Children’s sensitivity to the shared, conventional nature of word meanings makes their word learning more efficient and less prone to error. After reviewing the evidence in support of this claim, we suggest that children’s earliest appreciation of conventionality might be rooted in limitations in their theory-of-mind skills.

How an Appreciation of Conventionality Shapes Early Word Learning

Mark A. Sabbagh, Annette M. E. Henderson

Language enables social coordination and regulation among humans. By using the right words, speakers can make others aware of their otherwise private mental states and thus communicate their knowledge, desires, emotions, and intentions (H. H. Clark, 1996). However, it is not obvious how humans, especially young word learners, come to understand which words are likely to be the right ones. The relation between most word forms and their meanings is arbitrary (de Saussure, 1965); with few exceptions, there is nothing about a given word that naturally signals what that word might mean. For instance, there is nothing about the word cat that obviously or naturally signals that it refers to the purring, furry, four-legged, whiskered, domesticated mammals that can be found in many homes. Yet this fundamental truth about the relation between words and their referents does not square with our deep, everyday intuitions about the nature of language. For instance, although the relation is fundamentally arbitrary, even young infants recognize that only words, and not whistles, gestures, or other non-linguistic entities, constitute an appropriate way of referring to things (Namy & Waxman, 1998; Woodward & Hoyne, 1999). Moreover, there is a real sense in which, once established, word-referent links are correct in the prescriptive sense; infants as young as sixteen months also seem to believe that referring to a cat with any term other than the word cat is not only unusual; it is wrong (Koenig & Echols, 2003).
What is it, then, that enables words to transcend the inherent arbitrariness of their relations with their meanings and become prescriptive, correct facts about the world? The answer to this question lies in the basic claim that word meanings acquire their status as given facts (however tenuous) about the world through their everyday use by speakers within social contexts (Wittgenstein, 1953). More specifically, language users can parry the arbitrariness of word forms by relying on the fact that words will effectively communicate their intended meanings so long as those form-meaning associations are known, shared, and expected within a community of speakers. Within the psychological literature, this understanding has been called the principle of conventionality (E. Clark, 1988, 1990, 1993).

In addition to providing a much needed foundation for the everyday use of fundamentally arbitrary signs and symbols, there are at least three ways in which an appreciation of conventionality might aid children's acquisition of these signs (see E. Clark, 1992; Golinkoff, Mervis, & Hirsch-Pasek, 1994). First, conventionality provides for the extension of newly acquired word meanings to all members of the linguistic community. Second, conventionality provides a basis for lexical error avoidance. Third, conventionality may provide a basis for the contrast effect. This chapter first reviews the small but burgeoning body of evidence suggesting that children show these three benefits that an appreciation of conventionality might bestow on word learning. We conclude by addressing the question of where an appreciation of conventionality might come from. In particular, we advance the hypothesis that conventionality might emerge from limitations in young children's theory-of-mind skills.

**Speaker-Extension Effects**

An appreciation of conventionality might be particularly useful to children acquiring language through what we call the speaker-extension effect. Imagine a situation in which a young child is reading a picture book with a parent, who then provides a novel label for a novel object. Although the child hears the word used initially only by that one parent, it is very likely that the word will be known at least by others in the family, if not all speakers within the child's linguistic community. An assumption of conventionality—that word meanings are known, shared, and expected by other speakers—could potentially enable children to efficiently extend knowledge of word meaning to the other members of their linguistic community. That is, children would not have to wait to hear each individual speaker within a community use a particular word form before assuming that the word meanings are widely known. A growing body of evidence suggests that children appreciate that word meanings are shared across speakers quite early in development.

In one study, Diesendruck and Markson (2001) leveraged the disambiguation effect to provide evidence for conventionality in three- and four-
year-old children. The disambiguation effect (Markman & Wachtel, 1988; Merriman & Schuster, 1991) occurs when children are presented with two objects, one that is named and one that is unnamed, and are provided with an indirect label, for example, “Which one is the blicket” In this situation, young children typically select the unnamed object as the referent for the novel label. With this as their starting point, Diesendruck and Markson (2001) developed a situation within which an experimenter showed three- and four-year-old children two novel objects and then proceeded to label only one: “This is a fep.” A disambiguation test followed in which children were provided with an indirect label ("Which one is the wug?") by either the same experimenter who introduced the first label or a second unfamiliar experimenter. It was predicted that if children appreciate that word meanings are shared across speakers, they should show the disambiguation effect even when asked by the unfamiliar speaker. That is, children should assume that even the novel speaker (who has not been seen to use the word fep before) knows the name of the named object and thus assume that their novel label must be associated with the novel object (this second assumption may also have to do with conventionality, and we discuss this here). These findings seem to indicate that by the time children are three to four years old, they show the speaker-extension effect of conventionality—that is, they assume that words’ meanings are widely shared by speakers in the linguistic community.

Even more impressive is that the speaker-extension effect of conventionality may be present in very young word learners. Using a simpler paradigm, Henderson and Graham (2005) taught twenty-four-month-old toddlers the meaning of a new word in a nonostensive word learning paradigm. In a subsequent comprehension test, children were asked to select the referent of the novel label by either the same speaker who taught them the meaning of the new word (the same-speaker condition) or a second speaker who was not present during the learning phase (the different-speaker condition). It was predicted that if children assume that the knowledge of word meanings is shared by members of the same linguistic community, children in the different-speaker condition would be just as likely to select the target object as the children in the same-speaker condition. The findings were consistent with this prediction. There were no condition differences in the number of children who selected the target object. This pattern of results was also found in a follow-up study with children nineteen months old (Graham, Stock, & Henderson, 2006). Taken together, these studies suggest that nineteen-month-old children show the speaker-extension effect of conventionality.

Although the speaker-extension effect of conventionality is exceptionally useful in the case of language, there are other kinds of world knowledge for which such the application of such an effect would lead to error. For instance, speakers’ preferences and desires toward objects are not shared across speakers in the same way that object labels are. Thus, an intriguing question is whether children avoid applying the speaker-extension effect of
conventionality across these more idiosyncratic domains. This question has been explored in studies conducted by Diesendruck (Diesendruck, 2005; Diesendruck & Markson, 2001). For instance, in the Diesendruck and Markson (2001) study already described, the authors included a contrasting condition in which they showed that while preschoolers assume that knowledge of object labels is likely to be shared by members of the same linguistic community, they do not assume the same of idiosyncratic personal facts (“My uncle gave this one to me”). In a similar vein, Diesendruck (2005) showed that while knowledge of object labels (“This is a teega”) was likely to be widely shared, knowledge of proper names (“This is Teega”) was not. Very young word learners also show appropriate restriction of speaker-extension effects; twenty-four-month-olds do not assume that speakers will share preferences (Henderson & Graham, 2005), and nineteen-month-olds do not assume that speakers will share desires (Graham, Stock, & Henderson, 2006).

Together, the findings from the research suggest that nineteen-month-olds show the speaker-extension effects of conventionality. That is, they seem to assume that the meanings of object labels are shared by other members of the linguistic community. Impressively, the evidence suggests that children this age distinguish the shared nature of object labels from the non-shared nature of object preferences.

Error-Avoidance Effects

An assumption of conventionality allows speakers to sidestep the inherent arbitrariness of the relation between a word and its referent and judge particular word forms as being right or, equally important, wrong. In the terms of conventionality, a word meaning is “right” when it is likely to be shared by others and “wrong” when it is unlikely to be shared. Recognizing this, children might streamline their word learning by showing what we call the error-avoidance effect of conventionality. This effect is shown when children block the learning of a novel word meaning when there is some evidence that the word is not likely to be shared by other speakers.

One line of evidence suggesting that young children show the error-avoidance effects of conventionality comes from research showing that children do not learn words from speakers who are obviously ignorant of an object’s conventional label. In one study, Sabbagh and Baldwin (2001) investigated whether children learn the meaning of a new word in such a situation. In this study, children were taught the meaning of a novel object label by a speaker who appeared to be either knowledgeable or ignorant at the time of labeling. After a brief delay, children participated in a comprehension test in which they were asked to select the referent of the new word. Children who were taught the meaning of the new word by the knowledgeable speaker selected the object that had previously been labeled. In con-
Contrast, children in the ignorant-speaker condition did not select the object that had previously been labeled. These findings suggest that children do not learn new word-referent links from speakers who appear ignorant or uncertain at the time of labeling.

Follow-up research provided some evidence to suggest that the error-avoidance effect works at the level of encoding. That is, when confronted with an ignorant speaker, young children do not seem to encode word-referent links that are offered by that speaker. Evidence for this comes from a study by Sabbagh, Wdowiak, and Ottaway (2003) in which children were taught the meaning of a new word by a speaker who was either knowledgeable or ignorant of the correct meaning of the new word. After the first labeling episode, a second speaker entered the room and used the same label but referred to a different object. The authors reasoned that if children had some trace in memory for the word-referent link offered by the ignorant speaker, they would experience proactive interference and show some cost in learning an alternative from the knowledgeable speaker (e.g., Murnane & Shiffrin, 1991). However, findings from comprehension test data showed that there was no detectable cost of having heard the ignorant speaker’s word-referent link. Indeed, children performed just as well in that condition as they did in a control condition in which the ignorant speaker provided no word-referent link at all. Taken together, the evidence suggests that children avoid learning new word-referent links from ignorant speakers.

Not only do preschoolers seem to avoid learning new words from speakers who explicitly aver their ignorance, but they seem also to avoid learning from speakers who have been “wrong labelers” in the past. Koenig, Clément, and Harris (2004) showed that when presented with speakers who were either consistently correct (they called a ball a “ball”) or consistently incorrect (they called a ball a “shoe”), three- and four-year-olds preferred to learn new words from the consistently correct versus incorrect speaker (see also Koenig & Harris, 2005). These findings suggest that children can show the error-avoidance effects of conventionality even when judgments about speakers’ poor reliability are behaviorally based as opposed to explicitly mentioned.

Together these findings clearly show that when presented with ignorant or unreliable speakers, children do not learn words. Moreover, there is some evidence to suggest that these judgments about speakers affect the semantic encoding of the word-referent links. Within the context of this discussion, these findings raise the intriguing hypothesis that the error-avoidance effects of conventionality might be instantiated through a top-down filtering process. That is, when confronted with word-referent links that are unlikely to be conventional (because they issue from ignorant speakers) children block the kind of rapid semantic encoding that is typical of everyday word learning. Future research might be aimed at better understanding the cognitive mechanisms that support this kind of top-down filtering mechanism.
Contrast Effects

The speaker-extension and error-avoidance effects are predicated on an assumption of conventionality in that the assumption provides that word meanings are shared by others within the linguistic community. An equally important aspect of the assumption of conventionality is that certain word forms are expected to refer to certain referents (E. Clark, 1992). For example, once the word *dog* is acquired for the four-legged, barking, domesticated animal, an assumption of conventionality stipulates that speakers continue to use that same word form for the same object. This simple assumption might potentially lead to a fairly pragmatic and powerful tool for word learning. Specifically, by making the initial assumption that speakers use agreed-on forms to express particular meanings, children can subsequently assume that if a speaker uses a novel word, its meaning (or referent) must also be novel, or at least contrasting with something that is already known. Following the seminal work of Eve Clark (1988), we call this the contrast effect of conventionality.

Evidence that word learners generally expect new words to have meanings that contrast with other word meanings is in abundance. The bulk of this evidence comes from children’s performance in a disambiguation test, as previously described. Recall that in the disambiguation test, children are presented with two objects, one named and one unnamed, and then provided with an indirect label (“Which one is the *blicket*?”). If children expect novel words to refer to unnamed or contrasting referents, they should typically pick the novel referent in response to the indirect label. Many studies have shown that preschoolers show this effect (e.g., Markman & Wachtel, 1988; Merriman & Bowman, 1989), as do younger word learners (Graham, Poulin-Dubois, & Baker, 1998; Markman, Wasow, & Hansen, 2003).

The only trouble with this research is that selection of novel or contrasting referents is not uniquely predicted by the contrast effect of conventionality. Indeed, this effect can be taken as evidence for a number of principles of lexical development, including mutual exclusivity (Markman, Wasow, & Hansen, 2003), lexical gap filling (Momen & Merriman, 2002), novel-name-nameless-category (N3C) (Golinkoff, Mervis, & Hirsch-Pasek, 1994), and others. It is well beyond the scope of this brief review to describe the many subtleties of these proposals and use the evidence to adjudicate among them. However, we describe attempts that have been made to gain evidence that would disentangle the contrast effect of conventionality from some of these other proposals.

One dimension on which the contrast effect of conventionality might distinguish itself from alternative accounts concerns that of speakers’ knowledge. Because an appreciation of conventionality lies at the heart of the contrast effect, the contrast account predicts that children will be impelled to select the novel referent in the disambiguation test when it is safe to assume that the speaker knows the known object’s label. If an assumption of speaker
knowledge is not safe, then children would have no basis for showing the contrast effect of conventionality and should perform unsystematically. Importantly, none of the other accounts of children’s performance in the disambiguation test predict that manipulations of speaker knowledge should have a similar effect. Thus, experiments using the disambiguation test in which children’s perception of speaker’s knowledge of the familiar object label is either weakened or strengthened might help provide evidence for the contrast effect of conventionality.

One study that has attempted to weaken children’s perception of the speaker’s knowledge of familiar objects comes from the studies conducted by Gil Diesendruck (2005). In one study, children were taught either a novel common noun (“This is a teega”) or a novel proper noun (“This is Teega”) for a novel object. Children were then given the disambiguation test by another speaker who, critically, was either absent or present during the initial labeling phase. In line with the contrast effect of conventionality, Diesendruck (2005) reasoned that children would show the disambiguation effect in the common noun conditions regardless of whether the speaker administering the disambiguation test was present or absent during the initial labeling. This is because common nouns are general properties of languages and, as previous research has shown, are assumed to be shared among speakers. In comparison, Diesendruck reasoned that the disambiguation effect would work only for proper names when the disambiguation test speaker was present for the initial labeling. More important, children were not expected to show the disambiguation effect when the speaker was not present, because the speaker would have had no knowledge of the proper name that was used. These predicted results were obtained, thus showing that children’s performance in the disambiguation test depended on whether the known label was shared by the other speaker. Similar conclusions were drawn from a second study (Diesendruck, 2005) in which the disambiguation speaker’s knowledge of the “known” object was undercut by not knowing the language.

To clarify, if children have reason to doubt that the known object is truly known by the speaker, then there is no basis for conventionality—that is, there is no reason to assume that the speaker would use a preferred form if the speaker does not know that form. Operating from the contrast effect of conventionality should have a complementary effect: evidence that a particular form is known and preferred should actually boost children’s performance in the disambiguation test. This possibility was tested in a pair of studies conducted by Megan Saylor and colleagues (Saylor & Sabbagh, 2004; Saylor, Sabbagh, & Baldwin, 2002). These studies took as their basis children’s potential use of the contrast effect to acquire the names of parts. Three- to four-year-old children were shown pictures of familiar whole objects that had a novel part highlighted on them. The familiar whole object was one color, and the part was another. Children were then provided with an indirect label in one of two conditions. In one
condition, children were first provided with the name of the whole object and then the novel label (“See this butterfly? What color is the thorax?”). In the other condition, children were not first provided with the name of the whole object (“See this? What color is the thorax?”). In both conditions, children were presented with a novel word that could have potentially mapped onto the unnamed part. However, results showed that children systematically mapped the novel label to the novel part only when the novel label (thorax) was preceded by the familiar one (butterfly). This is consistent with the contrast effect of conventionality: only when the speaker provided evidence of knowing the conventional term for the whole object did children show the disambiguation effect and map the novel term to the novel referent.

These findings provide some evidence that preschool children’s performance in the disambiguation test may reflect the contrast effect of conventionality. That is, children seem to expect speakers to refer to familiar objects with their familiar names as long as those names are known. Thus, when speakers know the names of familiar objects and then use a novel word, children can assume that the meaning of the novel word contrasts with the meaning of the familiar one. The findings showing that either undermining or reinforcing speakers’ familiarity with the familiar word systematically undercuts or enhances children’s performance in the disambiguation test and strongly suggests that their performance is predicated on an assumption of conventionality.

**Theory of Mind and Conventionality**

To this point, we have discussed three ways in which children’s lexical development might be affected by an assumption of conventionality. As noted from the outset, an assumption of conventionality provides a fundamentally social basis for the use and understanding of words and language. This consideration has important implications for understanding the conceptual and cognitive developments that likely underlie young children’s abilities to assume conventionality in the linguistic domain. In particular, a number of researchers and theorists have noted that successful negotiation of this social world rests, at least in part, on having a theory of mind: an understanding of the ways in which observable actions can be understood in terms of internal mental states, such as knowledge, beliefs, desires, emotions, and intentions (Wellman, 1990). Some aspects of theory of mind clearly play a central role in the everyday engagement of conventionality. For instance, a key feature of conventionality is that knowledge of word meanings is shared by others within the linguistic community. Presumably, then, children’s judgments about conventionality would draw on their theory-of-mind skills to allow them to make appropriate judgments about others’ knowledge states.
For this discussion, however, what might be most interesting about the time course and trajectory of theory-of-mind development is the limitations that are characteristic of very young children, and in particular eighteen- to twenty-four-month-old word learners. For instance, prior to the age of about four, children do not seem able to reason correctly about how people act when they hold false beliefs—beliefs that do not comport with a current state of reality. Instead, these young children seem to act as if their own knowledge states were shared by others (Wellman, Cross, & Watson, 2001). Although research has shown that false belief reasoning may be especially challenging for peripheral reasons (Carlson, Moses, & Hix, 1998), there is general consensus among researchers that reasoning about interindividual diversity in epistemic states is fundamentally limited in children younger than three and a half years old. For instance, it is not until around three and a half years that young children systematically reason that someone who has not seen the contents of a generic box will not know what is inside (Wellman & Liu, 2004). These findings suggest that it is not until the preschool years, or perhaps even later (Friedman & Leslie, 2004), that children truly appreciate that their own knowledge may not be shared by others.

One might expect that because of these limitations in theory-of-mind reasoning, children’s appreciation of conventionality might be similarly limited. But in fact it is more likely that the opposite is true. That is, these limitations in children’s abilities to reason about others’ epistemic mental states may actually promote the development of an appreciation of conventionality. To illustrate, we take as our starting point for this argument the extension effect of conventionality. We have discussed the extension effect as the appreciation that young children have that enables them to assume that word meanings are shared by all of the speakers of a given linguistic community, even if not all speakers have evidenced knowledge of the word meaning themselves. Implicit in this formulation is that if young word learners did not have an appreciation of conventionality, they could potentially assume that words are known only to speakers who have evidenced such knowledge. For example, without conventionality, a child who has learned from a parent an appropriate use of the word dog might be hesitant to use the word with others who have given no reason to believe that they also know the term.

However, limitations in children’s theory-of-mind reasoning skills should actually prevent children from considering such no-conventionality scenarios. To adequately represent the particular no-conventionality scenario described, children would have to understand that the knowledge they have about a word meaning is not shared by others. Children have a difficult time judging others as ignorant when they themselves are knowledgeable about something (although see O’Neill, 1996 for some intriguing data to the contrary). Other no-conventionality scenarios might pose similar demands on young children’s theory-of-mind reasoning. For example, in a
no-conventionality world, children might believe that although they and their parents call dogs “dogs,” this designation might not be shared by others who could conceivably have their own designations. Although there is some evidence that reasoning about diverse beliefs is easier than reasoning about others’ ignorance (Wellman & Liu, 2004), there is consensus that the achievement does not happen prior to children’s third birthday.

The upshot is that children’s early understanding of conventionality might be a case of less is more. Children’s theory-of-mind limitations prevent them from wandering into befuddling no-conventionality scenarios and cause them to take the immature but adaptive stance that any word they know will be known by everyone else.

It is important to note that this hypothesis hinges on assumptions about the manner in which children represent knowledge about word forms. One possibility is that the label is about the referent itself, and thus have a representation along the lines of “this is called a ‘dog.’” A second possibility is that children represent the source of the label along with the label, and thus have a representation along the lines of “Mom calls this a ‘dog.’” If children’s representations of word-referent links are more like the former, then the theory-of-mind limitations would truly prevent them from reasoning about no-conventionality scenarios. That is, if young children believed “this is called a ‘dog,’” they would have difficulty considering the possibility that others would believe otherwise. However, if children’s word-referent links also encoded notions of source, limitations on theory-of-mind might not prevent children from considering a no-conventionality scenario. That is, the knowledge that children would impute to other people would be that of how children’s families use a given word, which would not necessarily say anything about how others would use the same word.

We believe that there are two reasons to believe that children’s word-referent links are represented more in terms of what objects are called as opposed to what people call objects. First, labels seem to license certain inferences about objects (Gelman, 2005). One such inference has to do with essentialism: words map onto nonobvious, underlying causal features of objects (Gelman, 1988). A dog is a dog not because it necessarily looks like other dogs, but because it must share some internal, usually invisible, underlying causal factor that makes it a dog. Labeling an animal as a dog licenses this particular inference. The second inference is that of genericity: for the most part, words map onto abstract kinds rather than specific instances of a kind (Prasada, 2000). This inference is similar to the hierarchical constraint, whereby children believe that words extend to all instances of a particular kind and are not limited to a special particular instance (Markman, 1990). These inferences would be greatly aided if children represented words as being about objects rather than about what people call objects. If the latter were true, these concomitant inferences would also have to be marked with source (Mom believes this has the underlying dog essence; Mom believes that all things that have this essence should be
called dogs). Such sophisticated source marking skills seem as if they are outside the range of young children.

A second reason to believe that object names are about the objects themselves rather than what people call objects is that there are limitations on preschoolers’ abilities to encode information about the contexts within which knowledge is acquired (Perner, 1991). For instance, prior to the age of four, children are known to have poor memory for the sources of their own knowledge (Taylor, Esbensen, & Bennett, 1994). After learning a new fact, children cannot remember who taught them that new fact. If they learn a new fact themselves, they cannot remember how (that is, through which modality) they learned that fact (O’Neill & Chong, 2001). Given these difficulties, it seems unlikely that young word learners would be able to remember who taught them the meaning of a given word.

Given these considerations, it seems likely that children believe that words reflect knowledge about what objects are called rather than about what people call objects. And so it seems quite possible that theory-of-mind limitations might go some distance to rendering the early effects of conventionality.

We conclude by proposing that the conventionality effects enabled by theory-of-mind limitations may set the stage for the emergence of a true understanding of conventionality. Theory-of-mind limitations allow children to use language in accordance with conventionality, and because words are generally widely known, children are rarely in error. Thus, children gain considerable direct evidence that they can readily interpret a given word form from the language. Sometime over the third year, as children gain insight into epistemic mental states, their repeated successes might be redescribed as evidence that knowledge of the language is shared. This insight might be the building block of a true appreciation of conventionality.

References


MARK A. SABBAGH is associate professor of psychology at Queen’s University in Kingston, Ontario, Canada. He is interested in the social, cognitive, and neural bases of children’s language and social-cognitive development.

ANNETTE M. E. HENDERSON is a doctoral candidate at Queen’s University in Kingston, Ontario, Canada. She is interested in the cognitive and conceptual mechanisms that support word learning in young children and their origins.