Mechanisms of Word Learning

Gil Diesendruck

Words are conventional arbitrary symbols used to communicate concepts. By school age, children have thousands of these. What are the mechanisms that allow children to accomplish this feat?

In answering this question, I will start with an introduction of the phenomenon, arguably emphasizing the most puzzling question word-learning theorists attempt to explain: What allows children to limit the possible meanings of words? I will describe six mechanisms proposed to answer this question, and discuss empirical challenges facing them. My goal is to define the narrower fields in which particular theoretical debates are being contested, and describe some of the most recent discoveries about word learning that any general theory needs to explain. Reflecting a bias in the literature, I will focus primarily on the learning of count nouns.

I will characterize the proposals as differing on two broad dimensions. The main one is the degree of specificity of the mechanism, that is, whether it is specific to the acquisition of words, or instead applicable to a variety of cognitive problems. A secondary dimension has to do with the source of the constraints, specifically, whether the constraints are endogenous or exogenous to the child. Figure 13.1 presents the placement of the six mechanisms on this two-dimensional grid.

The review of the mechanisms and challenges will lead to the conclusion that no single mechanism suffices to account for word learning, and that instead, the field needs approaches that critically integrate multiple mechanisms. Two such approaches will then be discussed. I will end with some general conclusions about where the field stands theoretically, and what are the empirical directions towards which the field should move.
Word learning is fast

It is commonly assumed that children start producing words by their first birthday. By 16 to 18 months of age, children are believed to have an average vocabulary size of 50 words. Most theorists of lexical development believe that around this time, children go through a “vocabulary spurt,” acquiring words at a much faster pace (Dromi, 1987). Recently, Bloom (2004) challenged this notion of a vocabulary spurt, arguing instead that lexical acquisition is a continuous process with a relatively stable rate of change. This debate notwithstanding, all word-learning researchers agree in their amazement regarding the move from the 50 words average vocabulary children have at 18 months of age to the estimated 10,000 words they have by age 6 (Anglin, 1993).

Part of the explanation for how children manage to acquire so many words is that they have good learning and memory capacities. Children can learn a new word associated with an object after a few exposures to the word, and can remember the word for a long period of time – a phenomenon dubbed “fast mapping” (Carey & Bartlett, 1978). Recent studies have further defined this phenomenon, showing that even 13-month-olds are capable of fast mapping (Woodward, Markman, & Firzsimmons, 1994), and 3-year-olds are just as good as adults in remembering such mappings (Markson & Bloom, 1997).
Word learning is universal

Prototypical North American middle-class children are to some extent directly bombarded with linguistic information. Parents tend to talk to their prelinguistic infants frequently, and in a particularly simplified, high frequency, and accentuated manner – often called motherese or baby-talk. Once children start talking, parents often engage them in language games, especially games that invite children to extend or express their emerging knowledge of the language.

The linguistic input children in other cultures receive differs from this prototypical North American middle-class speech. Specifically, cultures differ in the extent, manner, and emphasis of adults’ direct talk to children (Gathercole & Hoff, this volume; Hoff, 2006). As I shall review later, some of these differences in input do correlate with specific differences in children’s eventual linguistic knowledge. Nonetheless, the important conclusion for the present purpose is that despite these differences, children in all cultures studied acquire language at around the same ages, and with relatively similar efficacy.

Word learning is inductive

A popular belief is that word learning is one of the easiest aspects of language acquisition, because, presumably, all that children have to do is imitate adults. This belief implies that children are exposed to words in communicatively optimal contexts, in which an adult and the child are jointly focused on the same object or event, and the adult repeatedly utters the new word associated with the object or event. Fast mapping would then do the rest of the job.

The first major problem with this assumption is that most of the words children learn they encounter in much less optimal communicative contexts than the one just described. This is the case not only in non-Western cultures but also in typical Western middle-class families (Harris, Jones, & Grant, 1983). For instance, it is often the case that when children hear a word, its intended referent is not even visually available.

The second major problem with the popular notion is that even if the typical communicative context in which children encountered new words were optimal in the sense described earlier, there still would not be enough information in the input to allow a definite determination of the meaning of a word. This problem, first noted by Quine (1960), has to do with the fact that any word can have an infinite number of meanings (Poulin-Dubois & Graham, this volume). For instance, upon hearing a bottle of milk being labeled “bottle,” what in the situation can lead the child to rule out the color of the object, its material, its contents, the event itself, and a myriad of other possible concepts, as the referents of the word “bottle”? What would lead the child to necessarily infer that “bottle” refers to that specific kind of object? How does the child even establish what “specific kind of object” that thing is?

What these two problems highlight is that the information available to children in typical communicative interactions underdetermines the possible meanings of words. Thus, children have to do a lot of inductive work to figure out what it is that they are
supposed to imitate in adults’ speech. In what follows, I will describe some of the mechanisms arguably used by children in this inductive process.

Mechanisms of Word Learning

Input

While differences across cultures in children’s exposure to language do not seem to have a substantial impact in whether or not children acquire language, they do affect how and what exactly children acquire. One relatively robust finding in this regard is that across cultures, parents from higher socio-economic standing talk more to their children, and children from higher socio-economic standing have larger vocabularies (Gathercole & Hoff, this volume; Hoff, 2006). Longitudinal studies suggest that it is indeed properties of maternal speech to children – such as number of words – that most strongly predict children’s vocabulary size. In fact, it has been found that the frequency of specific words in maternal speech to children correlates with the frequency of these words in children’s vocabulary (Huttenlocher, Haight, Bryk, Seltzer, & Lyons, 1991).

Arguably, the most investigated issue regarding links between input and child language has to do with the composition of children’s vocabularies. Gentner (1982) found that the proportion of nouns was higher than the proportion of any other word class (e.g., verbs) in the vocabulary of children speakers of various languages. Based on this universal similarity, Gentner argued that this “noun bias” is a result of the relative conceptual simplicity of nouns over verbs. Since then, Gentner’s original findings and conclusions have been tested and challenged (Poulin-Dubois & Graham, this volume). On the one hand, countering Gentner’s conclusion, studies of the vocabulary composition of children speaking Korean and Mandarin Chinese have not found a noun bias. On the other hand, consistent with Gentner’s claim, investigations of the vocabulary composition of children speaking various Western languages have confirmed a prevalence of nouns. This debate notwithstanding, the studies on the noun bias reinforce the notion that the linguistic input children receive can explain certain individual differences in children’s vocabulary acquisition.

The remaining question on this matter is whether linguistic input can constrain how children infer word meanings. The argument underlying this possibility is that by being cooperative communicators, parents simplify the induction problem by providing children with numerous cues about the meanings of words, and the relationship between words (Clark & Wong, 2002; Nelson, 1988). One common labeling pattern found in parents’ speech to children is to refer to objects first and foremost by using basic-level count nouns (Callanan, 1985). It is possible that this type of exposure explains children’s tendency to map new words onto objects of similar kind (Waxman, 1991). A further example is that parents rarely use multiple labels to refer to the same object, and when they do it is usually qualified by clarifying expressions (Callanan & Sabbagh, 2004). Callanan and Sabbagh concluded that parents’ labeling tendencies seemed to endorse the development of an expectation in children that there is a “best” name for each object,
if not an assumption that each object can have only one name. Consistent with this latter possibility, studies found that how parents introduce nouns depends on the familiarity of the nouns to the child (Masur, 1997). Specifically, mothers typically question children about the referents of familiar names, but directly provide the names of novel objects. This parental strategy might encourage the development in children of an assumption that novel words refer to novel objects.

In sum, the “input proposal” suggests that not only the rate but also the manner by which children acquire words are significantly affected by the way their parents talk to them. The emphasis on parents as the source of the constraints makes this an exogenous mechanism, even if, eventually, children internalize the biases detected in the input. The fact that the constraining factor is linguistic input makes this a highly lexical-specific account.

Syntax

Another language-specific mechanism argued to help children figure out the meanings of words is their knowledge of syntax. The argument is that via an awareness of the different characteristic meanings associated with particular grammatical classes of words, children can substantially narrow down the meanings of words. For instance, by recognizing that a novel word they hear is a count noun, children can then search for the possible kind of object the word refers to. Similarly, children should search for the characteristic of an object when hearing a novel adjective, the particular individual referred to by a proper noun, or the unique action portrayed by a novel verb.

Evidently, for this mechanism to be effective, children need to be able to distinguish among the syntactic frames in which the different types of words appear. There is mounting evidence that, even before their second birthday, children indeed do make these distinctions, differentially extending words based on whether the words are syntactically framed as proper nouns, count nouns, or adjectives (Brown, 1957; Gelman & Taylor, 1984; Hall, 1994; see also Naigles & Swensen, this volume).

In a striking demonstration of children’s sensitivity to lexical form, Hall and colleagues showed 4-year-olds animate-looking novel entities, and described the entities either with a salient familiar adjective (e.g., “This is a red one”) or with a descriptive proper noun (e.g., “This is Mr. Red”) (Hall, Waxman, Bredart, & Nicolay, 2003). Children were told a story about an entity that lost its salient characteristic (e.g., it became blue). Children were then asked if the same entity, now with a different characteristic, or a different entity but with the same original salient characteristic, was the referent of the adjective or proper noun. Hall et al. found that children in the proper noun condition selected the same entity with a different characteristic, but children in the adjective condition selected the different entity with the same characteristic.

Extending this line of work, Hall and Graham (1999) examined whether children make different kinds of inferences about the relationship between two words depending on the lexical class to which they belong. The experimenter showed children pairs of familiar animals, and taught them either a novel proper name (“This dog is named Daxy”) or a novel adjective (“This dog is very daxy”) for one of the animals. Children
were then asked for the referent of a second novel proper name or adjective. Hall and Graham found that while children in the proper name–proper name condition tended to select the animal that lacked a proper name, children in the adjective–adjective condition tended to pick randomly between the two animals. The authors concluded that children seem to recognize that because proper names designate individuals, it is probable that two different proper names designate two different individuals. In turn, given that adjectives describe characteristics of individuals, it is plausible that two different adjectives nonetheless apply to the same individual.

**Lexical constraints**

Arguably, the most direct solution to the word-learning induction problem is that children are equipped with lexical constraints. Markman (1989) suggested that children have a set of three constraints that help limit the possible meanings of a word: a whole object bias (words refer to whole objects, not parts or features), a taxonomic bias (words refer to kinds of things, not individuals), and a mutual exclusivity bias (every object can have only one name). These three constraints work together and help children infer the meanings of words. For instance, if children are shown a familiar object, and an adult labels it with a novel name, mutual exclusivity leads children to overcome the whole object bias and to infer that the label probably refers to an unknown part or characteristic of the object (Markman & Wachtel, 1988).

Although Markman (1992) discusses the possibility that the origins of these biases are in general cognitive mechanisms, most of the studies attempt to show that the biases are unique to word-learning contexts. For instance, in traditional tests of the taxonomic bias the experimenter presents children a familiar object (e.g., a banana), labels it with a novel name, and then asks children to extend the name to either a taxonomically (e.g., an apple) or a thematically (e.g., a monkey) related object. Children typically extend the name based on taxonomic similarity. However, if asked to pick an object that “goes with” the target object, children pick based on thematic similarity (Imai, Gentner, & Uchida, 1994; Waxman & Namy, 1997). Furthermore, in line with the idea that these biases are especially dedicated for word acquisition, Markman and colleagues found that children respond in ways consistent with these biases from very early on in development (Littschwager & Markman, 1994; Markman, Wasow, & Hansen, 2003; see also Halberda, 2003).

In an attempt to capture more precisely the dynamic interactions among various biases, Golinkoff, Mervis, and Hirsh-Pasek (1994) proposed a system of six lexical principles, divided into two developmental tiers. One of the basic differences from Markman’s proposal is that some of the principles were themselves the results of word learning which, once developed, helped children further narrow the meanings of words. The first tier of principles consists of biases that get word learning started. They are: reference (words stand for objects), extendibility (words extend beyond the individual item labeled), and object scope (words refer to objects). The second tier consists of principles that evolve from the first tier principles. They are: conventionality (there are standard/expected words used to refer to things), categorical scope (words are extended on the basis of
taxonomic similarity), and novel-name nameless-category (new words refer to nameless objects). Despite the differences, this “developmental principles” approach shares with Markman’s proposal the emphasis on the lexical specificity of the principles, and the fact that the primary source of the constraints rests within the child rather than explicitly in the input (see also Merriman & Bowman, 1989; Waxman, 1991, for other proposals).

A final point regarding the lexical constraints account is that the constraints are not viewed as fully deterministic and encapsulated (Behrend, 1990; Woodward & Markman, 1998) but rather as default assumptions that can be overcome in the presence of contradictory evidence (for examples see Au & Glusman, 1990; Diesendruck & Shatz, 1997, 2001).

Attention and learning mechanisms

The least lexical-specific account of word learning stipulates that it can be explained by the general mechanisms of associative learning coupled with the operations of attention and memory. Indirect support for this account comes from studies revealing the sophisticated capacities of young infants to detect patterns in streams of linguistic signs (Gerken; Saffran & Thiessen, this volume).

The word-learning phenomenon addressed in most detail by such an account is children’s “shape bias,” that is, the tendency to favor extending object names based on shape similarity over similarity on any other dimension (e.g., color, material, size) (Jones, Smith, & Landau, 1991; Landau, Smith, & Jones, 1988). According to the attention and learning account, most object names to which English-speaking children are exposed are count nouns describing shape-based categories (Samuelson & Smith, 1999). Given this regular association between a count noun construction (e.g., “This is an X”) and the physical dimension of shape, every time children hear a novel count noun their attention automatically turns towards the shape of objects. In other words, general learning mechanisms pick out a regularity in the linguistic input to children, thus establishing a firm lexical-specific bias. Consistent with this account, a shape bias is manifested when 3-year-olds are asked to extend names, but not when asked to pick objects that go together (Landau et al., 1988). Furthermore, longitudinal studies show a relationship between children’s acquisition of a vocabulary of shape-based object names and their shape bias (Gershkoff-Stowe & Smith, 2004), and training studies reveal that as 2-year-olds get taught shape-based object names so their shape bias increases (Smith, Jones, Landau, Gershkoff-Stowe, & Samuelson, 2002). Also, connectionist networks develop a shape bias similar to that of children as a result of being trained in an input corpus of words similar to the ones children are exposed to (Samuelson, 2002). Importantly, while children’s initial word-learning bias is directed at shape – presumably reflecting the distributional frequency in the input – as children develop they come to notice further regular associations between different types of words in the input and their respective referents, thus developing biases tuned to these finer distinctions (Smith, 1999).

The importance of attention and memory in driving children’s learning of words was further illustrated in a study by Samuelson and Smith (1998). In that study, an experimenter showed 2-year-olds three novel objects while sitting on the floor. The
experiment then invited children to move to a table, where she showed children a fourth novel object. The experimenter and the child returned to the floor, the experimenter placed all four objects on a tray, and then asked the child to find the referent of a novel word. Samuelson and Smith found that 2-year-olds tended to select the object seen at the table in response to the experimenter’s request. The authors argued that given the distinctiveness of the “table” context, the object seen in that context stood out in children’s memory, and was consequently the most attention-grabbing.

In sum, this account builds on general cognitive mechanisms that are not specific to word learning, but which, in interaction with regularities in the linguistic input children are exposed to, may give rise to lexical biases.

Conceptual biases

The problem of inducing word meanings is massively underdetermined because a word can refer to any logically plausible concept. Psychologically, however, the problem could be more constrained if we were a priori biased to entertain only certain concepts. Specifically, children might have sufficient knowledge about the kinds of things that exist in the world so as to elaborate narrow hypotheses about the possible meanings of new words they encounter (Carey, 2001; Spelke, 1994). Among this knowledge are distinctions between individuated and non-individuated entities, between objects and actions, between animate and inanimate beings, and even between within-domain categories (e.g., between cats and dogs). Given this rich prelinguistic knowledge, upon encountering a novel word applied to a novel entity the child first decides upon the most likely conceptual category to which the potential referent belongs, and then maps the word onto that category.

In an investigation of this kind of bias, Soja, Carey, and Spelke (1991) showed 2-year-olds either novel solid objects or novel piles of stuff. The experimenter applied a novel word to such a target entity, and then asked children to generalize the word to either an entity of similar shape but made of different material, or an entity of different shape but made of the same material. It was found that when the word was applied to a solid object, children extended the word based on similarity of shape; when the word was applied to a pile of stuff, children extended the word based on similarity of material. Importantly, this pattern was found irrespective of whether children understood the syntactic difference between count nouns and mass nouns. The authors concluded that children’s conceptual differentiation of countable and non-countable entities predates the acquisition of the linguistic distinction, and in fact guides the acquisition of word meanings (see also Prasada, Ferenz, & Haskell, 2002).

Another type of conceptual knowledge children rely on has to do with the conceptual category to which the referent of a noun belongs. For instance, Diesendruck, Gelman, and Lebowitz (1998) found that the same type of information about two entities differentially affected 3- and 4-year-olds’ pattern of name extension depending on whether the entities were animals or artifacts. In particular, telling children that two animals had the same internal properties led children to accept the same name for the animals, but the same type of information applied to two artifacts did not have the same effect.
Finally, a number of studies show that young children are more likely to interpret a novel word applied to an object as a proper name if the object is animate rather than inanimate (Imai & Haryu, 2001; Sorrentino, 2001).

All these studies illustrate how the way in which children construe the world may influence the kinds of hypotheses they entertain about the meaning of a word. These are endogenous constraining factors, but they are not specific to lexical acquisition: children’s concepts also constrain how they categorize and reason about objects and events.

Pragmatics

According to the social–pragmatic account, when children hear a new word they are primarily interested in figuring out what is in the speaker’s mind (Akhtar & Tomasello, 2000; Baldwin & Moses, 2001; L. Bloom, 1998; P. Bloom, 2000). The argument is that by being sensitive to the communicative context, cues about the speaker’s behaviors, and a speaker’s state of knowledge and dispositions, children can substantially narrow down the plausible referential intents of a speaker and consequently the possible meaning of a word.

This social–pragmatic account presumes that by the time children start acquiring words they have some sensitivity to people’s mental states. Indeed, there is a growing body of research suggesting that by their second birthday children interpret people’s behaviors in psychological terms (Baldwin & Meyer, this volume). As Baldwin and Meyer review, the empirical evidence that children recruit this understanding of minds for word learning is also substantial. For instance, by 2 years of age children learn words primarily in communicative contexts, following and keeping track of speakers’ referential intents, and attending to the intentionality in speakers’ interactions with objects.

Various studies also show that from a young age children are sensitive to a speaker’s state of knowledge when inferring his/her referential intent. Akhtar, Carpenter, and Tomasello (1996) found that 2-year-olds interpret a speaker’s novel word as referring to an object that was novel to the speaker in a communicative context. This attentiveness to speakers’ knowledge is also manifested in children’s sensitivity to different types of words and speakers. For instance, 3-year-olds are aware that speakers might know the common names of novel objects even if they have never been exposed to them, but will not know the proper names of objects under the same circumstances (Birch & Bloom, 2002; Diesendruck, 2005). Moreover, children monitor speakers’ expressed knowledge of language. Thus, even before their second birthday (Koenig & Echols, 2003), but also after that (Sabbagh & Baldwin, 2001; Sabbagh, Wdowiak, & Ottaway, 2003), children infer that speakers who have been unreliable in their knowledge of words are likely not good sources for learning new words. By 4 years of age children draw the same kind of inference about speakers who speak a different language (Diesendruck, 2005).

These sensitivities to speakers’ intents and knowledge states have encouraged the development of pragmatic accounts of phenomena originally presented as examples of the operation of other mechanisms. One example has to do with the mutual exclusivity bias or novel-name nameless-category principle described before. These lexical
constraints explain why it is that children tend to select an object without a known label – as opposed to an also available familiar object – in response to an experimenter’s request for the referent of a novel label. The pragmatic account of this phenomenon is that children might reason that if the experimenter had wanted the familiar object, she could have simply asked for it by using its conventionally known name. The experimenter’s use of a different name probably indicates that she intended to refer to the other object (Clark, 1988, 1990). A series of studies now support this pragmatic account (Diesendruck, 2005; Diesendruck, Hall, & Graham, 2006; Diesendruck & Markson, 2001). First, children respond in the same way when asked for the referent of a novel fact – thus the inference is not about count nouns only, but about referential acts in general. Second, children respond in this way only if they are certain that the experimenter knows the information associated with one of the objects – thus the inference is not based only on what children know, but instead is modulated by what the speakers know. And finally, children respond in this way only if the linguistic information attached to the objects is clearly referentially constraining – thus the inference is not about all communicative acts.

Altogether, the emphasis on children’s understanding of speakers’ minds makes this an endogenous, not lexical-specific, account.

Empirical Challenges

In this section, I will describe some of the empirical challenges to certain conceptualizations of the mechanisms, particularly with respect to the specificity of certain word-learning phenomena, and the precise nature of the endogenous mechanisms.

How specific is word learning?

A number of proposals argue that children’s inferences of word meanings are driven by lexical-specific biases or constraints. Recent findings challenge this contention. For instance, according to the attention and learning account, children’s initial shape bias is specific to naming because it derives from the interaction of these general mechanisms with specific linguistic input (Smith, 1999). Contrary to this position, Diesendruck and Bloom (2003) found that 2- and 3-year-olds were as likely to favor shape over color or material when asked to pick objects “of the same kind” as a target object as they were when asked to pick objects with the same name as the target. In fact, this preference for shape when categorizing objects seems to precede children’s word learning, intimating that the bias may not derive from exposure to a lexical-specific regularity (Diesendruck, Graham, & Onysyk, 2004; Samuelson & Smith, 2005).

A second example regards lexical constraints. On the one hand, words more efficaciously facilitate object categorization by infants than do other auditory stimuli (Balaban & Waxman, 1997), 2-year-olds are more likely to accept a word as a referential symbol than they are to accept a gesture (Namy & Waxman, 1998), and 3-year-olds are more
likely to extend a word to taxonomically related objects than they are to extend a fact (Behrend, Scofield, & Kleinknecht, 2001; Waxman & Booth, 2000). On the other hand, when words are compared with other intentional acts, some of these advantages of words disappear (e.g., Childers & Tomasello, 2003; Markson & Bloom, 1997). Similarly, with regards to the mutual exclusivity bias or the novel-name nameless-category principle, Diesendruck and Markson (2001) found that 3-year-olds were as likely to select an object without a label in response to a request for the referent of a different novel label as they were to select an object without a fact in response to the request for the referent of a different novel fact.

The crucial questions coming out of this mixed pattern of comparative advantages of labels are to what extent, in which respects, and why are labels special? We have some clues for answers to these questions. For instance, sounds seem to function as referential signs only when paired with extensive intentional cues (Campbell & Namy, 2003). This raises the questions of the age at which and why labels are treated as intentional acts. A further example is that while facts might not be readily extended to other referents, certain novel actions clearly performed intentionally on objects are (Childers & Tomasello, 2003). Thus, what is the relationship between intentionality and extendibility? Finally, there are differences in how easily children learn to map certain kinds of facts to objects (Markson & Bloom, 1997), and in the kinds of assumptions children make about people’s knowledge of facts and words (Diesendruck & Markson, 2001). This raises the question of how is it that children come to develop distinct expectations about these various types of information.

More broadly, these challenges to the specificity of word-learning phenomena emphasize the need to better define the cognitive capacities responsible for children’s inferences of word meanings.

*What does the child bring to word learning?*

While speech directed to children can influence what words children acquire, and how they do so (see *Input* above), studies reveal that children are quite capable of going far beyond this type of input. For instance, studies reveal that 2-year-olds manage to learn the referents of labels by monitoring naming events in third-party conversations (Akhtar, Jipson, & Callanan, 2001), even if they are engaged in a distracting activity at the time the naming occurs, or if the label is embedded in a non-salient sentence position (Akhtar, 2005). In fact, children are as likely to map a word to its correct referent when directly taught the mapping as when only indirectly exposed to it (Jaswal & Markman, 2001).

What are the cognitive capacities that allow children to make these indirect mappings? For one, children’s rich conceptual beliefs certainly modulate their attention, such as to help them make adequate inferences. One example of such modulation has to do with the shape bias. Specifically, Jones et al. (1991) found that the addition of eyes to novel objects led 3-year-olds to extend names based not only on objects’ shape but also their texture. Jones et al. explained this finding by arguing that, in the input children are naturally exposed to, objects with eyes are regularly associated with commonality in both shape and texture, thus leading to the development of this biased association. An
alternative interpretation of this finding is that the presence of eyes indicated to children that the objects in question were animate beings, about which children have particular conceptual beliefs. Booth and Waxman (2002) showed 3-year-olds novel objects (with or without eyes), taught children novel names for the objects, and then asked them to extend the names to other objects. The critical manipulation was that, via stories, the experimenter conveyed to children that the objects were either animate or inanimate. Booth and Waxman found that what determined children's pattern of extension was not the presence or absence of eyes, but rather whether the objects were conceived of as animate or inanimate. Children extended the names of inanimate objects based solely on shape similarity, and the names of animate objects based on both shape and texture.

A recent study revealed a similar pattern of findings with 18-month-olds, indicating that from very early on children's name extensions are grounded in conceptual knowledge (Booth, Waxman, & Huang, 2005).

A second type of cognitive capacity that helps modulate children's inferences is their burgeoning understanding of minds. Recent studies suggest that what might seem like modulation via other mechanisms, such as attention and memory or syntax, might in fact be hiding an understanding of intentions. For instance, Samuelson and Smith (1998) argued that children's choice of a particular object presented to them at a table – rather than at the floor – as the referent of a novel label derives from the relative contextual salience of the object in the children's memory. Diesendruck, Markson, Akhtar, and Redor (2004) addressed an alternative interpretation of their findings, namely, that children's choice of the object derived from their interpretation of the experimenter's move to the table as a deliberate act reflecting the special status of that object to the experimenter. To test this hypothesis, we introduced a condition in which the change in context (i.e., from the floor to the table) was evidently accidental, as opposed to a condition in which it seemed intentional; and we introduced another condition in which the speaker asking the child for the referent of a novel name was not the experimenter but rather a different speaker who had no previous interaction with the objects. We found that in these two additional conditions, 2-year-olds did not associate the novel name with the object they were exposed to at the table. Thus, children's inference of the referent of the new name did not derive solely from the fact that an object was seen in a novel context, but rather from children's inferences about the intent underlying a speaker's interaction with the object.

An understanding of intentions also seems to modulate the role of syntax in driving children's inferences about word meanings (Diesendruck et al., 2006). In this study, children were taught either a novel prenominal (e.g., “This is a very daxy dog”) or a predicate adjective (e.g., “This dog is very daxy”) about one of two identical looking familiar animals. Children were then asked for the referent of a different novel prenominal or predicate adjective. Prenominal adjectives differ from predicate adjectives in that they restrict the reference of the noun attached to them. The question was whether children would respect this difference, and thus infer that the use of a different noun-restricting prenominal adjective likely indicated that the speaker had a different referent in mind. Or instead, whether children would treat both kinds of adjectives as describing characteristics of individuals, and thus conclude that the speaker in both conditions could be referring to either animal (as Hall & Graham, 1999, had found using only
Mechanisms of Word Learning

269

predicate adjectives). The results supported the first possibility: children in the prenominal condition selected the unlabeled animal, and children in the predicate condition selected randomly. That is, children selectively applied an inference of contrast based on their understanding of the semantic implications of different adjectival syntactic frames.

These studies indicate that children’s inferences about the meanings of words are informed not only by their sensitivity to syntactic cues or their basic attentional tendencies, but also by their conceptual beliefs and assumptions about speakers’ communicative intents given particular linguistic and behavioral contexts.

Integrative Approaches

The review of the six different mechanisms and the empirical challenges makes it apparent that no single mechanism can fully account for the relevant phenomena. One possible conclusion to be drawn from this review is that all these mechanisms contribute to word learning. However, in an attempt to maximize parsimony and explanatory power, a number of integrative proposals have been put forth. Each proposal highlights a different subset of the mechanisms that, in combination, are sufficient to account for the range of phenomena. The theoretical benefit of these approaches is that they make specific explanations and predictions about how children learn words in various contexts, based on the postulated relative weights of the different mechanisms at play. I will focus on two prominent such approaches.

Emergentist model

One of the most detailed integrative accounts of word learning is the so-called “emergentist coalition model” (Hollich, Hirsh-Pasek, & Golinkoff, 2000). The model postulates three basic assumptions about the word-learning process. The first assumption is that children rely on multiple mechanisms, such as their knowledge of syntax, their sensitivity to attention-grabbing salient aspects of the context, and their understanding of speakers’ intentions. The second assumption is that children differentially weight these mechanisms, and that the weights actually change with children’s development. For instance, early on in development, perceptual salience may be more heavily weighted than intentional cues. The final assumption is that as a consequence of development, children’s word-learning tendencies might become automatized. That is, word-learning principles emerge as a result of word learning, rather than being present from the outset. Once in place, however, these lexical principles enhance and expand children’s word learning.

Much of the evidence in support of the emergentist model comes from studies tracking how children’s attention to various cues changes with development (Hollich et al., 2000). For instance, while 12-month-olds seem to rely primarily on object salience over the direction of a speaker’s eye gaze to determine the referent of a novel word, 20-
month-olds are sensitive to speakers’ eye gaze but are still attracted by object salience, and it is only by 24 months of age that eye gaze determines children's referent choice. Arguably, as children’s understanding of intentions and of the symbolic nature of words develops, social cues become more heavily weighted among all the cues driving word learning. Further evidence for the model comes from studies showing that certain lexical principles appear to become active only after a substantial amount of vocabulary has already been acquired. In particular, the tendency to apply novel names to nameless categories (Mervis & Bertrand, 1994), and to link names onto categories more generally (Nazzi & Bertoncini, 2003), arguably emerges only after a vocabulary spurt. A similar case can be made for the role of syntax in word learning. Evidently, syntax can only become a factor after children have acquired a significant number of words. That is, syntax probably cannot help children in the earliest stages of acquisition.

Theory of mind, concepts, and syntax

P. Bloom (2000) argues that children learn words by relying on three basic capacities: an understanding of mental states, an understanding of the kinds of things that exist in the world, and knowledge of syntax. In Bloom’s view, children's learning of words is not driven by specially dedicated mechanisms such as lexical constraints but rather results from mechanisms children rely on for understanding a variety of their experiences, particularly social ones.

Evidence in support of Bloom's position comes from a number of studies already reviewed, for instance that word-learning responses presumably driven by lexical constraints or attentional mechanisms might actually be a product of intentional inference (Diesendruck & Markson, 2001) or conceptual understanding (Diesendruck & Bloom, 2003). Importantly, there is also evidence consistent with the idea that these basic capacities indeed interact so as to sufficiently account for word learning. For instance, Birch and Bloom (2002) demonstrated that 2-year-olds can (1) identify proper nouns based on syntactic information, and (2) draw pragmatically appropriate inferences about a speaker’s referential intent when using a proper noun – namely, speakers use proper names to refer to familiar objects.

Concluding Remarks

Most students of word learning agree that in order to account for the range of pertinent phenomena we need some integrative approach (see Hall & Waxman, 2004, for further proposals). One of the major challenges to such integrative approaches, however, is to provide a sufficiently detailed mechanistic account of how the various factors interact (see, e.g., Shatz, in press). How are the relative weights of the different mechanisms defined? Are they context-specific? Are they age-sensitive? If the answer to all these questions is affirmative, as they might turn out to be, then we may have to compromise with a fairly piece-meal “theory” of word learning. I want to end with a more optimistic perspective.
The common theme underlying integrative approaches is that multiple mechanisms, with varying weights, are involved in word learning. This idea opens the possibility that there are various ways in which these mechanisms can be combined for the task of word learning. In principle, nobody denies this possibility. De facto, however, it is still part of the mainstream research agenda to attempt to rule out mechanisms by demonstrating their irrelevance in specific contexts. What such an “eliminativist” agenda might end up revealing, however, is the relative weight of a mechanism in a specific context, or the degree of attunement of the mechanism.

Take for instance the research on the variety of word-learning populations. Parrots (Pepperberg & Wilcox, 2000), chimpanzees (Savage-Rumbaugh et al., 1993), and dogs (Kaminsky, Call, & Fischer, 2004) have been shown to understand, and some even produce, word-like symbols. Rico the dog, for instance, not only understands over 200 words, but also seems to map novel words onto objects for which he does not have a name. Recent work shows that infants as young as 14 months of age also can learn word-to-object pairings (Schafer & Plunkett, 1998; Werker, Cohen, Lloyd, Casasola, & Stager, 1998), and even 12-month-olds seem to honor mutual exclusivity (Halberda, 2003; Markman et al., 2003; Xu, Cote, & Baker, 2005). Last but not least, studies have found that differently from typically developing 18-month-olds, children with autism do not rely on the direction of eye gaze of a speaker to determine the referent of a word (Baron-Cohen, Baldwin, & Crowson, 1997). That is, they do not seem to recruit an understanding of intentions to acquire words. Nonetheless, despite this impairment, children with autism map novel words onto novel objects (Preissler & Carey, 2005). All these findings are taken to challenge the idea that typically developing children’s mapping of novel names to novel objects derives from their sensitivity to speakers’ intents.

There are two logically possible responses to this type of challenge. One is that all these populations have the basic understanding of minds presumed to be necessary for word learning. In fact, there are suggestions that this might indeed be the case (e.g., Onishi & Baillargeon, 2005; Tomasello, Call, & Hare, 2003); but these are all highly contentious claims (see Perner & Ruffman, 2005; Povinelli & Vonk, 2003).

A second type of defense – and the one I want to stress here – is that the cocktail of underlying mechanisms driving word learning in animals, infants, and children with autism might be different from the one used by typically developing children. For instance, it is possible that while Rico’s selection of novel objects derives from novelty preference, and that children with autism’s word learning is driven by associative training, typically developing children solve these problems by inferring speakers’ intents. An implication of this interpretation of the data is that instead of stopping at documenting the similarity on some outcome measure we should look at finer-grained consequences of such distinctive types of word learning. For instance, what kind of word-teaching paradigm works best for children with autism compared with for typically developing children? Can we find traces of differential acquisition processes in children’s uses of words? What are the assumptions that typically developing children, as opposed to other word-learning populations, make about words?

Another example of this eliminativist agenda is manifested in attempts to systematically trace when certain mechanisms are important for word learning. The rationale there is that if children are not sensitive to a certain type of cue (e.g., speakers’ eye gaze), then the underlying mechanism implicated by the cue (e.g., understanding intentions)
is probably irrelevant – and vice versa. This equivalence between “contextual cues” and “mechanisms,” however, requires caution. For instance, children might over-attribute intentionality in contexts where none was presumed to be involved. An example of this kind of dynamics can be found in studies reviewed earlier. Specifically, what Akhtar et al. (1996) interpreted as sensitivity to speakers’ knowledge state, Samuelson and Smith (1998) viewed as attention to novelty. What Samuelson and Smith then manipulated to test their hypothesis, Diesendruck et al. (2004) deemed as a manipulation of intentionality. This has important implications for how we interpret the developmental data, and how we design our studies. Lack of sensitivity to a cue might not mean that the capacity is unavailable, but rather that it is not tuned to that cue yet.

Addressing all these issues will bring us closer not only to a theory of word learning, but to an appreciation of the intricate links between this skill and the various other cognitive capacities that so uniquely characterize the human mind. That is the promise, and the appeal, of the study of word learning in children.

References


Mechanisms of Word Learning


