increasingly abstract models of 3-dimensional shape. These newly formed representations invite and support the substitution of geometrically appropriate objects for one another in play. These substitutions are predictive of later language development because later language is causally dependent on early language development.

Figure 4

Proposed causal connections among language development, object recognition, and symbolic play. Arrows indicate hypothesized causal relations.

Early language development (which consists mostly of acausal dependency on early language development. The recognition and these changes in object recognition lead, to symbolic play. These substitutions are predictive of causally related to language delay, but their absence is an easily detected signal of a problem in language acquisition. As shown in the figure, early learning of object names promotes (and is supported by) the formation of increasingly abstract models of 3-dimensional shape. These newly formed representations invite and support the substitution of geometrically appropriate objects for one another in play. These substitutions are predictive of later language development because later language is causally dependent on early language development.

Early language development (which consists mostly of learning object names) supports changes in visual object recognition and these changes in object recognition lead, along with other developments, to symbolic play. The absence of object substitutions in children’s play is thus a surface sign of a weakness in language learning.

There is no implication in this proposal that changes in visual object recognition are the only developments relevant to object substitutions. Striano and colleagues (Striano et al., 2001; Rakoczy, Tomasello & Striano, 2005) have shown that children’s object substitutions in thematic play are also strongly predicted by the frequencies with which their parents engage in such activities, and thus depend on a social model. The present results suggest a new and testable prediction about when social modeling will be effective. If children’s recognition of parents’ object substitutions depends on their extracting the abstract geometric similarity between the substituted object to the intended category, then only children who can recognize sparse representations of common categories should be influenced by social modeling. Likewise, although the present results do not provide evidence for a general symbolic function that links object substitutions and language, children’s object substitutions may nonetheless be relevant to developing symbolic behavior (Rakoszy, Tomasello & Striano, in press). The objects substituted in play are signs (not symbols) in that they resemble their referents, and in this way are like early gestures and scale models (Deloache, 1995; Iverson & Goldin-Meadow, 2005; Namy, Campbell & Tomasello, 2004).

Considerable research on gestures and scale models suggests that iconic signs, which are simple abstractions, serve as stepping stones in children’s understanding of symbols (DeLoache, 1995; Hoiting & Slobin, 2007; McNeil, 2005). The present results indicate that, for object substitutions, the relevant simplification concerns the representation of 3-dimensional object shape, which is also a core achievement of the human object recognition system. Finally, although developmental changes in visual object recognition may be essential to the earliest form of object substitutions in play – minimal objects with fundamental shape similarities to the real object – more advanced forms of object substitutions have also been observed (e.g. using a car as a hairbrush; see Watson & Fisher, 1977; McCune-Nicolich, 1981) that may emerge from these earlier advances but also involve other cognitive abilities.

These considerations and the present findings highlight the cascade that is developmental process – that development consists of many interacting and mutual dependencies across systems that may seem at first unrelated (Thelen & Smith, 1994). The results also focus attention on object recognition as a component of developmental change in what on the surface appears to be an unrelated competency. In this case, changes in visual object recognition matter to the emergence of object substitutions in play, and may be the source of the link of these object substitutions to language development. There may be other unsuspected consequences of ongoing changes in object perception and representation. The development of visual object recognition has not been well studied, particularly outside of infancy (see Gerhardstein, Shroff, Dickerson & Adler, 2005; Smith, 2009, for reviews). This is beginning to change, and there is increasing evidence for significant developments in late infancy and perhaps through early childhood (see Smith, 2009). Humans are visual animals and the present results suggest that the increasing sophistication of children’s visual object recognition is likely to play important roles in producing developmental changes in many cognitive domains.

Acknowledgements

This research was supported by a grant from the National Institute for Child Health and Development (R01HD 28675) to both authors. We thank Char Wozniak, Lynn Freeman, and Elizabeth Hannibal for their assistance in collecting and coding the data.

References


Received: 25 September 2010
Accepted: 25 February 2011