Capacities underlying word learning

Paul Bloom and Lori Markson

Two- and three-year-old children have poor motor control and bad manners; they are unrefined artists and inept dancers. However, they are strikingly good at learning the meanings of words. Current controversy focuses on the relative importance of different capacities in this learning process including principles of association, low-level attentional mechanisms, special word learning constraints, syntactic cues and theory of mind. We argue that children succeed at word learning because they possess certain conceptual biases about the external world, the ability to infer the referential intentions of others and an appreciation of syntactic cues to word meaning. Support for this view comes from studies exploring the phenomena of fast mapping, the whole object bias, the acquisition of nouns for entities belonging to different ontological kinds and the effect of lexical contrast. Word learning is not the result of a general associative learning process, nor does it involve specialized constraints. The ability to learn the meanings of words depends on a number of capacities, some of which are specific to language and unique to humans, others of which are potentially shared with other species.
Review

Box 1. The rate of word learning

It is often said that children start off learning words very slowly, about one or two a week, and then, after learning somewhat between 20 and 100 words, there is a sudden acceleration in the rate of word learning—sometimes called a vocabulary spurt, naming explosion, or word burst—that propels children to learn words at the rate of five, ten or even twenty new words a day. But vocabulary growth is actually less dramatic than this. While some children might show a sudden increase in the rate at which they learn words10, many do not, exhibiting instead a series of small bursts, a slow monotonic rise or a smoothly increasing exponential function11. Moreover, a vocabulary spurt will not be observed to take place when the learning rate is ten words a week, not a day.

To see how the rate of word learning changes, consider the estimates (see Table) from Fenson et al., based on parental reports of the vocabularies of children from 12 to 30 months, and from Anglin, based on comprehension studies with six-, eight- and ten-year-old children. (Anglin included only those words whose meanings could not be worked out using ‘problem solving’ strategies and hence must be learned.) Estimating vocabulary size is tricky for several reasons, so these numbers should be taken only as very rough estimates.

<table>
<thead>
<tr>
<th>Age</th>
<th>Average number of words per day</th>
</tr>
</thead>
<tbody>
<tr>
<td>12–16 months</td>
<td>0.3</td>
</tr>
<tr>
<td>16–24 months</td>
<td>0.8</td>
</tr>
<tr>
<td>23–30 months</td>
<td>1.6</td>
</tr>
<tr>
<td>30 months–4 years</td>
<td>3.6</td>
</tr>
<tr>
<td>6–8 years</td>
<td>6.6</td>
</tr>
<tr>
<td>8–10 years</td>
<td>12.1</td>
</tr>
</tbody>
</table>

Why do children gradually improve at word learning? There are several explanations, including the maturation of memory and attentional capacities, a growing sensitivity to different cues to a word’s meaning (such as its syntactic category and the discourse context in which it is used) and an increasing awareness of different entities that can be named. Another factor is access to new words; in particular, literate exposure to many more words than they would encounter through social interaction or watching television, and it is likely that the grammatical vocabularies of some English speakers (well over 100 000 words) could not occur without the ability to read.

Most adults do not learn several new words a day. This might be because we are not as good at word learning as we once were, but a simpler possibility is that we already know most of the words that our immediate environment has to offer. Unless we learn a new language, our only opportunities for word learning are proper names, archaic or technical terms, or new words that enter the language, such as ‘internet’, ‘dissing’ and ‘karaoke’.

References

morphology and syntax, in which children are clearly superior14. Children and adults were equally good, however, at remembering the arbitrary linguistically presented fact that they had learned about a novel object – that it was given to the experimenter by her uncle. A further experiment found that fast mapping has its limits. When taught the location of a sticker, "This goes here…" as a small sticker was placed on to one of the novel objects, both children and adults did relatively poorly when asked a month later to place a small sticker on the appropriate object (see Fig. 1). The question of what can and cannot be fast mapped is yet unresolved, but it is now clear that fast mapping is not restricted to language learning.

Object names
When confronted with the unfamiliar word ‘koba’ in the above study, children immediately understood it as a name for one of the objects in front of them. This leads to a rather vexing puzzle. Suppose a child hears a new word and determines that it describes a certain object in the world, for instance, a rabbit. Most likely, children learn this by attending to the referential focus of the speaker, as indicated by cues such as direction of gaze3,6. But the problem now arises that there are an infinite number of logically possible meanings for the word. It could refer to the color of the rabbit, its shape, its surface, the tail, the ears, the rabbit and the ground it is standing on, in motion, even ‘undetached rabbit parts’18. Children do not entertain such possibilities, however. In this situation, they will immediately take the word as naming the whole rabbit. More generally, there is a wealth of evidence showing that children and adults tend to interpret new words as referring to whole objects, not to parts of objects, properties of objects or the stuff that objects are made of15,19.

How is this bias best explained? One theory is that children possess a special constraint that guides them to view new words as names for objects15,20. An alternative is that this bias towards objects is the result of a more general fact about how people reason about the world. We are prone to think about the world in terms of whole objects and hence, when searching for the meaning of a word, are driven to favor the object interpretation15,21 (see Box 2).

The claim here is not merely that children parse the world into objects and, therefore, tend to take new words as object names. After all, children are also sensitive to motion and color but show no bias to interpret new words as naming motion or color. The proposal is instead that objects are highly salient, both linguistically and non-linguistically. In support of this, note that the very same focus on objects shows up in domains other than word learning. When preschoolers are shown an array of objects and asked to count, they show a strong tendency to count the objects, even when searching for the meaning of a word, are driven to favor the object interpretation15,21 (see Box 2).

**Box 2. What is an object?**

What precisely do we mean when we say that children have an object bias? We have an intuitive sense of what we mean by the term, so that a rabbit is an object, but the rabbit’s foot is not (at least, not if it is attached to the rabbit), nor is the rabbit and the tree that it is near to. But what are the precise criteria that guide children to view some things as objects and others as not?

One promising theory has been elaborated by Elizabeth Spelke on the basis of infant research20. The most important criterion is that objects follow the principle of cohesion; an object is a connected and bounded region of matter that maintains its connectedness and boundaries when in motion. If young infants see a portion of matter obey the principle of cohesion, they treat it as an object and expect other object principles to apply. (Those other principles state that objects are solid and do not pass through each other, they follow continuous paths through space, and they move if and only if they touch.)

Of course, it is not necessary actually to observe cohesion and bounded movement before concluding that something is an object; it is enough to infer that there could be such movement. Hence, we can parse stationary scenes into distinct objects when the gaps between entities imply that they will move independently, when the scene contains entities that we know from previous experience exist as separate objects (as with a rabbit that is in contact with a tree) or when Gomtsi cues, such as good continuity and sameness of color and texture, suggest that different entities have the potential for independent movement (as when we see a shiny red object resting on a flat green surface).

Pre-linguistic infants are sensitive to all of these cues, but only sometime after they are able to parse objects on the basis of motion.

**References**


they have little difficulty counting individuals such as sounds. When adults were shown three groups on a computer screen, each moving as a single unit, tracing paths across the screen, avoiding each other, and so on, they construed that the groups were individuals and interpreted ‘fendle’ as a collective noun. To test whether this is due to experience with real-world collections, like flocks of birds, Karen Wynn and I tested five-month-old infants in a similar experiment. We showed half of the infants two collections of three objects each and half of them four collections of three objects each. Each collection traced a vertical path up and down on a computer screen and each object moved independently within its collection, while also following the trajectory traced by the collection as a whole. Once infants were habituated to this display, they were presented alternately with two collections of four objects, and four collections of two objects, each moving horizontally back and forth on the screen. Infants looked reliably longer at the new number of collections, showing that they were sensitive to the number of collections of entities presented to them.

What do these cases have in common? It is interesting that they are all ways in which a collection is like an object – both are described with singular count nouns, treated as a single entity by others and move as bounded units. This is consistent with the view that principles of object parsing may extend to other non-object domains. Alternatively, humans may possess some set of more general principles that parse the world into categories, including not only objects and collections, but also shadows, holes, parts, sounds and actions. Further research into how children and adults learn words for non-object individuals will bear on this issue.

References

during explicit instruction. In one study, the experimenter carefully arranged three groups of objects in front of the subjects, with a picture-frame around each group, giving the impression that the groups were independent artistic creations. If the frames are then removed, and the display is described as ‘These are fendles’ children will tend to treat ‘fendle’ as a collective noun, suggesting that they can continue the groups as individual solely because they infer that they are thought of as individuals in the mind of the experimenter.

A third cue is movement. When adults were shown these groups on a computer screen, each moving as a single unit, tracing paths across the screen, avoiding each other, and so on, they construed that the groups were individuals and interpreted ‘fendle’ as a collective noun. To test whether this is due to experience with real-world collections, like flocks of birds, Karen Wynn and I tested five-month-old infants in a similar experiment. We showed half of the infants two collections of three objects each and half of them four collections of three objects each. Each collection traced a vertical path up and down on a computer screen and each object moved independently within its collection, while also following the trajectory traced by the collection as a whole. Once infants were habituated to this display, they were presented alternately with two collections of four objects, and four collections of two objects, each moving horizontally back and forth on the screen. Infants looked reliably longer at the new number of collections, showing that they were sensitive to the number of collections of entities presented to them.

What do these cases have in common? It is interesting that they are all ways in which a collection is like an object – both are described with singular count nouns, treated as a single entity by others and move as bounded units. This is consistent with the view that principles of object parsing may extend to other non-object domains. Alternatively, humans may possess some set of more general principles that parse the world into categories, including not only objects and collections, but also shadows, holes, parts, sounds and actions. Further research into how children and adults learn words for non-object individuals will bear on this issue.

References

certainly not, however, that children can only count objects; when there are no objects in sight, they have little difficulty counting individuals such as sounds. These considerations put the object bias in a different perspective. It may be that children have a range of ontological categories at their disposal. Objects are just very salient; it is hard for children not to focus on them. This explains why young children find it much easier to learn a name for a non-solid substance such as ‘water’, than to learn a solid substance name like ‘wood’. Children can learn ‘water’ without the distraction of a solid object, while learning ‘wood’ requires that children actively focus on a bounded object but think of it, not as an object, but as a portion of solid stuff. The acquisition of collective nouns presents an interesting intermediate case; these are words such as ‘family’ and ‘flock’ which refer not to one object, but to individuals that are composed of many objects (see Box 5). Common nouns, proper names and other ontological categories

Once children know that a word refers to a whole object, they must then determine whether it refers to a kind, as with a common noun such as ‘dog’, or to a specific individual, as with a proper name like ‘Fido’. One important factor is the type of entity: two-year-old children know that words...
Box 4. Syntactic cues to word meaning

The fact that the syntactic category of a word can provide cues to its meaning is not a lucky accident; it follows naturally from the way that language works. Rules of syntax allow people to combine words to express a potential infinity of sentence meanings, and this combinatorial power rests in part on the correspondence that exists between syntactic categories and conceptual categories. To take a very simple example, nouns often refer to things and verbs often refer to actions, so one can combine the two to form a sentence that states that certain things tend to perform certain actions, as in: ‘Dogs bark’.

Over the last decade, there have been many demonstrations of the role that syntactic knowledge has in word learning. Summing up several studies, young children can use a number of cues to word meaning (see Table 1).

While there is little doubt that a sensitivity to syntactic cues plays some role in word learning, there remains substantial debate about how children become sensitive to these cues, how early in development they apply and how important they are for the acquisition of certain word meanings.

References

- Pinker, S. (1994) How could a child use verb syntax to learn verb semantics? Lingua 92, 277–410

<table>
<thead>
<tr>
<th>Syntactic cue</th>
<th>Usual type of meaning</th>
<th>Examples</th>
<th>Refs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Noun is a noun of a category</td>
<td>Individual member of a category</td>
<td>cat, forest</td>
<td>a,b</td>
</tr>
<tr>
<td>These are feps</td>
<td>Multiple members of a category</td>
<td>cats, forests</td>
<td>a,c</td>
</tr>
<tr>
<td>This is Fep</td>
<td>Specific individual</td>
<td>Fido, John</td>
<td>b</td>
</tr>
<tr>
<td>This is a non-individuated thing</td>
<td>Action with one participant</td>
<td>sleeps, stands</td>
<td>d,e</td>
</tr>
<tr>
<td>John feps</td>
<td>Action with two participants</td>
<td>hits, kisses</td>
<td>e</td>
</tr>
<tr>
<td>This is a fep</td>
<td>Spatial relationship</td>
<td>on, near</td>
<td>g</td>
</tr>
<tr>
<td>The dog is fep</td>
<td>Comparative</td>
<td>Quantifier</td>
<td>same, fine</td>
</tr>
</tbody>
</table>

A sensitivity to syntactic cues plays some role in word learning. The original study was carried out by Roger Brown, who showed preschoolers a picture of a strange action being performed on a novel substance with an unfamiliar object. One group of children was told: ‘Do you know what a a? In this picture, you can see a a’ (count noun syntax); a second group was told: ‘Have you seen any a? In this picture, you can see a a’ (mass noun syntax); and a third group was told: ‘Have you seen a a? In this picture, you can see a a’ (verb syntax). The preschoolers tended to construe the count noun as referring to the object, the mass noun as referring to the substance, and the verb as referring to the action. Subsequent research has found that syntactic cues can guide young children to learn words belonging to a range of different ontological categories (see Box 4).

Lexical contrast

Eventually, children do learn words that describe objects but which are not object names. One non-syntactic cue that could help children learn the meanings of such words involves lexical contrast. When children already have a name for an object, they tend to assume that another word, presented in the presence of that object, has a different meaning. When three-year-old children are shown a familiar object, such as a cup, and an unfamiliar object, such as a pair of tongs, and asked to: ‘Show me the dax’, they tend to select the unfamiliar object as the referent of the novel label.

It seems that children hold a specifically linguistic assumption that describing individuals that have an intrinsic specialness, such as people and pets, are more likely to be proper names than words that describe things that are seen as interchangeable members of a kind, such as trucks and bees. Another factor is syntax. For example, when exposed to a word that describes a doll, children younger than two can use the syntax of the word to work out its meaning. If a word is used in a context such as: ‘This is a…’, they take it as naming the individual, as in ‘Stella’; if it is used in a context such as: ‘This is a…’, they take it as naming the kind, as in ‘Dolls back’. A sensitivity to syntactic cues plays a more general role in word learning. The original study was carried out by Roger Brown, who showed preschoolers a picture of a strange action being performed on a novel substance with an unfamiliar object. One group of children was told: ‘Do you know what a a? In this picture, you can see a a’ (count noun syntax); a second group was told: ‘Have you seen any a? In this picture, you can see a a’ (mass noun syntax); and a third group was told: ‘Have you seen a a? In this picture, you can see a a’ (verb syntax). The preschoolers tended to construe the count noun as referring to the object, the mass noun as referring to the substance, and the verb as referring to the action. Subsequent research has found that syntactic cues can guide young children to learn words belonging to a range of different ontological categories (see Box 4).
Review

Bloom and Markson – Word learning

72

• How does word learning by young children differ from word learning by older children and adults?

• What can developmental disorders such as specific language impairment and autism tell us about the relationship between word learning and other aspects of language development?

Outstanding questions

Table 1. The scope of the lexical contrast assumption

<table>
<thead>
<tr>
<th>Condition</th>
<th>What children were told and asked</th>
<th>Which object they tended to choose</th>
</tr>
</thead>
<tbody>
<tr>
<td>Names</td>
<td>This is a bem (Object A)</td>
<td>Object B</td>
</tr>
<tr>
<td></td>
<td>Can you show me a job?</td>
<td></td>
</tr>
<tr>
<td>Facts</td>
<td>My sister gave this to me (Object A)</td>
<td></td>
</tr>
<tr>
<td>one speaker</td>
<td>Can you show me the one that dogs like to play with?</td>
<td>Object B</td>
</tr>
<tr>
<td>Facts</td>
<td>My sister gave this to me (Object A)</td>
<td></td>
</tr>
<tr>
<td>two speakers</td>
<td>Can you show me the one that dogs like to play with?</td>
<td>No preference</td>
</tr>
</tbody>
</table>

*Three-year-olds were presented with two unfamiliar objects in all conditions.

Table 1. The scope of the lexical contrast assumption

- Children have a general bias to focus on objects but, nonetheless, they quickly learn words that refer to the attributes or actions of objects.
- Is the ability to fast map present in non-linguistic creatures, such as young infants and non-human primates?
- What cues other than syntax and lexical contrast can help them override this object bias?
- Is the acquisition of new words from overheard speech as effective as word learning when young children are active participants in the communicative interaction?
- Even 15-month-old infants tend not to accept two names for the same object. Is it plausible that this quite precocious understanding is based on an understanding of the communicative intentions of others?
- How does word learning by young children differ from word learning by older children and adults?
- What can developmental disorders such as specific language impairment and autism tell us about the relationship between word learning and other aspects of language development?

- The children tended to infer that a second, different name (“Can you show me a job?”) referred to the other, unlabeled object, replicating previous studies. In another condition, a different group of children were told a novel fact about one of the objects (“My sister gave this to me”) and were then asked to select the referent of a different fact (“Can you show me the one that dogs like to play with?”). Again, children in this condition tended to choose the other object, the one that they had not been told the fact about. Presumably, this is because children reasoned that if the experimenter had intended to describe the first object, she would have referred to it by stating the original fact – she would not have introduced a different fact.

- This account predicts that children should be less inclined to produce such a response in a two-speaker scenario, where the second speaker does not share mutual knowledge with the child. That is, if one speaker tells the child: “My sister gave this to me” about one object, and then a different speaker, who is new to the discourse context, enters the room and asks: “Can you show me the one that dogs like to play with?”, children should now choose each of the objects with equal frequency. As predicted, three-year-old children selected among the two objects randomly. This supports the hypothesis that children’s interpretation of the referent of a speaker’s utterance is based upon their understanding of how people communicate with one another; it is not the product of a special lexical constraint.

Summary and implications

The studies reviewed above support the view that young children’s remarkable ability to learn words emerges from more general cognitive capacities: intentional, conceptual and syntactic. Such capacities explain phenomena such as fast mapping, the whole object bias, the acquisition of names for entities belonging to different ontological kinds and assumptions about lexical contrast.

Our claim is that while other aspects of language acquisition, such as the learning of phonology, morphology and syntax, involve dedicated cognitive modules, word learning does not. The ability to learn the meanings of words emerges instead from a host of other capacities, some of which are shared by other species (such as object parsing), some of which are parasitic on language learning and hence unique to humans (such as the use of syntactic cues to word meaning), and others for which human uniqueness is an open question (such as the ability to fast map).

We suspect that the central capacity underlying word learning is theory of mind. Here, this was argued to underlie children’s inferences about lexical contrast, and elsewhere it has been proposed to account for various other phenomena in word learning, including how children work out what a new object name refers to. Word learning can be accomplished, at least to a limited extent, without the ability to fast map and without sensitivity to syntactic cues. But an understanding of the notion of referential intent may be essential. Interestingly, it is exactly this capacity that is lacking in non-human primates. The ability to contemplate the thoughts of others may be the engine that drives word learning, and the emergence of this capacity may be a
Acknowledgements
Preparation of this review was supported by a grant from the Spencer Foundation. We are grateful to Karen Wynn, Lila Gleitman and an anonymous reviewer for helpful comments on an earlier draft.

References
7 Gleitman, L.R. (1990) The structural sources of word meaning. Languag. Acquisition 1, 3–51
18 Quinn, W.C. (1940) Word and Object. MIT Press

Acknowledgements
Preparation of this review was supported by a grant from the Spencer Foundation. We are grateful to Karen Wynn, Lila Gleitman and an anonymous reviewer for helpful comments on an earlier draft.

References
7 Gleitman, L.R. (1990) The structural sources of word meaning. Languag. Acquisition 1, 3–51