Constructions as Categories of Language

Adele E. Goldberg1
Princeton University

Devin Casenhiser
York University

Tiffani R. White
Princeton University

Draft of paper for *New Ideas in Psychology*

---

1 Correspondence should be addressed to Adele Goldberg at adele@princeton.edu.
Abstract
What causes children to categorize distinct utterances they hear into a constructional generalization? That is, what makes subjects create a constructional category instead of treating each utterance as a distinct unrelated idiom? One simple factor that encourages the learning of abstract categories is shared concrete similarity. When instances share concrete attributes, learners are more likely to categorize them together, and moreover are more likely to attend to their more abstract commonalities (Gentner & Medina, 1998; Markman & Gentner, 1993). This paper reports results that confirm the prediction that presentation of items with concrete shared similarity early in training enhances language learning.
Children must generalize over the utterances they hear so that they can creatively produce and understand utterances they have never heard before. One factor that enables language to be used creatively is the fact that, within a given language, formal patterns correlate strongly with the meanings of the utterances in which they appear (see Goldberg, Casenhiser & Sethuraman, 2005, for a measure of such correlations). These correlations can be thought of as sets of "linking rules" or as argument structure constructions (Goldberg, 1995).

Constructionist approaches to language argue that generalizations about form and meaning—constructions—are learned. Much research in this tradition has emphasized how little evidence of generalization there is in young children’s productions (for reviews see, e.g., Tomasello, 2000; Dabrowska, 2004). Surprisingly little data has been found that has identified particular facilitory or inhibitory factors in learning constructions, beyond varying overall exposure (Vasilyeva, Waterfall, & Huttenlocher, 2006, but cf. Childers & Tomasello, 2001). And yet generalization is part and parcel of learning a language: knowledge of language does not consist of a set of unrelated item-based facts, but is instead a rich interconnected network, containing both specific and general knowledge. Therefore the question of how learners form generalizations is of central interest.

There has been much work on learning generalizations within artificial grammar learning paradigms (Gomez, 2002; Hudson Kam & Newport, 2005; Marcus, Vijayan, Bandi Rao, Vishton, 1999, Saffran, 2001; Saffran, 2002; Saffran & Wilson, 2003; Valian & Coulson, 1988). However, there have been virtually no previous studies that have trained children to map a novel word order onto a novel meaning: exactly the task that the child faces when naturalistically learning language.

Certain traditions within linguistics have led researchers to expect to find learning abilities that are specific to the domain of language. However, clear-cut instances of such domain-specific abilities have proven elusive. Recent investigations have revealed that many abilities that had been thought to be domain-specific are in fact domain-general. For example, the ability to track transitional probabilities, used to determine word boundaries in ongoing speech has been found in cotton top tamarin monkeys (Hauser, Newport, & Aslin, 2001) as well as humans (Saffran, Johnson, & Newport, 1996). Moreover, humans track transitional probabilities when they are exposed to a series of tones or pictures as well as sounds (Saffran, Johnson, Aslin, Newport, 1999; Fiser & Aslin, 2002). Categorical perception of phonemes, once thought to be a language-specific ability (Eimas, Siqueland, Jusczyk, & Vigorito, 1971) has been shown to exist in chinchillas as well as humans (Kuhl & Miller 1975). The Perceptual Magnet effect (Kuhl, 1991) has been discovered in birds as well as in humans (Kluender, Lotto, Holt, & Bloedel, 1998). The syllable has been argued to necessarily be part of a specific-to-language genetic endowment (Yang, 2004), but others have countered that syllabification is a by-product of more general motor-level properties (MacNeilage & Davis, 2005). Hauser, Chomsky and Fitch (2002) go so far as to suggest that the only domain-specific aspect of language may be the recursive interfaces to phonology and semantics; moreover, they concede that even this may not be specific to language since the ability to plan and the ability to understand others’ states of mind also appear to rely on recursive interfaces. In fact, starling songbirds have recently been shown to be capable of learning to classify acoustically recursive patterns of the form $A^nB^n$ (Gentner, Fenn, Margoliash, & Nubaum, 2006) (cf. also Jackendoff & Pinker, 2005 for arguments that recursive interfaces are not specific to language).

---

2 Pinker and Jackendoff (2005) and Jackendoff and Pinker (2005) argue that domain-specific abilities have in fact evolved for the purpose of language, but their examples are primarily aimed at processes involved in speech perception and articulation. Although, as noted above, much of speech perception and articulation is at least arguably domain-general, whether or not there turn out to be any domain-specific abilities related to speech is not crucial to our purposes, since our concern is with grammar learning.
This is not to say that there are no differences between humans and non-human primates. There obviously are. Humans learn language and other primates do not. But language is just one way that humans differ from other primates. It is therefore far from obvious that the abilities that enable us to acquire language are biologically determined abilities that are specific to language. Non-human primates do not share our impressive aptitude to understand others’ intentions in cooperative contexts (Tomasello, 1999; 2003), nor do they share our strong instinct to imitate conspecifics (Tomasello 1999; Horner & Whiten 2005). As Grice (1975) observed, the ability to understand another’s communicative intentions in cooperative contexts is a prerequisite to linguistic abilities. As Tomasello (1999) has emphasized, the human predilection to imitate is another important prerequisite for language insofar as a myriad of idiosyncratic facts about each language, from the way particular sounds are produced, to the form of words to idioms to idiosyncratic aspects of a particular language’s grammar must be learned by imitating the speech that is heard.

Given the present controversy surrounding domain-specific linguistic abilities, we take the null hypothesis to be that language is “a new machine built out of old parts” (Bates, 2004; Elman et al.,1996). That is, our working assumption is that the grammars of languages are learnable on the basis of domain-general cognitive processes. Given this perspective, we can investigate well-understood domain-general processes in order to address the question of how (and if) such processes may provide an account of how language is learned. Such research can provide valuable insights into cognitive processes in general and how these processes function in relation to language. They also may offer a better understanding of the nature of language itself—regardless of whether some aspect of language eventually requires recourse to domain-specific processes.

I. Categorization

This paper focuses on one ability that is uncontroversially domain-general: the ability to form categories. People perform rudimentary acts of categorization thousands of times each day. Each time we open a new door, we categorize it, with its unique marks, doorknobs and color as an instance of a door; this is what enables us to know what to do with it. Infants have been shown to be able to form categories based on perceptual properties as early as 4 months (Colombo, McCollam, Colden, Mitchell, & Rash,1990; Eimas & Quinn, 1994; Oakes, Coppage, & Dingel, 1997; Quinn & Eimas, 1998; Quinn, Eimas, & Rosenkrantz, 1993; Rakison & Butterworth, 1998), to form categories based on relational properties by 10-months (Casasola & Cohen, 2002; Casasola, Cohen, & Ciarello, 2003), to form prototypes (Younger, 1985), and learn the correlations among features that determine the boundaries of categories (Younger & Cohen, 1986). Therefore it is clear that the capacity to categorize develops early enough to account for language learning in principle; moreover, the ubiquity of categorization in early cognition suggests that it is an inevitable process by which children analyze their surroundings.

Neither children nor adults categorize randomly or completely. We do not form a category that includes, for example, rooms that have exactly two windows and two doors. At the same time, we do form a category of rooms that contain both a refrigerator and a stove (the category of *kitchens*). Why is it that we sometimes form categories and other times not, and what is it about argument structure patterns in particular that encourages young children to learn them? We form categories in order to make sense of the world around us. Exactly *which* categories are formed is determined to a great extent by the usefulness of a potential category in predicting how the environment will behave. Being able to categorize the form of a ball as distinct from a surface is the first step in being able to predict what will happen when the ball rolls beyond the end of the table. Function plays an obvious and important role in determining what is categorized and affords similar predictive value to the category. Knowing what a kitchen is, for example, allows us to make many relevant predictions: it is the room where meals are prepared, it also contains a
sink and a stove, there is typically only one such room per living space. This is information that wouldn’t necessarily be gleaned from the form of the room alone. When it comes to language, clearly the relevant predictive tasks are to predict meaning, given the form (comprehension) and to predict form, given meaning (production). Thus given the fact that constructions relate critical aspects of form and meaning, their acquisition is well-motivated. Accordingly, earlier work has emphasized the role of prediction in language learning, and has shown that constructions are good predictors of overall sentence meaning, useful for conveying “who did what to whom” (Bencini & Goldberg, 2000; Goldberg, Casenhiser, & Sethuraman, 2005).

A second explanation for the what and why of category formation is somewhat more mechanistic. Exemplars are more likely to be classified together if one exemplar reminds the learner of another (Ross et al., 1990). A number of studies involving non-linguistic categorization tasks have demonstrated, for example, that instances that share concrete similarities are more likely to be classified together; moreover, learners are more likely to notice additional abstract similarities among items when those items share concrete similarities as well. Markman and Gentner (1993) have observed that when similarity comparisons are made between two items, subjects are more likely to attend to shared abstract relational structure between the items as well via a process they refer to as structural alignment. In one study (Kotovsky & Gentner, 1996), children ages 4, 6 and 8 performed a triads choice task in which they were shown a standard and two alternative choices; subjects were asked which of the two choices the standard best “went with.” One choice shared a relational property with the standard (e.g., if the standard was xXx, one choice might share the relational property of symmetry: oOo). The other choice contained the same items, but without the shared relational structure (e.g., ooO). Kotovsky and Gentner’s results demonstrated that children were better able to identify the item with shared relational structure if the standard and choices shared concrete similarities (e.g., xXx was easier to match with oOo than with 121) (cf. also Markman & Maddox, 2003; Goldstone, 1994 and related findings in research on analogies by Gick & Holyoak, 1983; Holyoak & Koh, 1987).

Since both the surface as well as the thematic/conceptual properties are relational, construction learning is a good candidate for structural alignment. The input can help to focus the learner’s attention on the relevant features of the category by including concrete similarities between exemplars. These similarities serve as reminders to previous exemplars. In the following sections, we detail a number of previous studies that have demonstrated relevant effects in construction learning, and contribute to this body of literature by presenting a new set of experiments designed to further test the hypothesis that construction-learning is akin to general category-learning.

2. The role of categorization in learning argument structure constructions

Argument structure constructions are pairings of form and meaning that provide the means of expressing simple propositions in a language. Examples of argument structure constructions are provided below (see Goldberg, 1995 for discussion):

<table>
<thead>
<tr>
<th>Meaning</th>
<th>Form (example)</th>
</tr>
</thead>
<tbody>
<tr>
<td>X causes Y to receive Z</td>
<td>Subject Verb Object1 Object2 (She faxed him a letter)</td>
</tr>
<tr>
<td>X causes Y to become Z</td>
<td>Subject Verb Object ResultPhrase (She talked herself silly)</td>
</tr>
<tr>
<td>X causes Y to move to Z</td>
<td>Subject Verb Object PrepositionalPhrase (She sneezed the foam off the cappuccino)</td>
</tr>
<tr>
<td>X moves to Y</td>
<td>Subject Verb PrepositionalPhrase</td>
</tr>
</tbody>
</table>
The meanings of argument structure constructions are clearly relational, and so these constructions are candidates for categorization by way of structural alignment. Research on general categorization makes a clear prediction: the abstract relational structure of a construction should be learned more easily if instances of the construction witnessed during training share concrete similarities.

Surprisingly little work has trained subjects to map a novel abstract form onto a novel meaning: exactly the task that the child faces when naturalistically learning language. However, a new experimental paradigm has been developed which has been used to begin to address this gap (Casenhiser & Goldberg, 2005; Goldberg, Casenhiser, & Sethuraman, 2004). The paradigm involves training subjects on a novel construction. In a series of experiments, learners were exposed to 16 instances of the novel construction: their total exposure was less than 3 minutes in duration. The meaning assigned to the novel formal pattern was that of APPEARANCE: an entity appears in a location (a meaning novel for English phrasal patterns). Subjects watched a set of short video clips in which they saw objects appear in or on various locations. Each video clip was accompanied by an audio description whose syntactic form was composed as follows:

\[(1) \text{ noun phrase}_{(\text{theme})} \quad \text{noun phrase}_{(\text{location})} \quad \text{nonsense verb}\]

The audio description was given in the simple present tense at the start of the scene and was repeated in the past tense at the end of the scene. Given a scene where a flower appears in a queen’s hand, subjects heard, *The flower the queen fegs...the flower the queen fegged*. The entity named by the first noun phrase appeared in the place named by the second noun phrase. That is, the novel construction involved a non-English word order, Subject Object Verb, together with an abstract meaning that is novel for English constructions: something appears in a location.

Following training, the children performed a forced-choice comprehension task; they were asked to match an audio description to one of two video clips displayed simultaneously on a computer screen. One scene showed an object appearing in or on a particular location and the other showed the same object interacting with or acting on that location. For example, given the audio description *the sailor the pond naifoed*, one scene showed a sailor sailing his boat onto a pond (i.e., he begins off camera and sails into the scene) while the second scene showed the sailor sailing his boat around the pond (having been on camera the entire time). The audio description used the novel appearance construction, so the correct answer would be the first scene. The task is reminiscent of the preferential looking paradigm, the main difference being that our subjects provided an unambiguous behavioral response, pointing to the matching scene instead of simply looking longer at one scene than another.

Previous experiments have demonstrated that shared concrete similarity among instances of a novel construction does lead to better recognition of the abstract relational structure of the construction at test (Casenhiser & Goldberg, 2005; Goldberg et al., 2004). The shared concrete similarity that was used involved a shared nonsense verb. That is, in one previous study, six year olds were randomly and equally divided into three conditions: a SKewed FREQUENCY TRAINing condition, a BALANCED FREQUENCY TRAINing condition and a control condition (Casenhiser & Goldberg, 2005). In both training conditions, overall type and token frequency of the novel verbs used was held constant. In the balanced frequency training condition, subjects heard 5 different novel verbs, each with a relatively low token frequency of 1 or 2 (1-1-2-2-2). The skewed frequency condition was designed to test the hypothesis that children’s learning of a novel construction would be aided if a single verb appeared in a disproportionately large number of instances of the novel construction, in that way, providing more shared concrete similarity.
between items than in the balanced frequency condition. Therefore in this skewed frequency condition, subjects again hear the same 5 novel verbs, but this time one novel verb had an especially high token frequency of 4, while the other novel verbs were recorded once each (4-1-1-1-1). Each scene was repeated exactly twice in each condition, so that a total of 16 clips were witnessed. Any difference among groups can only be attributed to a difference in the linguistic input that subjects were exposed to, as all three conditions watched exactly the same video.

Results supported the prediction that shared concrete similarity would lead to more accurate categorization: children in the skewed frequency condition performed better than children in the balanced condition. Children in the balanced frequency training group outperformed those in the control group as well, suggesting that skewed frequency facilitates, but is not necessary for learning a novel construction. The scores of the balanced and the skewed frequency groups were also significantly greater than would be expected from chance performance, while the score of the control group did not differ significantly from chance. The differences between groups can only be attributed to a difference in the linguistic input, as all conditions watched exactly the same video.

The reason why token frequencies of particular novel verbs were varied instead of using some other sort of shared similarity across items is that the actual input children receive tends to be skewed in just this way. That is, the argument structure constructions children hear tend to be skewed disproportionately towards a single verb, even when they potentially occur with a much broader range of words or meanings. For example, Goldberg et al. (2004) investigated speech from mothers to young children in the Bates corpus from CHILDES (Bates, Bretherton, & Snyder, 1988 in MacWhinney, 1995). They reported that a particular construction is typically dominated by the use of that construction with one particular verb. For example, go accounts for a full 39% of the uses of the “intransitive motion” construction ((Subj) V Obj Obl[path/loc]) in the speech of mothers addressing 28-month-olds in the Bates corpus. This high percentage is remarkable since this construction is used with a total of 39 different verbs in the mother’s speech in the corpus; the figures for three constructions are given below in Table 1.

<table>
<thead>
<tr>
<th>Construction</th>
<th>Mothers</th>
<th>Total Number of Verb Types</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Subj V Obj</td>
<td>39% go (136/353)</td>
<td>39 verbs</td>
</tr>
<tr>
<td>2. Subj V Obj Obl</td>
<td>38% put (99/259)</td>
<td>43 verbs</td>
</tr>
<tr>
<td>3. Subj V Obj Obj2</td>
<td>20% give (11/54)</td>
<td>13 verbs</td>
</tr>
</tbody>
</table>

Table 1. Corpus Study (Goldberg et al., 2004): 15 Mothers’ most frequent verb and number of verb types for 3 constructions in Bates et al. (1988) corpus

Clear motivation exists for speakers to use certain verbs more frequently than others. If we compare for example, go with amble, or put with shelf, it is clear that go and put are more frequent because they apply to a wider range of arguments and therefore are relevant in a wider range of contexts (Bybee, Perkins, & Pagliuca, 1992; Heine, 1993; Zipf, 1935).

It is not claimed that any particular verbs are necessarily the very first verbs uttered. Longitudinal studies have suggested that they might be (Ninio, 1999); but see Campbell and Tomasello (2001) for evidence that they are not always the very first verbs.

When looking beyond argument structure constructions to other sorts of constructions, we find that they too tend to be represented by skewed input, in that utterances tend to be disproportionately represented by instances that involve the same word or small set of words (Brenier & Michaelis, 2004; Cameron-Faulkner, Lieven, & Tomasello2003 Diessel, 2002;

3. Testing further predictions

*Does presenting all skewed instances first facilitate learning?*

The notion that shared concrete similarity should enhance learning because the shared similarity serves to remind learners of previous instances leads to a further prediction about the order of presentation. In particular, presentation of all of the skewed examples first, before other novel verbs are introduced should have an additional facilitative effect. That is, if hearing examples that share some concrete similarity (e.g., the same novel verb) encourages learners to form a category, witnessing those examples early in training should encourage them to form the category earlier. This in turn would lead to easier assimilation of other instances to the category, resulting in a better learned category. That is, after hearing all eight examples that contained the same novel verb, subjects would, by hypothesis, have formed a tentative category that they could use as a standard for the other instances, thereby forming a stronger category representation.

An indication that order of presentation does have an effect on categorization comes from the non-linguistic categorization literature. Elio and Anderson (1984) set up two conditions relevant to the current discussion. In the “centered” condition, subjects were initially trained on closely related instances, with the study sample growing gradually to include more members of the category. In the “representative” condition, subjects were trained on a fully representative sampling from the start. In both conditions, subjects were eventually trained on the full range of instances. Elio and Anderson demonstrated that categories were learned more accurately in the centered condition; the representative condition yielded poorer typicality ratings and accuracy during the test phase on new instances. Elio and Anderson observe, “The superiority of the centered condition over the representative condition suggests that an initial, low-variance sample of the most frequently occurring members may allow the learner to get a ‘fix’ on what will account for most of the category members” (p. 25). They go on to note that “a low-variance sample, in which there is a maximum amount of similarity among items, is particularly conducive to forming strong category generalizations” (p. 28). Similar results were found by Avrahami et al. (1997) who demonstrated that subjects learned categories better when presented with several ideal positive cases followed by borderline cases than if they were presented with sequences that emphasized category boundaries from the start.

To test the hypothesis that the order of presentation affects the learning of an abstract novel construction, we included in a new experiment, two training conditions. The SKEWED FIRST condition heard eight instances of the same novel verb (*moopo*) before hearing the other eight instances involving four other novel verbs presented twice each; the SKEWED RANDOM condition heard the same eight instances of the same novel verb randomly interspersed among the same other eight instances. (The skewed random condition corresponds to the SKEWED condition in Casenhiser & Goldberg, 2005).

*Do instructions to learn the nonsense verbs have an effect on learning an abstract construction?*

In a 2x2 design, half of the skewed first and half of the skewed random conditions were placed in an EXPLICIT INSTRUCTION condition and the other half of subjects were placed in an IMPLICIT INSTRUCTION condition. Those in the former were instructed to try to figure out the meanings of

---

3 The study involved descriptions of people belonging to one of two clubs, with members’ descriptions varying on five 4-valued dimensions.

5 Stimuli in this experiment also consisted of non-linguistic stimuli: variable sized semi-circles with variably oriented radial lines.
the nonsense verbs, while those in the latter were simply told to pay attention to the video clips. It is possible that increased attention to the individual novel verbs during training would lead to worse performance at test, since the test items involve different novel verbs. On the other hand, it is possible that subjects would try to learn the novel verbs in any case, so that varying the instructions would make no difference.

Can learners distinguish the novel construction from novel instances of the transitive construction at test?
Previous work on adult subjects (Goldberg et al., 2004) demonstrated that they could identify new instances of the novel construction at test. However, it is possible that subjects were simply primed to choose scenes of appearance by the training, without attending to any particular aspects of the novel form. In order to determine that subjects did attend to the form of the novel construction, it is necessary to show that they could distinguish the novel construction from some other construction at test, choosing scenes of appearance only for the novel construction and not for the foil construction. Previous work has revealed that children are able to accurately distinguish the novel construction from novel transitives (Casenhiser & Goldberg, 2005); we aimed to replicate that finding in the current study with adults.

Thus, in the test phase of the present experiment, subjects heard six new instances of the novel construction and six instances of the already familiar transitive construction. Both instances of the novel construction and the transitive construction involved new novel verbs, so each utterance was unfamiliar to the learner. Each test item included both a scene of appearance and a semantically transitive scene. Subjects’ task was to choose the scene that matched the description they heard; if they heard the novel construction, the accurate choice was the scene of appearance; if they heard a transitive construction, the accurate choice was the semantically transitive scene.

Is learning simply an effect of priming “appearance”?
In previous work, we used a control condition in which subjects watched the same training film without sound (Casenhiser & Goldberg, 2005). However, it is conceivable that simply hearing the entities in the scene named would, by focusing attention on the relevant entities in the scenes during training, lead to better performance at test, without subjects actually attending to the form of the construction. The inclusion of transitives at test will serve as one indication as to whether the form of the construction is attended to, but it is conceivable that subjects could accurately discriminate the novel construction from the transitive construction purely on the basis of previous knowledge of the transitive construction. If the training on the novel construction is necessary, then training involving other constructions or lexical items should not lead to above chance performance at test.

For this reason, we included in the present experiment two new control conditions, in addition to using the no-sound control. In the TWO-NP condition, subjects heard just the two NPs (the theme and the location) in randomized order during the training (no novel verbs). In the INTRANSITIVE condition, subjects heard instances of the familiar intransitive motion construction used with five nonsense verbs (e.g., *The sun vaks in the sky*) during training. The TWO-NP control condition is useful to determine whether hearing the entities named serves to focus attention sufficiently on scenes of appearance; the intransitive control condition encourages subjects to notice that scenes of appearance are involved insofar as the intransitive motion construction can readily be used to describe just such scenes (e.g., *The sun appears in the sky*).

In all control conditions, the testing (including transitives) is the same as in the training conditions (skewed first and skewed random). If subjects in the training conditions are truly learning the form of the novel construction on the basis of exposure to the construction during training, neither exposure to two noun phrases nor exposure to novel instances of a familiar construction should facilitate the identification of the novel construction at test.
3.1 Method

Participants
One hundred twenty-six English-speaking undergraduates at Princeton University were paid $6 for their participation.

Procedure
A computer program was prepared with a different training section for each condition. Initially, subjects in all conditions went through a forced choice pre-testing procedure with English sentences in order to ensure that they understood the task. The pre-testing clips had two similar scenes next to each other on the screen, one that corresponded with the audio track and one that did not. For example, one of the pre-testing clips had the audio track, *the king and queen are bowing*, and subjects had to pick between one scene in which a king bows at a queen and another scene in which both a king and queen bow. All subjects completed the pretest with at least five out of six correct.

The training film for all conditions contained eight different clips that demonstrated a scene of appearance. The clips were shown twice. In the experimental training conditions, the clips were paired with an audio sequence that described the scene using the novel word order we created. The subjects heard the sentence in the present tense as the appearance was occurring in the scene and in the regular form of past tense after the appearance had occurred. For example, *the bug the table moopo-s . . . the bug the table moopo-ed* was played as a bug appeared on a table.

The skewed first condition witnessed eight sentences containing the nonsense verb *moopo* at the beginning of the training, followed by the other eight sentences containing four other nonsense verbs. The skewed random order condition witnessed the 16 clips in a randomized order: the eight sentences containing the nonsense verb *moopo* were interspersed with the other novel verbs in a randomized order.

Within each of these conditions, half the subjects received implicit instructions to watch the videos, and the other half received the explicit instructions to try to learn the novel verbs in the sentences. In all three control conditions, the clips were shown in a random order, and subjects were told to pay attention to the films. See table 1 for a complete list of the sentences used in the experimental conditions, table 2 for a complete list of the sentences used in the two NP control and table 3 for a complete list of the sentences used in the intransitive control.

<table>
<thead>
<tr>
<th>Training scene displayed on video</th>
<th>Corresponding audio track</th>
</tr>
</thead>
<tbody>
<tr>
<td>A rabbit appears on a hat</td>
<td>The rabbit the hat moopoed</td>
</tr>
<tr>
<td>A monster crawls out of a cloth</td>
<td>The monster the cloth keeboed</td>
</tr>
<tr>
<td>A frog drops onto a box</td>
<td>The frog the box moopoed</td>
</tr>
<tr>
<td>A king drops into a chair</td>
<td>The king the chair vakoed</td>
</tr>
<tr>
<td>A sun rises into a sky</td>
<td>The sun the sky fegoed</td>
</tr>
<tr>
<td>A queen rolls into a stage</td>
<td>The queen the stage sutoed</td>
</tr>
<tr>
<td>A bug appears on a table</td>
<td>The bug the table moopoed</td>
</tr>
<tr>
<td>A ball rolls into a room</td>
<td>The ball the room moopoed</td>
</tr>
</tbody>
</table>
Table 1. Training stimuli for the experimental conditions. In the fixed order conditions, all the moopo clips were viewed first, followed by the rest of the clips in randomized order; in the random condition, the clips were viewed in randomized orders. All 8 clips were viewed twice in every condition so that each subject witnessed 16 clips.

<table>
<thead>
<tr>
<th>Scene displayed on video</th>
<th>Corresponding audio track</th>
</tr>
</thead>
<tbody>
<tr>
<td>A rabbit appears on a hat</td>
<td>The rabbit the hat</td>
</tr>
<tr>
<td>A monster crawls out of a cloth</td>
<td>The cloth the monster</td>
</tr>
<tr>
<td>A frog drops onto a box</td>
<td>The box the frog</td>
</tr>
<tr>
<td>A king drops into a chair</td>
<td>The chair the king</td>
</tr>
<tr>
<td>A sun rises into a sky</td>
<td>The sun the sky</td>
</tr>
<tr>
<td>A queen rolls into a stage</td>
<td>The queen the stage</td>
</tr>
<tr>
<td>A bug appears on a table</td>
<td>The table the bug</td>
</tr>
<tr>
<td>A ball rolls into a room</td>
<td>The ball the room</td>
</tr>
</tbody>
</table>

Table 2. Training stimuli for the noun phrase control condition. The clips were viewed in a random order with implicit instructions. All 8 clips were viewed twice in every condition. Theme and locative NPs were heard in counter balanced order.

<table>
<thead>
<tr>
<th>Scene displayed on video</th>
<th>Corresponding audio track</th>
</tr>
</thead>
<tbody>
<tr>
<td>A rabbit appears on a hat</td>
<td>The rabbit mooped in the hat</td>
</tr>
<tr>
<td>A monster crawls out of a cloth</td>
<td>The monster keebed out of the cloth</td>
</tr>
<tr>
<td>A frog drops onto a box</td>
<td>The frog mooped onto the box</td>
</tr>
<tr>
<td>A king drops into a chair</td>
<td>The king vaked in the chair</td>
</tr>
<tr>
<td>A sun rises into a sky</td>
<td>The sun feged in the sky</td>
</tr>
<tr>
<td>A queen rolls into a stage</td>
<td>The queen suted onto the stage</td>
</tr>
<tr>
<td>A bug appears on a table</td>
<td>The bug mooped on the table</td>
</tr>
<tr>
<td>A ball rolls into a room</td>
<td>The ball mooped into the room</td>
</tr>
</tbody>
</table>

Table 3. Training stimuli for the intransitive control condition. The clips were viewed in a random order. All 8 clips were viewed twice in every condition.

Upon completion of the training phase, subjects were given a forced-choice task. They were asked to identify which of two scenes corresponded to the audio track sentence. Each test item paired a new scene of appearance with a similar scene that did not involve appearance. The audio tracks for the testing phase contained new novel verbs not used during training. To select a scene, subjects clicked the mouse on a box below the scene, and then clicked on a box with the word “accept” to finalize their choice. Subjects performed six forced choice tasks on the novel word order and then performed six forced choice tasks on the English transitive construction. Responses were automatically entered into an output file each time a subject clicked “accept”.

Subjects were debriefed, paid, and thanked for their participation upon completion of the experimental phase.

Results

A two-way between groups ANOVA with Instruction (Implicit versus Explicit) and Order (Fixed versus Random) revealed that there were no significant differences for Instruction (F(1, 68) = 1.05, p = .309) and no significant interaction of Instruction x Order (F(1, 68) = .148, p = .702). There was, however, a significant effect for Order of presentation (F(1, 68) = 4.199, p = .044). The Implicit and Explicit Instruction groups were accordingly collapsed for the remainder of the analyses.
As represented in Figure 1, an ANOVA conducted on the five resulting conditions (two training conditions and three controls) revealed a significant main effect of condition: $F(4, 121) = 4.426, p = 0.002$. Planned comparisons revealed that subjects in the skewed first condition performed significantly above those in the skewed random order $p = 0.026$. Subjects in the skewed first condition also performed better than each control condition: comparison with no sound $p = 0.001$; comparison with two noun phrase control $p = 0.015$; comparison with intransitive control $p = .003$. Subjects in the random condition performed marginally better than the those in the no-sound control ($p = .08$).

When comparing the results found in the current study to Casenhiser and Goldberg (2005), the skewed random order condition (Casenhiser and Goldberg’s skewed frequency condition) had a less robust effect, as it was only marginally higher than the no-sound control. Although this is not an exact replication of previous research, Goldberg et al.’s (2004) experiment with adults used slightly different films, and their films contained more scenes of pure appearance without any manner. For example, in one scene a dot appears on a king’s nose, in another, a bug appears on a table. These films also contained an identical “bling” sound as the entity appeared, providing an additional concrete similarity. These slight differences could explain why adult subjects in Goldberg et al. were better at test for the scenes of appearance than subjects in the current study.

A previous experiment on six year olds used identical materials to the current study and found a significant difference between their skewed frequency condition (=current skewed random) and the no-sound control (=our no-sound control) (Casenhiser & Goldberg, 2005). It is possible that children perform better than adults on these tasks; however it is premature to draw this conclusion because a statistical comparison of the children in the previous study and the adults in the current study is not entirely equitable given slight differences in the designs. Additional studies that directly compare adults and children will reveal whether there are any age effects.

Comparisons to chance revealed that subjects in the two experimental conditions performed above chance: skewed random order ($t(35)=2.53, p =.02$), skewed first order ($t(35)=5.45, p =.001$). Subjects in the three control conditions did not perform differently from chance: no sound control: ($t(17)= -.48, p =.64$), the two noun phrase ($t(17)=1.13, p=.27$), and the intransitive ($t(17)=.34, p= .73$).
Turning to performance on the transitive construction (see figure 2), no differences between groups were anticipated across conditions, since, even though the test items contained novel verbs, all subjects were already familiar with the abstract transitive construction. As expected, a two way between subjects ANOVA with Instruction (Implicit versus Explicit) and Order of presentation of stimuli (Fixed versus Random) as the between subjects factors revealed no significant differences for Instruction (F(1, 68) = .023, p = .879), Order (F(1, 68) = 1.879, p = .175) or the Instruction x Order interaction (F(1, 68) = 1.879, p = .175). All conditions performed above chance at identifying semantically transitive events when they heard the novel transitive utterances (Fixed: t(35) = 8.270, p = .001; Random: t(35) = 4.340, p = .001; No-Sound Control t(17) = 10.000, p = .001; Control Intransitive t(17) = 6.553, p = .001; 2NP Control t(17), p = .005).
Figure 2. There were no significant differences in performance on the transitive sentences for any condition.

Discussion
There were four main goals in the present study. The first was to examine the effects of varying the order in which the same linguistic stimuli are shown. The present work demonstrates that subjects who heard eight examples of scenes of appearance with the novel verb *moopo* first, followed by another eight instances involving four other novel verbs, were more accurate at identifying new instances of the novel construction at test than subjects who heard the eight examples with the novel verb *moopo* interspersed randomly among the rest of the instances during training. That is, presenting the skewed input at the beginning of training boosts the accuracy at test over presenting the skewed input randomly throughout training. This finding is consistent with the idea that subjects are forming a category of the construction: they are facilitated when initial instances remind them of immediately preceding instances.

A second objective was to determine whether adults would be able to distinguish the novel construction from novel instances of the transitive construction at test. We found that they were. Subjects demonstrated that they were able to distinguish new instances of the novel construction, involving new novel verbs and new scenes of appearance, from new instances of the transitive construction, also involving new novel verbs, replicating previous work with six year old children (Casenhiser & Goldberg, 2005) and four year old children (Goldberg, in preparation). That is, performance on both the novel construction and on the transitive construction was above chance in the training conditions. In order to distinguish the novel construction from the transitive construction in training conditions, learners had to recognize something about both the form and meaning of the novel construction.

As expected, performance on the transitive construction (only) was above chance in the control condition, since subjects already knew the transitive construction and were not exposed to the novel construction.
A third objective of the study was to introduce two new control conditions to ensure that subjects were not simply primed to choose scenes of appearance after having to attended to scenes of appearance during training. Subjects in none of the control conditions performed significantly differently from chance (NoSound control t(17) = -.483, p = .636; Intransitive Control t(17) = .345, p = .734; 2NP Control t(17) = 1.13, p = .274).

A final goal of the study was to examine the effect of instructions on linguistic categorization. We considered the possibility that asking subjects to learn the meaning of the nonsense verbs would direct attention away from the generalization across nonsense verbs, leading to worse overall performance. Varying the instructions did not create a significant difference between groups. Verbal accounts of subjects during debriefing suggest that subjects were trying to learn the nonsense verbs even in the implicit instruction condition where they were instructed simply to “pay attention to the film clips.” This may well be why explicit instructions to learn verb meaning had no effect.

Conclusion

One factor that enables language to be used creatively is the fact that formal patterns correlate strongly with the meanings of the utterances in which they appear. Of particular interest is the identification of facilitative factors in learning these correlations (i.e., constructions).

Previous experiments have demonstrated that individuals are able to learn to differentiate the form and meaning of a novel construction from that of a known construction after short exposure to construction exemplars (Casenhiser & Goldberg, 2005; Goldberg et al., 2004). We here interpret this data in the light of non-linguistic research which has demonstrated that shared concrete similarity among instances of a novel category facilitates learning of the category (Markman & Gentner, 1993; Kotovsky & Gentner, 1996). We also present a new experiment that explores the implications of such a model in terms of the order of presentation of instances. If hearing examples that share some concrete similarity encourages learners to form a category, witnessing those examples early in training should encourage them to form the category earlier. This in turn would lead to easier assimilation of other instances to the category. Results reported here bear out this prediction.

The current experiment fits into a growing body of research that aims to determine the extent to which linguistic categorization involves domain general processes. We have found that children are able to generalize beyond their input with minimal exposure to form a linguistic category, and that they are better able to do so when the input is structured in such a way that early examples share concrete similarity. This finding is consistent with the idea that generalizations in language are akin to non-linguistic generalizations. At the same time, it is clear that we cannot yet conclude that no domain-specific abilities are required for language learning. What we hope to have demonstrated, however, is that lessons learned in other domains of cognitive psychology may well provide insight into our impressive ability to learn language.

References


