The role of children’s explanatory construals in the structure of the lexicon: Cross-linguistic evidence from polysemy.

Mahesh Srinivasan¹* and Hugh Rabagliati²,³*

¹University of California, Berkeley
²Harvard University
³University of Edinburgh

* These authors contributed equally to the project.

Correspondence:

Mahesh Srinivasan
Department of Psychology
University of California, Berkeley
Berkeley, CA 94720-1650
Email: srinivasan@berkeley.edu

Or

Hugh Rabagliati
Department of Psychology
University of Edinburgh
Edinburgh, UK EH1 3NS
Email: hugh.rabagliati@ed.ac.uk
Abstract

We propose that the structure of polysemy across languages is a consequence of how children approach the task of word learning. Specifically, we propose that children have a set of expectations about how words can be used flexibly, which derive from universal and early-developing cognitive biases through which children construe and explain the world. These biases scaffold word learning, making some senses easier to acquire than others, for children of all linguistic communities. Over time, these easily-learnable senses propagate, creating similar patterns of senses across different languages. To test the theory’s predictions about cross-linguistic regularity and variation in polysemy, we conducted a large-scale survey, assessing whether 26 distinct patterns of polysemy found in English (e.g., animal for meat, material for artifact, etc.) were attested in 14 other languages. Strikingly, we found that almost all patterns of polysemy were present in other languages, although the specific senses that participated in those patterns (e.g., the artifact sense of glass) sometimes varied. We argue that, while this pattern of constrained variation is predicted by our theory, it is difficult to explain for theories that solely invoke pragmatic reasoning, on one hand, or language-specific conventions, on the other.

Word count: 194

Keywords:
Polysemy, pragmatics, cross-linguistic semantics, conceptual development, explanatory construals
All natural languages include a repertoire of words to express a large set of basic ideas, from concrete concepts of animals, objects, and materials, to more abstract notions like events and beliefs. Interestingly, however, rather than labeling each idea with a unique word, languages systematically lump multiple related ideas – or *senses* – under a single label, a phenomenon called polysemy (Breá, 1987). For instance, in English, the same words are often used to label an animal or its meat (e.g., *chicken, lamb*, etc.), or a material and an artifact derived from that material (e.g., *glass, tin*, etc.). Here, we ask what causes certain senses to be lumped together, by exploring cross-linguistic regularity and variation in polysemy.

We propose that the structure of polysemy is, in part, a consequence of how children approach the problem of learning mappings between words and concepts. In particular, children may sidestep the task of learning mappings one-by-one, and instead expect that words will label multiple concepts in systematic and constrained ways. Children’s expectations about word senses, we argue, are the product of a set of universal cognitive biases through which humans flexibly construe the world. Critically, if this hypothesis is correct, these cognitive biases should also constrain the nature of polysemy across languages, such that forms of polysemy that match these biases should be readily learnable and thus widespread cross-linguistically.

Below, we first describe previous proposals concerning constraints on polysemy, and how they are related to our own approach. Then, we review evidence suggesting that young children might learn word meanings using a set of cognitive biases through which they construe the world, and describe the surprising parallels between these biases and the nature of polysemy in English. Finally, we test our proposal by examining whether cross-linguistic variation in polysemy, assessed in a new large-scale survey, shows the signature marks of learning using these cognitive biases.

### 1.1 Constraints on polysemy

Theories of constraints on polysemy make a crucial assumption: that there are constraints on polysemy, and that senses are not randomly lumped together. In support of this assumption, linguists have identified a number of systematic patterns by which senses are organized (see, e.g., Copestake & Briscoe, 1995; Ostler & Atkins, 1992; Pustejovsky, 1995). Table 1 presents examples of some of the patterns that are found in English. As can be seen, these patterns often include senses that cross different semantic categories, alternately labeling people, animals, objects, substances, actions, and more. Some of these patterns invoke metaphorical relations, including when body part names are used to label parts of objects (e.g., “the chair’s arm is broken”). These patterns sometimes also include senses that cross grammatical categories, as when words are used to label objects and substances as nouns, and actions involving those objects as verbs (e.g., “He buttered the bread”, “She shoveled the snow”).

*Table 1. Examples of polysemy in English.*
Patterns and Participating Senses | Examples
--- | ---
Animal for Meat (chicken, turkey, fish, etc.) | The *chicken* walked on the grass / The *chicken* was well-salted
Material for Artifact (glass, tin, iron, etc.) | There is broken *glass* on the floor / She drank milk from the *glass*
Object for Representational Content (book, magazine, DVD, etc.) | The *book* is very light to carry / The *book* is very interesting
Container for Contents (pot, bowl, box, etc.) | She washed the *pot* after dinner / She stirred the *pot* with a spoon
Body Part for Object Part (leg, arm, back, etc.) | He broke his *leg* last year / That chair has a broken *leg*
Artist for Product (Picasso, Camus, Mozart, etc.) | *Picasso* was born in 1881 / That museum has a *Picasso*
Place for Institution (White House, Wall Street, City Hall, etc.) | The *White House* is being renovated / The *White House* should make a decision
Place for Event (Vietnam, Waterloo, Woodstock, etc.) | *Vietnam* shares a border with China / He championed civil rights during *Vietnam*
Substance for Placing Substance at Goal (butter, salt, water, etc.) | He bought some *butter* from the store / He is going to *butter* the bread
Instrument for Action Involving Instrument (shovel, hammer, rake, etc.) | She has a red *shovel* / She is going to *shovel* the snow

At least in English, however, these patterns of polysemy differ in important ways, and in particular, in how freely they permit generalizations. In a number of cases, patterns of polysemy can be easily extended to create new senses. The animal for meat pattern provides a good example of this. We can easily extend this pattern to label the meat of animals that aren’t typically thought of as edible. Thus, it sounds natural (though culinarily odd) to say “he ate some *seagull*.” However, not every pattern is so easily extendable. While *glass* and *tin* both describe materials and artifacts, it sounds distinctly odd to say “He bought a *plastic*,” even though we know that plastic is a material out of which many artifacts are made. These two types of patterns – generative and non-generative patterns – are typically referred to in the literature as *regular* and *irregular* polysemy, respectively (see e.g., Apresjan, 1974; Ostler & Atkins, 1992).

In sum, two general features characterize polysemy: senses are organized based on patterns, and some patterns are more generative than others. What kinds of mental representations and processes might account for these features of polysemy? Most work has focused on two potential sources of constraints. Some have looked at the role of pragmatics and conceptual structure, arguing that attested senses of polysemous words do not need to be learned, but are instead derived from context, and reflect how we reason about the world around us (e.g., Fauconnier, 1985; Nunberg, 1979, 1995; Papafragou, 1996; Wilson, 2003). Others have focused on the development of linguistic conventions, arguing that the senses of polysemous words are conventions that must
be learned and stored in memory by members of a linguistic community (e.g., Klein & Murphy, 2001; Lehrer, 1990; Murphy, 1997, 2007; Pinker, 2007). We discuss each of these ideas below, and how they might account for the features of polysemy described above: that senses are organized into patterns, which have differing levels of generativity. To our knowledge, no theory solely ascribes the structure of polysemy to either pragmatics or conventions, but theories do differ in where they place their emphasis.

Under proposals relying more on pragmatics, speakers do not store all of the individual senses of polysemous words in memory, but instead derive these senses on-line, from a single represented meaning. Ruhl (1989) provides perhaps the most extreme example of this idea, arguing that most polysemous words actually only have a single meaning that captures the essence of the concept, and that can be pragmatically adjusted to suit the surrounding context. For example, a core meaning of glass could denote a material that can be used to form solid objects, and context would then be used to fill in the details, such as whether the word is being used to label the material itself (as in “He bought a sheet of glass”), or instead an object composed of that material (as in “He poured water into the glass.”).

In order to make contextual adjustments to core meanings, these theories assume that listeners and speakers reason pragmatically, based on their knowledge of the intentions of interlocutors, as well as their more general knowledge of the world. The structure of polysemy is therefore, in part, a function of the structure of concepts. For example, Nunberg (1995) proposes that senses can be shifted based on a principle of noteworthiness: when context sets up a noteworthy relationship between a core meaning and another possible sense, that sense becomes plausible.¹ For example, the material composition of a glass drinking vessel may be a noteworthy aspect of glasses, allowing the word glass to label both the material and artifact. Critically, then, language users do not need to learn the different ways that glass can be used. Instead, these different senses follow naturally from conceptual relations like noteworthiness.

Under pragmatic theories, conceptual structure accounts not only for how senses are derived from core meanings, but also for the fact that polysemy is organized into patterns, and that these patterns are generative. For example, under Nunberg’s theory, patterns arise when the same noteworthy relationship is seen over and over again in the world. Just as the material composition of glasses may be noteworthy and explain the flexibility of glass, the material composition of tins, sponges, and irons may also be noteworthy and explain the systematic alternation of tin, sponge, and iron. Generativity also follows naturally from pragmatic theories, because a new sense of a word can always be derived, so long as it stands in a noteworthy relationship to a core meaning of

¹ To explain how senses are derived, other, similar theories appeal to factors such as relevance (Papafragou, 1996; Falkum, 2010; Wilson, 2003), idealized cognitive models (Fauconnier, 1985; Lakoff, 1987), and cue validity (Nunberg, 1979).
the word. That is, when a relationship is noteworthy in a novel context – e.g., that some meat being served is from a seagull – a novel sense can be coined (e.g., “That seagull could use some salt”).

However, given that there are many noteworthy relationships between concepts, a potential difficulty for pragmatic theories is that they may be too generative to account for the relatively constrained nature of polysemy. In particular, while pragmatic theories can explain the senses we do use, they have trouble explaining why we do not use many other senses. For instance, under pragmatic theories, it is difficult to explain why a word like glass can only label one kind of artifact (i.e., drinking vessels), as opposed to the myriad other artifacts that are also noteworthy for being composed of glass material (e.g., mirrors, windows, etc.). To take another example, although English permits the use of names for animals to label their meats, it does not allow animal names to label products derived from those animals. Thus, although it is certainly noteworthy that eggs are laid by chickens, chicken cannot label an egg. These facts suggest that a theory that appeals only to conceptual structure and pragmatic reasoning may be too unconstrained to explain how polysemy is actually realized.

Citing the apparent arbitrariness of polysemy, a number of accounts have proposed that the majority of word senses are not derived via conceptual structure, but are instead conventions that members of a linguistic community must learn, one-by-one (Lehrer 1990; Murphy, 1997, 2007; Pinker, 2007). These theories propose that each sense is initially coined by an individual speaker and learned by individual listeners. When these conventions are useful, they become more frequent and stable in the language. Conceptual structure therefore plays a much more limited role in conventionalization accounts than it does in pragmatic accounts. In particular, concepts are thought to provide only weak constraints on how senses are coined: speakers and listeners have to grasp the relationship between the new and old sense, but otherwise, these senses could be related in any number of ways. For example, by this account, there is no principled reason as to why glass labels drinking vessels, as opposed to windows, or why chicken labels chicken meat but not an egg: these are merely facts about language that speakers must master.

Conventionalization theories also do a good job of explaining how senses can drift apart from one another, semantically. In particular, once a new sense for a word has been coined and has entered the language, its relationship to other senses does not need to remain transparent. This means that senses can remain in a language long after the initial communicative motivation that created them has died away. An example of this is the English use of iron to describe a tool for pressing clothes. When the sense was coined, clothes were mainly pressed using large pieces of flattened iron (i.e., flatirons). However, that technology is now obsolete, and iron can now be used to label pressing machines that do not contain any iron. Similarly, it is possible for glasses to be made of plastic, to land on water, and to shelf books on a windowsill (see Clark & Clark, 1979; Kiparsky, 1997).
Critically, however, while conventionalization accounts do a good job of explaining arbitrary properties of polysemy, and the presence of fossilized senses like *iron*, they do not naturally account for the two features of polysemy that we laid out earlier. In particular, the presence of patterns, that incorporate multiple participating words, does not naturally follow from a theory in which there are only weak conceptual constraints on the senses that can be formed. Further, the presence of generativity is unexpected, if senses have to be individually stored in memory, and cannot be derived. In response to these points, Murphy (2007) has speculated that patterns emerge when senses are coined via analogy. For example, the inspiration for labeling chicken meat using its animal name may come from comparisons to other uses of animal names, like *fish* or *lamb*. Additionally, after multiple related senses have entered a language, speakers may form *lexical rules* that characterize existing patterns and that are generative, allowing language users to produce and understand novel uses of words along the same patterns (e.g., such that new animal names can label meat; see also Copestake & Briscoe, 1995; Strigin, 1998). Thus, by this account, generative patterns of polysemy arise over the history of a language. Patterns may vary in generativity, however, to the extent that it is more or less difficult for language users to abstract lexical rules that characterize those patterns. Thus, it may be relatively difficult to draw analogies between the material and artifact uses of terms like *glass*, *tin*, and *iron*, such that a generative lexical rule cannot be formed for the pattern.

Our approach draws on the ideas of these two different theories, but also introduces a novel element. Whereas pragmatic accounts argue that polysemy reflects conceptual structure, and whereas conventionalization accounts argue that polysemy reflects facts about language that must be learned, we propose something different: polysemy has the structure it does because it has evolved to fit children’s early cognitive biases, and thus, to be easy for children to learn. Our theory is similar to conventionalization accounts, in that children have to learn individual senses, and cannot simply derive them based on salient or noteworthy relations. But, reminiscent of pragmatic accounts, we also argue that the ease with which children learn senses depends on how well these senses correspond with children’s early conceptual structure, which places constraints on the kinds of senses that will emerge as a language evolves. The combination of these two processes results in a set of predictions about cross-linguistic variation in polysemy that differ importantly from those of pragmatic and conventionalization theories. We describe these predictions in section 1.4, and test them using a large-scale cross-linguistic survey. But first, we present our proposal in detail.

### 1.2. Polysemy as a learning tool

By adulthood, humans can use words to express an extraordinary range of basic concepts. If each of these basic concepts had to be represented by a separate word, then building such a vast lexicon would be a slow and difficult process for children. This
is because mappings between word forms and concepts are arbitrary (Saussure, 1959), and so each mapping has to be learned on its own. We propose that polysemy arises as a way of speeding up this learning process. In particular, after learning one meaning for a word, children should be well equipped to guess its other senses.

The proposal that polysemy aids learning might initially seem surprising: shouldn’t it be confusing to learn a language whose words conflate different concepts? On the contrary, we suggest that, because the senses of polysemous words are related, learning one sense of a word provides a clue for learning its other senses. Imagine, for example, a parent who would like their child to pick up a glass from the dining table. The child has not yet learned the name of this object, but has learned that glass can label a kind of transparent material. If the parent asks the child to “Pick up the glass”, the child may reason that the referent of glass may be related in some way to the transparent material, and may constrain her hypotheses on this basis. However, without such polysemy, the parent would have to use a different word for the drinking vessel, e.g., “pick up the dax”, which could potentially refer to any of the objects on the table. This suggests that learning multiple senses should be much easier than learning multiple words.

In addition to providing children with clues about what new senses of words mean, polysemy may also allow children to spontaneously infer these senses. Under our theory, children may have to learn the senses of some polysemous words one-by-one (as is argued by conventionalization accounts of polysemy). However, when many of these senses follow a similar pattern, children could make a higher order generalization, that would allow them to predict and infer senses for subsequently learned words. In particular, generalizations could be made for systematic patterns of polysemy, in which senses alternate in similar ways. For example, a child could learn that chicken can label an animal and its meat, and that words like lamb and fish alternate in the same way. Having noticed this systematicity, children would be in a position to make a higher-order generalization, that words can label both animals and their meat. Such a generalization would greatly simplify the process of learning word meanings, because children would only need to learn one label to express two concepts. For example, having learned that seagull labels an animal, children could infer that this word can also label the derived meat, even in the absence of ostensive evidence that it can.

In sum, we suggest that polysemy simplifies the process of word learning by providing children with clues about new senses of words, and even by allowing children to spontaneously infer these senses. This proposal assumes that children are not confused by the systematic flexibility presented by polysemy, but instead find it intuitive. In the following section we review evidence that from early in life, children conceptualize the world in a way that allows them to not only understand the conceptual relations between polysemous senses, but also to expect words to alternate along these relations. Importantly, this suggests that the structure of polysemy is actively shaped by children’s expectations about the flexibility of word meaning. The next section provides evidence
for these points, while section 1.4 lays out what this might mean for the structure of polysemy across languages.

1.3. Polysemy and conceptual development.

Consistent with our proposal that children find polysemy intuitive, there is good evidence that, from early in development, children are able to learn multiple senses for words, and even expect words to be used flexibly. By at least age four, children are sensitive to the relations between the senses of polysemous words (e.g., the use of book to label an object or its abstract content), and distinguish them from unrelated homophones (e.g., the use of bat to label an animal or baseball equipment; Srinivasan & Snedeker, 2011, in press). They also show a sophisticated ability to determine the correct meaning for an ambiguous word in context (Rabagliati, Pylkkanen, & Marcus, 2013). Moreover, even before age four, children appear to expect words to be used in innovative ways. For example, children creatively use words for space to describe time (“Mommy, can I have some reading behind dinner”; Bowerman, 1983), words for instruments to describe actions involving those instruments (“Don’t broom my mess”; Clark, 1982), and words for abstract content to describe objects (e.g., agreeing that a movie can be round; Rabagliati, Marcus & Pylkkanen, 2010). Importantly, these innovations are related to attested polysemous senses (e.g., the use of broom is similar to attested uses of hammer and shovel), suggesting that children have formed higher-order generalizations, allowing them to coin novel senses. Consistent with this, recent evidence indicates that four- and five-year-olds spontaneously expect new words to be used flexibly according to existing patterns (e.g., to label instruments and functional uses of those instruments, see Srinivasan & Barner, 2013a).

Together, the findings described above confirm that children can guess and learn new senses. But what drives these abilities? For children to guess the new senses of a polysemous word, they first have to understand how the concepts labeled by that word are both related and distinct. For instance, to infer the different senses of glass, children have to understand that a drinking glass is made from glass material. Further, to use these senses correctly, children have to understand how they differ from one another. For example, when glass is used in its material sense, it can label any entity that is composed of glass, irrespective of its particular form or function, but when glass is used in its artifact sense, it labels an entity with a specific form, that was created with a particular function in mind (Bloom, 1996; Malt, in press; Malt & Johnson, 1992). Thus, to learn the different senses of polysemous words like glass, children need to be able to flexibly conceptualize entities in multiple ways, differentially focusing on properties such as form, function, material composition, and origin.

Under a number of theories of conceptual development, this sort of flexibility may be something that children have to learn. For example, empiricist and constructivist theories, in which children form concepts by tracking frequencies and correlations amongst the observable features of entities (e.g., Rogers & McClelland, 2004; Rosch &
Mervis, 1975; Smith, 2001), predict that children have to discover abstract notions, such as intended function, that underlie polysemy. Discovering these notions is not easy, because they are not readily observable. This poses a potential problem for our proposal that children find polysemy intuitive, because these abstract features are critical for distinguishing between polysemous senses (e.g., between the material and artifact senses of glass). Thus, if children cannot construe the world flexibly from the earliest stages of conceptual development, polysemy should pose a hurdle for them, rather than facilitating their development of a lexicon.

However, an alternative approach to conceptual development, which has received much empirical support in recent years, starts with the idea that children are flexible from the start, and are driven to seek out multiple abstract explanations for why entities have the properties they do (a proposal whose intellectual heritage can be traced back to Aristotle; see Moravscik, 1975, 1981). By this account, children do not blindly keep track of frequencies and correlations among properties, but are instead guided by larger frameworks, or intuitive theories, through which they understand the world (e.g., Carey, 1985; Gelman & Wellman, 1991; Gopnik & Meltzoff, 1997; Keil, 1989, 1994a; Leslie, 1994; Murphy & Medin, 1985; Prasada, 2000). These frameworks may help children understand how an entity’s properties might be related to one another, and how this depends on the domain an entity belongs to – e.g., whether it is an artifact, substance, natural kind, and so on. Thus, guided by an intuitive theory of artifacts, children could understand that an object, like a drinking glass, has the function it does due to its process of creation (e.g., it may hold liquid because it was designed to do so), could understand its form by invoking its function (e.g., it may be hollow so that it can hold liquid), and could understand some of its other properties by invoking its material (e.g., it may be transparent and fragile because it is made of glass). An intuitive theory of substances, in contrast, may specify different relations between these explanatory factors: e.g., to understand why a pile of sand has the properties it does, it may not make sense to appeal to its intended function, but instead to its material composition.

Critically, these explanatory frameworks could shape not only how children make sense of entities in the world, but also how they expect words to describe the world. In particular, the application of different intuitive theories would allow children to construe even the same entity in different ways, and this flexibility should allow children to learn the different senses of polysemous words. For example, knowing that a cylindrical glass object was designed to hold liquid, children could conceptualize it as an artifact and learn this sense of the word glass. But if they thought it was created by accident and had no designed function, they could conceptualize it as a substance, and thus learn the material sense of glass.

But what are the grounds for believing that children are conceptually flexible, and that this helps them learn word meanings? While a full review is beyond the scope of this paper, the evidence suggests that from early in life, children can readily shift their perspectives of an entity, and can do so by attending to the explanatory factors we have discussed above, including function, process of creation, form, and material composition.
To begin, young children are attuned to the functions of objects. When children as young as two-and-a-half years are exposed to a new object and shown how it can be used, they appear to assume that the object cannot be used to perform another function, and thus, is exclusively for a certain purpose (Casler & Kelemen, 2005). Consistent with this sensitivity to function, several studies have indicated that preschool-aged children prefer to extend novel nouns between objects that are used for a common purpose (e.g., Kemler Nelson, 1995), rather than between objects that are merely perceptually-similar (but see Landau, Smith & Jones, 1988; for discussion, see Bloom, 2000).

However, children’s focus on function is not absolute: they can use information about how an entity was created to shift between different perspectives of that entity. For example, Gelman & Bloom (2000) showed that three-year-olds can construe the same object in terms of its intended function as an artifact, or instead as a material, depending on the information they are provided. Specifically, children will label a sharp piece of plastic as a knife if told that it was intentionally shaped that way, but will label it as plastic if told that it was created accidentally. This is exactly the kind of flexibility that would help children learn the artifact and material senses of polysemous words like glass.

Similar flexibility has also been observed in other domains, and may arise even before children acquire relevant aspects of language. For example, even before children acquire a syntactic mass-count distinction, they are able to alternately focus on the forms of objects, and on their material compositions. When a substance is non-solid, children treat material composition as more central to its kind membership than shape, and vice versa for solid substances (Soja, Carey, & Spelke, 1991). Prior to learning language, children are also able to alternately construe the same physical entity as an inanimate object or intentional agent. For instance, infants as young as 12 months will follow the “gaze” of a faceless object, if it interacts with them in a contingent, “animate” way (Johnson, Slaughter & Carey, 1998).

Together, these studies and others suggest that children’s acquisition of knowledge about the world is guided by a set of theories that allow them to flexibly construe entities in different ways. This raises the possibility that these same theories could also guide children’s early expectations about the flexible meanings of words. Interestingly, while this proposal has not previously been explored, independent support for it comes from linguistic theories of the mature lexical-semantic structures that children eventually develop by adulthood. Several scholars have proposed that lexical items include explanatory schemes that give rise to and constrain a variety of linguistic phenomena, including polysemy (Keil, 1994b; Moravscik, 1981, 1990; Pustejovsky, 1995; Prasada, 2000). These schemes are similar to the theories proposed in developmental work, in that they also appeal to explanatory factors like form, function, material composition, and origin. For example, the systematic alternation of words like glass between materials and products made from those materials, and of words like chicken between
animals and meat, may stem from an ability to view an entity both as an object with a particular form (and intended function, in the case of a glass), and as something composed of a particular material (Moravcsik, 1981). Other forms of polysemy, such as the use of words like *hammer* to label objects (as nouns) and actions supported by those objects (as verbs), or the use of words like *book* to label objects and their informational content, may depend on an ability to construe an entity both as an object and in terms of the function of that object. Function may also be critical to explaining semantic type coercion phenomena: “enjoy the book” may mean “enjoy reading the book”, but “enjoy the rock” may seem poorly-formed, because we tend to attribute functions to books, but not rocks (Pustejovsky, 1995).

The overlap between theories of how children construe the world and theories of lexical semantics provides a tantalizing hint that children’s early conceptual representations may constrain their hypotheses about the senses of polysemous words. If so, children’s conceptual representations could also constrain the structure of polysemy across languages. In particular, if children learn senses by invoking different explanatory factors, then they will find it easier to learn senses that correspond to the application of these modes of explanation. Therefore, as a language changes and evolves over generations of learners, it may develop patterns of easily learnable senses that correspond to these modes of construal. These would include polysemous alternations such as material for product, animal for meat, and complement coercion; exactly those forms of polysemy that lexical semanticists have developed models to capture.

In sum, our theory provides an intriguing account of how the development of polysemy in children might explain the structure of polysemy in a language. However, the data as of now are not conclusive, because most lexical semantic theories of polysemy have only focused on the polysemy found in English. If polysemy is truly constrained by early-developing cognitive biases, these biases should also constrain polysemy in other languages. The remainder of the paper evaluates this proposal. We lay out what different approaches to the structure of polysemy – including our own – predict about its cross-linguistic variation, and then test these predictions using a new, large-scale cross-linguistic survey.

### 1.4. Cross-linguistic predictions of different accounts of polysemy

Thus far, we have described previous accounts of the structure of polysemy – including pragmatic and conventionalization accounts – and also introduced our own proposal, that polysemy is constrained by early-developing cognitive biases that shape how children learn senses. In this section, we outline the predictions that each of these three theories make about cross-linguistic variation in polysemy.

If the senses of polysemous words are not learned conventions, but are instead contextually derived via pragmatic reasoning and conceptual structure, we would expect a great deal of uniformity in how polysemy is expressed across different languages. In
particular, because speakers of different languages are likely to share broadly similar conceptual repertoires, they should generally find the same relationships to be salient or noteworthy, and should thus derive senses from core meanings in similar ways. This would predict that the same patterns of polysemy should be present in different languages (e.g., animal for meat, material for product, etc.). It would also predict that the same senses should participate in these patterns across languages, because these senses all stand in equally noteworthy relationships to their core meanings. For example, the relations between glass material and a drinking vessel, and between tin material and a cookie tin, are similarly noteworthy, so these senses of glass and tin should be equally possible across languages. Thus, if the senses of glass are possible in a language, so should be the senses of tin and iron. Of course, due to cultural differences, speakers of different communities may find different relationships noteworthy – and thus may use different senses. For instance, cultures that do not use glass vessels to drink from would not have that particular sense of glass. But, by and large, pragmatic theories predict a great degree of uniformity with respect to both senses and patterns across languages.

By contrast, conventionalization theories of polysemy predict that language- and culture-specific variation in patterns and senses should be the norm, not the exception. According to these theories, conceptual structure plays only a limited role in how senses are coined and learned. Thus, languages should be free to develop senses in unique ways, to address their communicative needs. For instance, in some languages, it may have been communicatively beneficial to create a convention by which glass can label a drinking vessel, but this might not have been true of other languages. Further, if patterns of polysemy arise as speakers draw analogies between existing senses to coin new ones, languages may vary greatly in the patterns they include. In particular, the first few senses coined – which could vary dramatically by language – could determine which patterns arise as the language evolves. Of course, conventionalization accounts need not predict incommensurability with respect to polysemy across languages, because different cultures may have similar communicative needs. But in general, these accounts predict a great degree of variation with respect to both senses and patterns across languages.

Our own account carves out a middle ground, and predicts that while patterns of polysemy should be robust across languages, the senses that participate in these patterns may vary (see Figure 1). Our theory is similar to a conventionalization theory in that children must learn senses as conventions, and cannot simply derive them based on factors like noteworthiness or salience. Critically, however, we also propose that the senses that children learn are guided by cognitive modes of construal that children of all linguistic communities share. Given this, children of different cultures and language groups should be biased to learn senses in very similar ways. Therefore, as languages evolve, they should develop similar patterns, containing easily learnable senses that are adapted to these modes of construal. Patterns should only be absent from a language when they have not had time to evolve, or when other forms in the language block them
(e.g., morphemes or compounds that express the same alternations in meaning – which may also have developed to match children’s cognitive biases).

Figure 1. How modes of construal might constrain the structure of polysemy over the course of language development and change.

However, because modes of construal may in some cases provide only loose constraints on learning, our theory predicts that the actual senses that participate in patterns should vary across different languages (see Figure 1). In particular, when patterns are relatively unpredictable, languages may differ with respect to the specific senses that are coined and learned as conventions. Pattern predictability describes the degree to which we can predict one sense from another, given a particular pattern. For example, knowing that a dax is a type of animal, we can be confident in predicting that the application of the animal-meat pattern will result in a sense that describes some meat. By contrast, knowing that a dax is a material of some kind, it would be hard to predict the output of a material-artifact pattern with any degree of accuracy, as the range of possible artifacts could be very large. Thus, across languages, pattern predictability should have an impact on how senses are coined. While senses described by predictable patterns are likely to be robust across languages, senses described by less predictable patterns are likely to be less robust across languages, because there
will be many possible senses that could fit the pattern. For instance, when labeling an artifact, *glass* might not label a drinking vessel, but instead a window, a mirror, or other artifacts made of glass. Because each of these senses would be similarly learnable by children (given their modes of construal), they should each stand a similar chance of emerging in a language, allowing different languages to develop different senses.

How might we test this hypothesis, that pattern predictability determines cross-linguistic variability in senses? Unfortunately, measuring pattern predictability is difficult, because it is a description of the world and not of people’s intuitions. For example, measuring pattern predictability would require us to count the number of things made of glass, or sponge, or iron, etc., across cultures. But our theory can still be tested because pattern predictability may also play another role, in determining when children are able to make higher-order generalizations (Figure 1). As discussed in Section 1.2, when children realize that multiple words, like *chicken*, *lamb*, *fish*, are all polysemous in the same way, they could make a generalization, e.g., that all names for animals can describe meats. This generalization could allow them to form generative patterns of polysemy, and thus coin new senses, such as “tasty *seagull*”. Making these generalizations will be simpler when a pattern is predictable, because the surface relations between senses will be straightforwardly comparable (e.g., the meats of different animals appear similar to one another and are used for similar purposes). By contrast, when a pattern is not predictable, the surface correspondences between pairs of senses will be more difficult to align (e.g., a glass, a tin, and an iron are all made of different materials and are used for different purposes), and so it will be harder to coin a single generalization that can capture how all pairs of senses are related. Consequently, it may be more efficient to store each sense as a separate lexical item (for evidence consistent with this distinction, see Rabagliati & Snedeker, 2013).²

The above considerations therefore suggest that pattern predictability may determine both whether a pattern’s participating senses are variable across languages and whether the pattern is generative. More generally, relative to pragmatic and conventionalization theories, our theory generates three unique predictions about the nature of polysemy across languages:

1. Patterns of polysemy should be robust across languages.

2. For some patterns, the same participating senses should recur across languages, while for other patterns, participating senses should be more variable.

² Explanations like this are quite prominent in the study of language development, and have been applied to explain phenomena such as the acquisition of past-tense inflection, see Marcus et al., 1992; Pinker, 1991; for a recent example, see O’Donnell, Snedeker, Goodman & Tenenbaum, 2011).
3. Patterns whose senses show the least variability across languages should also be the most generative.

1.5. Existing cross-linguistic data on polysemy

To evaluate the above predictions, we would need to assess a large number of patterns of polysemy across a large set of languages, with many examples per pattern. However, to our knowledge, such a survey has not yet been conducted. As we review below, studies that have explored a large set of languages have often focused on just one pattern of polysemy, while studies that have assessed a large set of patterns have often focused on a limited set of languages, or have probed only a small number of examples for each pattern.

For example, Boyeldieu (2008) provides a cross-linguistic analysis of the use of the word *animal* to mean *meat*. Greenberg (1983) was the first to note that many of the Niger-Congo languages of west and southern Africa collapse the meanings *animal* and *meat* into a single word. The same conflation is also found in two Tibeto-Burman languages (Matisoff, 1978, reported in Boyeldieu, 2008) and in Warlpiri, where the same word is used for meat and edible animals. Boyeldieu argues for a conventionalized, culture-driven explanation for why this polysemy emerged. But because this claim is based on only the flexibility of a single word and a limited number of languages, it provides a weak foundation for drawing general conclusions about the nature of polysemy across languages.

Studies conducted on a broader scale have found evidence for both cross-linguistic regularities in patterns of polysemy, as well as variation in the particular words that participate in those patterns, consistent with our proposal. Perhaps the best known of these studies is Viberg’s (1984) survey of perception verbs, which was conducted across 52 languages. Based on his data, Viberg proposed a hierarchy for characterizing how, across languages, verbs for one sensory modality can be extended to describe another. For example, verbs that originally described *seeing* were often extended to describe *hearing*, but not vice versa, and in turn, verbs that meant *hearing* were sometimes further extended to describe *smelling*, but again, not the reverse. This hierarchy can be thought of as a skeleton that permits languages to extend verb senses in certain ways, but not others. However, extensions are not required, allowing for cross-linguistic variation.

Away from verbs of perception, Seto (1996, referenced in Peters, 2003) conducted a study exploring cross-linguistic variation in the container-contents pattern of polysemy, which is evidenced by the use of the English word *kettle* to label a container (*cast-iron kettle*), as well as the contents of that container (*boiling kettle*). Seto found that the pattern was attested in a wide set of languages, including Korean, Mongolian, Javanese, Italian and English, providing evidence for a potential universal. This follows from our
Polysemy Across Languages

proposal, because this pattern is highly predictable (e.g., all containers should have contents).

Although the studies of Viberg (1984) and Seto (1996) explored a large set of languages and assessed the flexibility of multiple words, they each focused on only a single pattern of polysemy, and so their findings may not generalize to other patterns of polysemy. To resolve this issue, some researchers have conducted broader surveys, assessing multiple patterns of polysemy across languages. For instance, Kamei & Wakao (1992) found differences in how English, Mandarin Chinese and Japanese speakers rated sentences that exhibited several different patterns of polysemy. The results provided interesting evidence of cross-linguistic variability. Thus, for example, Mandarin speakers judged that the use of a word for a producer to describe their product, as in John read Mao, was unacceptable.

For our theory, the findings of Kamei & Wakao (1992) are surprising, because the producer-product pattern is quite predictable and should thus be robust cross-linguistically. For example, knowing that Mao labels a producer should constrain the product sense of Mao to a limited set of referents in the world – i.e., to those works created by Mao. However, it is not clear whether the observations from Kamei and Wakao can be generalized beyond the examples they tested for each pattern of polysemy. For instance, subjects’ responses could have been affected by cultural norms involving the specific tested examples: e.g., Chinese individuals may not want to refer to Mao in careless ways. Responses to specific examples could also have been exceptions to larger patterns. Thus, a survey that asked English speakers to rate “The man ate pig” or “The man ate cow” might erroneously conclude that English does not use the same labels for animals and the meat derived from them. But in fact, this pattern is common English. These exceptions can be explained through the presence of synonymous terms, like beef and pork, which block the regular pattern in the same way that went blocks the regular past-tense form of go.

These considerations suggest that it is critical not only to assess a large number of patterns of polysemy across languages, but also to probe a large set of examples for each pattern. Peters (2003) attempted to do this by comparing polysemous senses in English, Dutch and Spanish, using a large dataset: the cross-linguistic thesaurus known as EuroWordNet (Vossen, 1998). Peters reported that a number of patterns of polysemy (e.g., plant for food) are shared across the languages, though not every polysemous word found in English is also attested in Dutch or Spanish. These findings match the predictions of our theory, which similarly predicts that, while patterns of polysemy should be cross-linguistically robust, individual senses may be more variable. However, Peters’ findings are limited, in that his data were confined to three languages. Additionally, it is possible that Peters’ findings resulted from the structure of EuroWordNet. Specifically, the English thesaurus had by far the broadest coverage, raising the possibility that many extant English senses simply weren’t listed in the less extensive thesauri of the other languages.
In summary, although previous cross-linguistic studies have provided intriguing data, they have not yielded strong conclusions about the nature of polysemy across languages. In particular, on the basis of extant data, it is difficult to evaluate the claims of our theory, that while polysemy patterns are robust across languages, individual senses are variable across languages, and that this follows from the predictability of different patterns. Thus, to test these predictions, we conducted a new assessment of how polysemy varies across languages, which differed from previous surveys in several critical ways. First, in contrast to previous studies, which focused on testing a limited number of patterns of polysemy across a large set of languages or vice versa, our study assessed both a large set of languages (15 including English) and a large set of polysemy patterns (27 patterns found in English). Second, to provide a rigorous test of whether each language included a particular pattern of polysemy, we tested a large number of examples for each pattern (from three to seventeen), including not only examples of participating senses in English (e.g., chicken and lamb for the animal for meat pattern), but also examples of exceptions in English (e.g., cow and pig), which might not constitute exceptions in other languages. We also allowed participants to report examples that we did not ask about, and supplemented these data by consulting dictionaries. Third, to explore the potential role of pattern predictability in explaining cross-linguistic variation in senses, we probed participants’ judgments about the generativity of each pattern.

2 Methods

2.1 Participants
Our participants were native speakers of languages other than English, but who could all read, write, and speak English to a reasonable degree of fluency. 36 participants, drawn from 14 different languages, completed the survey (for details about these participants, see Table 2). For most languages, we collected responses from multiple participants. In general, we recruited additional participants for a language if our existing participants lacked a background in linguistics. Thus, although we only had 1 participant each for Hungarian, Italian, and Turkish, we were confident in their judgments because they had received training in linguistics. However, we excluded 3 participants (speakers of Arabic, Sindhi and Marathi), because they lacked training in linguistics, and because we were unable to get additional speakers to confirm the accuracy of their judgments. Finally, four native English speakers also participated, to validate our judgments about polysemy in English. Participants were drawn from the communities around Harvard University and the University of California, San Diego, and included students as well as full-time researchers.

Table 2. Background information about our participants and references to consulted dictionaries

<table>
<thead>
<tr>
<th>Native Language</th>
<th>Number of participants</th>
<th>Participants with</th>
<th>Average age</th>
<th>Average number of</th>
<th>Dictionary analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Language</td>
<td>Linguistics Background</td>
<td>Began Learning English</td>
<td>Years Speaking Fluent English</td>
<td>Source</td>
<td></td>
</tr>
<tr>
<td>----------</td>
<td>------------------------</td>
<td>------------------------</td>
<td>-----------------------------</td>
<td>--------</td>
<td></td>
</tr>
<tr>
<td>English</td>
<td>4</td>
<td>1</td>
<td>0</td>
<td>19</td>
<td>NA</td>
</tr>
<tr>
<td>Cantonese</td>
<td>3</td>
<td>0</td>
<td>4</td>
<td>10</td>
<td>dict.youdao.com</td>
</tr>
<tr>
<td>Farsi</td>
<td>3</td>
<td>1</td>
<td>2</td>
<td>19</td>
<td>NA</td>
</tr>
<tr>
<td>French</td>
<td>1</td>
<td>1</td>
<td>14</td>
<td>4</td>
<td>(Harrap's, 2001)</td>
</tr>
<tr>
<td>Hindi</td>
<td>3</td>
<td>1</td>
<td>5</td>
<td>15</td>
<td><a href="http://www.hinkhoj.com/">http://www.hinkhoj.com/</a></td>
</tr>
<tr>
<td>Hungarian</td>
<td>1</td>
<td>1</td>
<td>11</td>
<td>7</td>
<td>(Országh, Futász, &amp; Kövecses, 1998)</td>
</tr>
<tr>
<td>Indonesian</td>
<td>2</td>
<td>0</td>
<td>8</td>
<td>8</td>
<td>(Stevens &amp; Schmidgall Tellings, 2004)</td>
</tr>
<tr>
<td>Italian</td>
<td>1</td>
<td>1</td>
<td>5</td>
<td>16</td>
<td>(Reynolds, 1981)</td>
</tr>
<tr>
<td>Japanese</td>
<td>2</td>
<td>1</td>
<td>6</td>
<td>14</td>
<td><a href="http://www.css">http://www.css</a> e.monash.edu.au/~jwb/cgi-bin/wwwjdic.cgi?1C</td>
</tr>
<tr>
<td>Korean</td>
<td>2</td>
<td>0</td>
<td>5</td>
<td>16</td>
<td>(Martin, Lee, &amp; Chang, 1967)</td>
</tr>
<tr>
<td>Mandarin</td>
<td>3</td>
<td>2</td>
<td>7</td>
<td>8</td>
<td>dict.youdao.com</td>
</tr>
<tr>
<td>Russian</td>
<td>2</td>
<td>2</td>
<td>5</td>
<td>16</td>
<td><a href="http://en.bab.la/dictionary/english-russian/">http://en.bab.la/dictionary/english-russian/</a></td>
</tr>
<tr>
<td>Spanish</td>
<td>3</td>
<td>0</td>
<td>6</td>
<td>13</td>
<td>(Galimberti Jarman, Russell, Rollin, &amp; Carvajal, 2008; Velázquez de</td>
</tr>
</tbody>
</table>
2.2 Materials

We chose 27 patterns of polysemy in English that had previously been identified in the literature (see, e.g., Clark & Clark, 1979; Copestake & Briscoe, 1995; Ostler & Atkins, 1992; Pustejovsky, 1995). For each pattern, we identified multiple examples of English senses that followed each pattern, and in some cases, we identified examples of exceptions to these patterns. Thus, for the animal-meat pattern, we included not only participating examples like chicken and lamb, but also exceptions like cow and pig. As we describe below, we first asked participants to translate the base sense of an example (e.g., the animal meaning of chicken), before asking them to judge whether that word also had an extended sense (e.g., the meat meaning of chicken). We drew on previous accounts, as well as our own intuitions, to decide the base meanings for each pattern.

Table 3 lists the patterns and words assessed in our survey, along with the shorthand notation for each pattern, which we use in our graphs. Place for event polysemy was added after the survey was completed by the Italian and Spanish speakers, and so it is excluded from some of our analyses, as indicated below.

Table 3. The patterns, participating words, and blocked words assessed in the survey

<table>
<thead>
<tr>
<th>Pattern Used (Base sense listed first)</th>
<th>Short Name</th>
<th>Participating Words (Starred words were used on the list task)</th>
<th>Examples (Used in the survey)</th>
<th>Blocked words</th>
</tr>
</thead>
<tbody>
<tr>
<td>Animal for fur derived from animal</td>
<td>AnFur</td>
<td>Mink, chinchilla, rabbit, beaver, raccoon*, alpaca*, crocodile*</td>
<td>The mink drank some water / She likes to wear mink</td>
<td>Sheep, cow, goose, elephant, oyster</td>
</tr>
<tr>
<td>Animal (or object) for personality property</td>
<td>AnPro</td>
<td>Chicken, sheep, pig, snake, star*, rat*, doll*</td>
<td>The chicken drank some water / He is a chicken</td>
<td></td>
</tr>
<tr>
<td>Animal for meat</td>
<td>AnMe</td>
<td>Chicken, lamb,</td>
<td>The chicken drank</td>
<td>Cow, pig,</td>
</tr>
<tr>
<td>Derived from animal</td>
<td>Artifact for activity involving artifact</td>
<td>Body part for object part</td>
<td>Building for people in the building</td>
<td>Predicate for predicate with verbal complement (complement coercion)</td>
</tr>
<tr>
<td>---------------------</td>
<td>----------------------------------------</td>
<td>---------------------------</td>
<td>----------------------------------</td>
<td>------------------------------------------------------------</td>
</tr>
<tr>
<td>fish, shrimp, salmon*, rabbit*, lobster*</td>
<td>Shower, bath, sauna, baseball,</td>
<td>Arm, leg, hand, face, back*, head*, foot*, shoulder*, lip*, heart*, eye*, tongue*, wing*</td>
<td>Church, factory, school, airplane,</td>
<td>Begin, start, finish, try</td>
</tr>
<tr>
<td>some water / The chicken is tasty</td>
<td>The shower was leaking / The shower was relaxing</td>
<td>John’s arm was tired / The arm was discolored</td>
<td>The church was built 20 years ago / The church sang a song</td>
<td>John began reading the book / John began the book</td>
</tr>
<tr>
<td>deer, calf, sheep</td>
<td>Classroom, racket, toilet</td>
<td>Hair, calf, wrist, bones</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Instance of an entity for kind of entity it is</td>
<td>InsKnd</td>
<td>Tennis, soccer, cat, dog, class*, dinner*, chair*, table*</td>
<td>*Tennis was invented in England / Tennis was fun today</td>
<td></td>
</tr>
<tr>
<td>-----------------------------------------------</td>
<td>--------</td>
<td>----------------------------------------------------------</td>
<td>------------------------------------------------------</td>
<td></td>
</tr>
<tr>
<td>Location for placing object in location</td>
<td>Location</td>
<td>Bench, land, floor, ground, box*, bottle*, jail*</td>
<td>The <em>bench</em> was made of pine / The coach <em>bench</em>ed the player</td>
<td>Garage, oven, hive</td>
</tr>
<tr>
<td>Object/substance for placing object/substance at goal</td>
<td>LocGoal</td>
<td>Water, paint, salt, butter, frame*, dress*, oil*</td>
<td>The <em>water</em> is cold / He <em>watered</em> the plant.</td>
<td>Blanket, shirt, ring, letter</td>
</tr>
<tr>
<td>Object/substance for taking object/substance from source</td>
<td>LocSrc</td>
<td>Milk, dust, weed, peel, pit*, skin*, juice*</td>
<td>The <em>milk</em> tastes good / He <em>milk</em>ed the cow.</td>
<td>Lint, fleas</td>
</tr>
<tr>
<td>Material for artifact</td>
<td>MatArt</td>
<td>Tin, iron, china, glass, linen*, rubber*, nickel*, fur*</td>
<td>Watch out for the broken <em>glass</em> / He filled the <em>glass</em> with water</td>
<td>Copper, aluminum, silver, clay, cement, wool, yarn, cotton</td>
</tr>
<tr>
<td>Object for color</td>
<td>ObjCol</td>
<td>Orange, violet, peach, rose, gold*, amber*, lavender*, turquoise*</td>
<td>She ate an <em>orange</em> / She has an orange <em>t-shirt</em></td>
<td>Yellow, green, pink, black, brown, purple, white, blue, scarlet</td>
</tr>
<tr>
<td>Occupation for a role played in action</td>
<td>OccRol</td>
<td>Boss, nurse, guard, tutor,</td>
<td>My <em>boss</em> is nice / He <em>boss</em>ed me around.</td>
<td>Chef, lawyer, priest</td>
</tr>
<tr>
<td>Place for an event</td>
<td>PlEv</td>
<td>Vietnam, Korea, Waterloo, Iraq</td>
<td>It is raining in Vietnam / John was shot during Vietnam</td>
<td></td>
</tr>
<tr>
<td>Place for an institution</td>
<td>PlIns</td>
<td>White House, Washington, Hollywood, Pentagon, Wall Street*, Silicon Valley*, Supreme</td>
<td>The <em>White House</em> is being repainted / The <em>White House</em> made an announcement</td>
<td></td>
</tr>
</tbody>
</table>
## 2.3 Procedure
The survey was hosted online using the LimeSurvey package. Participants completed it at their own pace, and in a place of their choosing. Upon completion, they received a gift voucher.

The survey was broken into 27 sections, based on the 27 English patterns that we evaluated. Each section began with a description of the pattern under investigation, instructing participants about the kinds of words they would be making judgments about. For example, before beginning the animal for meat section, participants read: “In this section, we will ask you a series of questions about words that, in English, can be both the name for an animal, and the name for the meat from that animal. In addition, we will

<table>
<thead>
<tr>
<th>Court*</th>
<th>Plant for food or material</th>
<th>PlntFd</th>
<th>Corn, broccoli, coffee, cotton, lettuce*, eggs*, oak*, pine*</th>
<th>The large field of corn / The corn is delicious</th>
<th>Grape, orange, apple, olive, chickpea</th>
</tr>
</thead>
<tbody>
<tr>
<td>Substance for portioning of that substance</td>
<td>Portion</td>
<td>Water, beer, jam</td>
<td>She drank some water / She bought three waters</td>
<td>Book, car, toy, hat</td>
<td></td>
</tr>
</tbody>
</table>

| Publishing institution for product created by institution | Prd | Newspaper, magazine, encyclopedia, Wall Street Journal*, New York Times*, People* | The newspaper is badly printed / The newspaper fired three employees | Book, car, toy, hat |

| Artist for product created by artist | PrPr | Writer, artist, composer, Shakespeare, Dickens*, Mozart*, Picasso* | The writer drank a lot of wine / The writer is hard to understand | |


| Object for something that is visually or functionally-related | VisFun | Beam, belt, column, stick, bug*, leaf* | Most of the weight in the structure rests on the beam / There was a beam of light | |
ask you about words that can be the name of an animal, but not the name for the meat of that animal.”

After reading the instructions for a section, participants answered a series of questions designed to investigate whether attested English senses (and exceptions to patterns) existed in their language. For each pattern, participants received the following kinds of items:

1) *Judgments about participating English senses.* These items tested whether translation-equivalents to participating English words were similarly polysemous in the tested language, using several measures. An example for the word *chicken* is provided in Figure 2.

![Figure 2](image)

*Figure 2.* An example of the judgments probed for the word *chicken*, in the Animal for Meat pattern.

As can be seen, for each tested word, participants were first given an explanation of the base sense of the English word (e.g., the animal sense of *chicken*), as well as an example of a sentence in which this sense of the word was used (e.g., “The *chicken*
drank some water.”). Participants were then asked to provide a translation of the critical word, as it was used in the sentence (see Figure 2).

Following this, participants were given an explanation of the extended English sense, as well as a sentence in which the English word was used in the extended sense (e.g., “The chicken is tasty”). They were then asked whether the translation of the base sense they had identified could be similarly extended, and provided a naturalness rating on a 1 (Not Natural) to 5 (Perfectly Natural) scale (see Figure 2).

Next, participants made an implicit binary judgment about the acceptability of extending their translation of the base sense. If they deemed this extension “natural”, they were asked to translate an English sentence using the extended sense (“The chicken is tasty”) into one box, and if they deemed the extension unnatural, they were asked to provide an alternative translation of the sentence, in a second box (see Figure 2). We refer to participants’ decision of which box to use as a binary acceptability judgment. When reading the instructions, participants were told to note down if the translation of the second sense involved adding a morpheme to the original word, or the creation of a compound.

Finally, participants were asked to list any other additional senses of the word they had translated for the base sense of the target word (e.g., “If there is another way of using “chicken” that immediately comes to mind, please describe it below”; see Figure 2). This was to probe for the existence of other senses within the same pattern (e.g., the use of glass to label a mirror, as opposed to a drinking vessel), as well as other, possibly language-specific patterns of polysemy.

2) Judgments about additional examples. For some patterns, we also included an additional list of English words that participated in the pattern (see Table 3 for an indication of which patterns included these questions). In the interests of time, participants were not asked to give naturalness ratings. Instead, we simply asked participants to provide a translation for each example, and then asked them to judge whether their provided translation also had an extended sense in line with the pattern in question.

3) Judgments about English exceptions. These items tested whether senses that are exceptions to patterns in English (e.g., the use of cow to label beef) nevertheless appear in other languages. For some patterns, we could not identify any exceptions to patterns, but for others, we provided participants with a list of exceptions, and asked them questions about these exceptions (see Table 3). Participants provided translations for each exception and indicated whether this translation could be extended (e.g., whether their translation of cow could label beef), just as they did for the judgments about additional examples. If the translation did not have an extended sense, participants were asked to provide the distinct word in their language that corresponded to this extended sense.
4) **Free recall of participating senses and exceptions.** With these items, we asked participants to provide a list of any additional words in their languages that either did or did not participate in the pattern in question.

5) **Generalization judgments.** Finally, we tested whether participants were willing to coin novel senses based on the pattern in question. Participants read about a newly coined word that corresponded to the base sense of the pattern, and then rated whether this word could be felicitously extended (on a 1-5 scale). For example, for the animal for meat pattern, participants read: “Imagine that a new animal was discovered called a “dax”. Imagine that a person was eating the meat derived from this animal and found it to be tasty. How acceptable would it be to say, in your language, that “The dax is tasty?” We included generalization questions for all patterns, except for the complement coercion pattern, for which we could not easily construct a candidate novel example.

2.4 **Dictionary analysis**

We complemented participants’ responses by exploring whether additional senses, for each pattern, were listed in bilingual dictionaries (see Table 2 for references). To do so, we took, for each language, one participants’ translations of base senses (e.g., of the animal sense of chicken), and noted whether additional senses for those words were listed in dictionaries, and whether those senses fit the target patterns. We were unable to perform a dictionary analysis for Farsi: Our Farsi respondents all transliterated into the Roman alphabet, and online attempts to transliterate them back into the Persian alphabet (using http://www.behnevis.com/en/) did not produce words that could be found in Farsi dictionaries.

3 **Results**

A full description of our results is beyond the scope of this paper. Here, we confine our analyses to the current hypotheses of interest. Our full dataset (along with analysis scripts) is available upon request. Below, we present our analyses, organized into sections according to the hypotheses they address.

3.1 **Hypothesis 1. Patterns of polysemy should be present across languages.**

As noted above, if children’s expectations about the flexibility of word meaning are based on a universal set of cognitive biases through which they construe the world, we would expect patterns of polysemy to be adapted to these biases. This predicts that the patterns observed in English should also recur across other languages.

To test this hypothesis, we evaluated whether each of the target patterns was present in each of the tested languages. We considered a pattern to be present in a language if one of the following criteria was met:
1. If participants judged the extended sense of an example as acceptable in one of the judgment tasks (i.e., in the binary acceptability judgment task, or in the judgment tasks about additional examples and exceptions).³

2. If participants listed a word from their language that participated in the probed pattern, in the free recall task.

3. If the dictionary analysis uncovered a polysemous word in the language that fit the pattern.

Figure 3. Evidence regarding the existence of patterns across languages. A white box indicates evidence for a pattern in a language, and a grey box indicates no evidence for the pattern. The figure is ordered such that languages with evidence for more patterns are toward the top, and patterns that are attested across more languages are toward the right. The place for event pattern is excluded from this analysis (see footnote 5). Table 3 contains a legend for the pattern names.

Figure 3 depicts the results for each of the patterns, across each of the languages. As can be seen, the data are broadly consistent with our prediction that English patterns of polysemy are also robust across other languages. In particular, only 23 pattern/language combinations were unattested out of a total of 390 possible combinations, a rate of 6%. This rate of absent pattern/language combinations is much

³ Naturalness ratings were not used in this analysis because they were redundant with participants’ binary acceptability judgments.
lower than would be expected by a theory in which all patterns are language-specific conventions. Further, although our criteria considered a pattern to be present in a language when one example of that pattern could be identified, most patterns were not limited to one participating example. Indeed, as indicated in Table 4, native speakers often indicated that there were multiple participating examples in their language, for each pattern. Together, these data provide support for our hypothesis that the patterns observed in English are also present in other languages, perhaps because these patterns are adapted to universal cognitive biases.

Table 4. Average number of examples judged or listed by native speakers as acceptable, for each pattern and language (data from the dictionary analysis and for place for event are not included; see footnote 5).

But what can we make of the fact that some patterns appeared to be more attested across languages than others? For example, patterns like plant for food (“The corn is delicious”) and container for representational contents (“The book is funny”), were robust across languages (Figure 3), and participants were able to report many participating examples (Table 4). In contrast, other patterns, like locatum verbs describing sources (“John milked the cow”) and location verbs describing where objects are placed (“She boxed the books”) were more absent across languages, with fewer reported examples. One question raised by these data is whether some patterns were reliably more absent across languages than others, or instead whether observed variability across patterns was an artifact of limitations in our methods. For example, it is possible that we did not find evidence for some pattern-language combinations not because they do not exist, but instead because our participants forgot the appropriate translations or because we could only assess a limited number of examples per pattern.4

4 The latter seems unlikely, because there was no correlation between the number of examples assessed per pattern and the proportion of languages that pattern was absent in (r(24)=.02, ns).

<table>
<thead>
<tr>
<th>Language</th>
<th>Spanish</th>
<th>Farsi</th>
<th>Turkish</th>
<th>Russian</th>
<th>Japanese</th>
<th>Indonesian</th>
<th>Hindi</th>
<th>French</th>
<th>Cantonese</th>
<th>Korean</th>
<th>Italian</th>
<th>Hungarian</th>
<th>Vietnamese</th>
<th>Mandarin</th>
</tr>
</thead>
<tbody>
<tr>
<td>Location</td>
<td>3.5</td>
<td>3.5</td>
<td>3.0</td>
<td>0.0</td>
<td>0.5</td>
<td>0.5</td>
<td>0.3</td>
<td>0.4</td>
<td>0.3</td>
<td>0.0</td>
<td>0.4</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>Location</td>
<td>2.7</td>
<td>6.7</td>
<td>0.0</td>
<td>1.0</td>
<td>2.0</td>
<td>2.0</td>
<td>2.0</td>
<td>4.0</td>
<td>1.0</td>
<td>2.0</td>
<td>4.0</td>
<td>4.0</td>
<td>2.0</td>
<td>1.0</td>
</tr>
<tr>
<td>Location</td>
<td>1.0</td>
<td>0.5</td>
<td>0.0</td>
<td>3.0</td>
<td>1.5</td>
<td>1.5</td>
<td>3.0</td>
<td>4.0</td>
<td>1.0</td>
<td>3.0</td>
<td>4.0</td>
<td>4.0</td>
<td>4.0</td>
<td>3.0</td>
</tr>
<tr>
<td>Location</td>
<td>0.7</td>
<td>5.0</td>
<td>0.0</td>
<td>0.0</td>
<td>2.0</td>
<td>2.0</td>
<td>2.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>Location</td>
<td>0.4</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
</tbody>
</table>
If the observed absent-pattern language combinations were due to limitations in our methods, and do not reflect that some patterns are reliably more absent than others, then they (i.e., the gray squares in Figure 3) should be randomly-distributed across both languages and patterns. To explore if this was the case, we used a resampling approach to estimate how likely it would be, given the observed number of absent patterns in each language, that a single pattern would be absent across multiple languages. Thus, we took the data from each language, and within that language, randomly shuffled the pattern labels (e.g, “material for artifact”) so that each label was associated with a randomly-chosen data point. Consequently, there was a 1 out of 26 chance that the “material for artifact” label would be associated with the data for the “material for artifact” pattern, as opposed to the data corresponding to the other 25 patterns. We did this for each language, and then recorded the largest number of languages for which a single pattern was absent (e.g., in our original dataset this number would be 6). We repeated this process 10,000 times, and then compared the resulting empirical distribution of “most absent” languages for patterns to our actual dataset.

The most absent pattern in our dataset was the set of locatum verbs describing sources, which was absent in 6 languages (see Tables 3, 4, Figure 3). In our simulations, it was very rare for a pattern to be absent in 6 or more languages: this occurred only 1 out of 10,000 times. The probability of a pattern being absent in 6 or more languages of our sample is therefore approximately .0001, which meets standard criteria for statistical reliability, even when we correct for multiple comparisons, suggesting that this pattern is less attested than we might expect by chance. The next most absent pattern, was the set of verbs describing the location something goes to – these were absent in 4 languages. Our simulations suggest that this result was also quite improbable \( p = 0.024 \), although this is not significant if we correct for multiple comparisons. But after this, the probability of a pattern being absent in fewer than 4 languages by chance, which was true of several other patterns in our dataset, was above standard thresholds for statistical significance.

The analysis described above suggests that the absent pattern-language combinations in our data were not entirely randomly distributed. For example, an unexpectedly high number of languages lacked locatum source verbs. Although this could be explained by limitations in our methods, we think that this is unlikely. Instead, we believe that there is a good reason for the high number of languages that did not include locatum source verbs, which we set out in the discussion section.

\[5\] We excluded the place for event pattern from this analysis, as we did not have data from Spanish and Italian for this pattern.

\[6\] To ensure that this result was not driven by data from the dictionary analysis (which potentially may not reflect speakers’ intuitions), we excluded these data and repeated the analysis. The results were the same.
Our analyses do not speak to whether those patterns that were absent in fewer than four languages in our data are indeed absent in those languages. On one hand, as noted above, it remains possible that we did not find evidence for those patterns because we failed to ask participants about actual, participating examples. On the other hand, it is also possible that our data do reveal genuine cross-linguistic variability in the presence of patterns. For instance, it is worth noting that most of the patterns for which there were absent languages involved senses that cross grammatical categories, appearing in the syntax as nouns or verbs (e.g., in the case of locatum and location verbs), or as count or mass nouns (e.g., in the case of grinding and portioning). Thus, these patterns may not have been present in some languages due to restrictions those languages place on syntactic flexibility.

To sum up, although our data provide evidence of some cross-linguistic variation with respect to patterns of polysemy, it is important to note that most patterns were generally robust across languages. Indeed, all patterns were present in multiple languages, and most were present in nearly all probed languages.

3.2 Hypothesis 2: For some patterns, participating senses should recur across languages, while for other patterns, participating senses should be more variable.

Although our theory predicts that patterns of polysemy should be present across languages, it also predicts that these patterns may not be instantiated in the same ways across languages. In particular, our theory predicts that the less predictable a pattern is – i.e., the less strongly it constrains the referents of its different senses – the more likely it is that languages will develop diverging conventions from one another regarding which senses may participate in these patterns. This would predict, for example, that while a predictable relation, like animal for meat, should have very similar participating senses across languages, a less predictable relation, like material for artifact, may have more variable senses across languages. In this section, we focus on whether senses are indeed variable across languages, and whether this depends on the pattern in question.

To examine whether participating senses in English are also present in other languages, we focused only on those items that all participants received, and in particular, on subjects’ judgments of participating English examples (e.g., excluding data from the free recall task). Our analyses incorporated two measures: 1) naturalness ratings (i.e., 1-5 ratings of whether an extended sense was natural) and 2) binary acceptability judgments (i.e., whether extended senses were judged acceptable or not). For each language and each probed sense, we computed average naturalness ratings and binary acceptability judgments across respondents for that language. Figure 4 depicts the mean naturalness ratings across languages, separated by pattern.

To get a broad sense of the degree to which English senses are also attested in other languages, we first explored whether the naturalness and acceptability judgments for English senses (i.e., by our native English-speaking controls) predicted judgments of
these senses in each of the other 14 languages. Interestingly, these judgments were generally uncorrelated, as only 2 of 14 of the pair-wise correlations for the naturalness ratings reached significance (Mean $r(98) = .13$, range: $0.02 - 0.26$; mean $p = .26$, range: $.01 - .84$), while 1 of 14 correlations for the binary acceptability judgments reached significance (Mean $r(98) = .11$, range: $-.03 - .23$; mean $p = .34$, range: $0.03 - .89$).²

![Figure 4. The mean naturalness ratings of senses for each pattern, by language. Average rating is color-coded. Higher scores — which are progressively whiter — indicate that probed senses was rated as being more acceptable, while lower scores — which are progressively bluer — indicate that senses were less acceptable. Grey squares with crosses indicate missing data. The heatmap is ordered such that the languages with the highest average rating are toward the top, and patterns with the highest average rating are toward the right. Table 3 contains a legend for the pattern names.](image)

This analysis suggests that there is large variation in the senses that participate in polysemy patterns across languages. But to what extent does this variation depend on the specific pattern in question? As can be seen in Figure 4, for some more stable patterns, ratings of senses across languages were more similar to one another than they were for other, less stable patterns. For example, the plant for food pattern was

---

² We excluded the place for event pattern from this analysis, as it was not tested in Italian and Spanish.
very stable across languages. The animal for meat pattern was a little more variable, because a number of languages label the meat by compounding some animal names with the word for “meat” (e.g., Farsi, Chinese, Japanese, Korean, Indonesian). However, respondents often felt that use of the animal name on its own was still acceptable: e.g., a Korean respondent wrote that “In Korean, one can say the chicken is tasty. However, it is more proper to say the chicken meat is tasty”). Finally a number of patterns, like material for artifact, were instantiated in very different senses across languages. Thus, iron, tin and glass often did not name the same artifacts in English as they did in other languages. However, as noted in the previous section, the pattern was still attested across different languages, though applied in different ways: e.g., a respondent noted that the Russian word for rubber can also name a car tire.

Our theory proposes that this variation in sense acceptability across patterns results from variation in the predictability of different patterns. For less predictable patterns, languages may develop different conventions, that license different participating senses. However, an alternative possibility is that this variation is due to chance, and is thus randomly-distributed. As a test of this second possibility, we used a resampling analysis to ask whether the distribution observed in Figure 4 could be a product of random variability. In particular, we asked whether differences between more and less stable patterns were larger than would be expected if we had computed our averages over random groups of senses, as opposed to grouping senses by their pattern. That is to say, are there real differences between patterns?

To conduct this analysis, we first measured the difference between more and less cross-linguistically stable patterns. To do so, we created a stability score for each pattern, by averaging the mean sense acceptability ratings for each pattern across languages (excluding English), and then ranked these patterns by their stability scores. We then regressed the rank order of each pattern against its stability score. The regression slope served as our measure of difference in cross-linguistic stability: the steeper the slope, the larger the difference between more and less stable patterns.

If patterns play no role, then the steepness of this regression slope should be no different when senses are randomly assigned to different patterns. We therefore compared our observed slope against a distribution of simulated regression slopes that did not invoke patterns. Specifically, within each language, we first shuffled pattern labels for the probed senses to create random groups of senses, e.g., such that “material for artifact” could come to be associated with an animal for meat word like chicken. Then, we calculated regression slopes based on the stability scores for these random groups of senses. We repeated this process 10,000 times, and then compared the regression slope derived from our data – which was based on grouping senses by

---

8 Still, this was not without exceptions: e.g., while most languages use the same word for the cotton plant and material, Hungarian differentiates the two (gyapot describes the plant; pamut the material).
patterns – to the empirical distribution of slopes – which used random groups of senses. Strikingly, the regression slope derived from our data was steeper than each of the 10,000 slopes derived from the simulations, indicating that differences in sense acceptability between more and less stable patterns were indeed larger than would be expected by chance.

Together, these analyses provide support for our hypothesis that, although patterns of polysemy are generally present across languages (as shown in the previous section), there is large variation in the senses that participate in these patterns across languages, with some patterns being more variable than others. The following section explores why some patterns might be more stable, and in particular, whether this depends on pattern predictability.

3.3 Hypothesis 3. Patterns whose senses show the least variability across languages should also be the most generative

Our theory predicts that, when a pattern’s predictability is high (i.e., such that the referents of its different senses are tightly constrained), that pattern should be stable across languages, such that the same senses appear across languages. Our theory also predicts that when a pattern’s predictability is high, language users will find it easier to make a higher-order generalization about the senses that instantiate that pattern (e.g., realizing that any word for an animal can label its meat), allowing them to coin novel senses that follow that pattern. Our final analysis tested whether this predicted association holds true: Are more generative patterns also more likely to be stable across languages?
Figure 5. The mean acceptability ratings of novel senses for each pattern, by language. Higher scores – which are progressively whiter – indicate that the novel sense was rated as being more acceptable, while lower scores – which are progressively bluer – indicate that this sense was less acceptable. Grey squares with crosses indicate missing data. Table 3 contains a legend for the pattern names.

We measured each pattern’s generativity using the survey’s generalization question, which asked participants to rate the acceptability of a novel sense on a 1 to 5 scale. Figure 5 depicts the average acceptability ratings of these novel senses, for each language and pattern. As can be seen, although there was considerable cross-linguistic variation in ratings of the novel senses across different patterns, there was also consistency in the ratings of novel senses within patterns. To confirm this, we correlated ratings for novel English senses with ratings of these senses in each of the other 14 languages, and found reliable correlations in each case (mean r(24) = .58, range: .42 - .79; mean p = .008, range: .000003 - .04). This is especially striking, in light of the fact that acceptability ratings of existing English senses were not correlated with acceptability ratings of those senses in other languages (see previous section). This suggests that cross-linguistic variability in how senses are instantiated may indeed be related to universal intuitions about how easy it is to coin new senses for patterns.

Finally, we conducted an analysis to directly address our primary question: Does the generativity of a pattern predict the cross-linguistic stability of its senses? To create a generativity score for each pattern, we averaged together generalization ratings from all of the non-English languages, for each pattern. Thus, higher generativity scores indicate that participants, across languages, judged the novel senses as more acceptable. To create a cross-linguistic stability score for each pattern, we averaged the naturalness ratings of the different senses from that pattern, across all of the non-English languages. Higher stability scores thus indicate similar senses across languages.

Note that we excluded the Place for Event pattern from this analysis, as it was not assessed in Italian and Spanish, as well as the complement coercion pattern, for which we did not test generalization.
Figure 6. The relationship between a pattern’s generativity (i.e., the acceptability of novel senses across languages, plotted on the y-axis) and its cross-linguistic stability (i.e., the acceptability of attested English senses across languages, plotted on the x-axis).

Figure 6 plots stability scores against generalization scores. Strikingly, the relationship between these two variables was very strong, \( r(24) = .91, p < .001 \). Thus, patterns that were more stable across languages were also more generative.\(^{10}\) This is consistent with

\(^{10}\) Readers may be concerned that this analysis risks circularity. For example, stability scores might be lower when a pattern is rarely used across languages, and so generativity scores would also be correspondingly lower. To assuage this concern, we repeated the analysis by comparing cross-linguistic stability to English generativity (where all the patterns are known to exist). Again, we observed a robust correlation
our prediction that generativity and cross-linguistic stability should be related, because they are each determined by a common factor: pattern predictability. Specifically, when a pattern is more predictable, languages will develop similar conventions regarding which senses participate in the pattern. At the same time, when a pattern is more predictable, the relation between each participating word’s senses will be easier to abstract and thus generalize to novel senses.

4 Discussion

Polysemy is a quintessential feature of language, and provides a window into the relationship between lexical and conceptual structure. Here, we have proposed that the structure of polysemy is a consequence of how children approach the task of building a lexicon. Specifically, we have argued that children come to the task of word learning with a set of expectations about how words can be used flexibly, and that these expectations follow from early-developing cognitive biases through which children construe the world. Over generations, the structure of polysemy within a language adapts to meet children’s expectations about word meaning, and thus develops patterns of easily learnable alternations between senses. This in turn facilitates the acquisition of the lexicon, because by learning one sense of a word, children are well-equipped to guess its other meanings. Thus, in our view, children actively shape the structure of polysemy and, in return, polysemy makes it easier for children to acquire a lexicon.

We used a large-scale survey to test our theory’s predictions about regularity and variation in polysemy across languages. By our account, because children’s expectations about polysemy are guided by universal cognitive biases, we should see very similar patterns of senses across languages. However, our theory also predicts that these patterns will be instantiated in different ways across languages, with some patterns having more variable senses across languages than others. To explain differences between patterns, we introduced the idea of pattern predictability, based on the number of possible senses for a word that conform to a particular pattern, and argued that only predictable patterns were likely to have similar senses across languages.

Our data broadly confirmed these predictions. Across fifteen languages and 27 patterns of polysemy, we found very few instances where a language showed no evidence for having a particular pattern. The robust presence of patterns across languages is therefore consistent with the idea that these patterns are a result of universal and early-developing conceptual biases. In addition, we found that the individual senses of some polysemy patterns were much more stable across languages than others. Patterns with the most stable senses were also the most generative, consistent with our hypothesis that regularity in senses is driven by pattern predictability.

\( r(24) = .65, p<.01 \), consistent with the proposal that differences in polysemy across languages are due to pattern predictability.
In the remainder of this discussion, we describe how our data speak to other theories of polysemy, discuss avenues for future cross-linguistic work and aspects of our data that need further explanation, and present other predictions of our theory concerning lexical and conceptual development.

4.1 Implications for other theories of polysemy

Our data have important implications for both conventionalization and pragmatic approaches to polysemy. In particular, neither class of theories does a good job of explaining why on the one hand, the same patterns of polysemy are present across languages, while on the other hand, the senses that instantiate those patterns can be variable.

According to conventionalization theories, both senses and patterns should be variable across languages. Because conceptual structure plays a limited role in how senses are coined and learned, languages should develop senses in their own particular ways. Further, because patterns are formed as speakers coin new senses by analogy to older senses (see Murphy, 2007), languages should ultimately develop quite different patterns, based on the first senses that emerge. This prediction is not consistent with our finding that the same patterns of polysemy recur across languages. Moreover, because conventionalization theories predict that senses should be variable across languages, they also have trouble explaining why some of the patterns we assessed were quite stable across languages (e.g., animal for meat). Together, these considerations suggest that, although senses may be learned as conventions, this process may be constrained by cognitive biases that children of all linguistic communities share.

Our data are also difficult to explain for pragmatic theories, which hold that polysemy is a direct reflection of conceptual structure. On these theories, both patterns and senses should be uniform across languages. In particular, because speakers of different languages should generally find the same relationships to be noteworthy, they should derive senses from core meanings of words in similar ways. However, although we found that the same patterns of polysemy are generally present across languages, the senses that participated in these patterns were at times quite variable. For instance, our data reveal that while English uses glass to describe a material and a drinking vessel made from the material, this artifact sense of glass is not used in a number of other languages, such as Spanish (where glass is instead used to label a car window), or Russian (where glass can label a mirror). Meanwhile, Turkish, Cantonese, Farsi, Japanese and Korean demand that speakers use a whole phrase, the equivalent of “glass cup”). These facts are difficult to explain on a pragmatic theory, because there is little reason to believe that children learning Russian find the relationship from glass material to glass cups less noteworthy than the relationship to mirrors (and vice versa for children learning English). Thus, the presence of language-specific senses suggest
that at least some senses are not pragmatically derived from core meanings, but instead may be conventions that speakers have to learn.

Because theories that ascribe polysemy solely to conventions or pragmatics face problems, one possibility is that a hybrid theory, that combines these mechanisms together, may do a better job of explaining our data. For example, under a hybrid theory, those patterns that have stable senses across languages (i.e., forms of regular polysemy like animal for meat) could rely on pragmatic mechanisms, while those patterns that have variable senses across languages (i.e., forms of irregular polysemy like material for artifact) could be composed of stored conventions. However, such a theory would also have to motivate the difference between regular and irregular patterns of polysemy. Perhaps the most obvious way to do this would be to propose that relations encoded by irregular patterns (e.g., material for artifact) are on the whole less noteworthy than relations encoded by regular patterns (e.g., animal for meat). This is of course an empirical question, but we doubt that this claim is uniformly true. When you look at a drinking glass, its material is clear and evident. By contrast, when you look at some chicken meat, its relationship to the animal is not perceivable. This suggests that metrics such as noteworthiness are not satisfactory for explaining the difference between regular and irregular polysemy, and indeed Rabagliati, Marcus & Pylkkänan (2011) provide evidence that this is the case.

We suspect that our theory provides a better explanation for this difference. By our account, senses that instantiate both kinds of patterns must be learned by children as conventions (i.e., neither are pragmatically derived), but learning is constrained by universal cognitive biases. Senses are more cross-linguistically variable for irregular patterns not because they involve less noteworthy relations, but because these patterns are less predictable, allowing different languages to develop different senses to fit these patterns.

4.2 Directions for future cross-linguistic work

To our knowledge, our study provides the broadest cross-linguistic survey of polysemy to date (i.e., in terms of the number of languages and polysemy patterns assessed). However, a number of exciting future directions still remain, which we describe below.

First, although we tested a variety of languages, we nonetheless only covered a small fraction of the languages spoken across the world. For instance, we did not assess any African languages, nor did we include any Semitic languages. The latter are particularly worthy of further study, because they include morphological paradigms through which roots can be adjusted to encode different meanings, which children master early in life (see, e.g., Berman, 1999, for the case of Hebrew). Although our study used a working definition of polysemy in which the same word form is used to label different meanings, the functions served by polysemy bear a clear resemblance to those served by morphological paradigms (see e.g., Pylkkänen, Llinas & Murphy, 2006). Indeed, in
some cases, polysemous alternations are mirrored by parallel morphological rules – e.g., just as one can say “He drank a bottle of whiskey”, using bottle as a measure of the contents of a bottle, one can also say “He drank a bottleful of whiskey”, using the suffix –ful to indicate a measure (Copestake & Briscoe, 1995). Thus, future research should explore how morphological phenomena are related to polysemy across languages.

Second, although we tested a large number of patterns of polysemy, time constraints meant that, on average, we assessed only eight examples per pattern (see Table 3). In principle, this aspect of our sample could have skewed our results in two ways. First, it may have led us to overestimate how cross-linguistically stable the senses of certain patterns actually are. That is, it is possible that the limited number of senses we tested were somehow special, and more stable across languages than other potential senses. Second, our sampling may have led us to conclude that patterns were absent when they actually were not. For instance, we found no evidence for portioning (“She bought three waters”) in Vietnamese or Hindi, but this could be because we failed to assess actual, participating examples.

Third, and related to the above, because we focused specifically on patterns of polysemy found in English, our data cannot speak to the cross-linguistic regularity of patterns that might exist only in other languages. Thus, it remains possible that these unexplored patterns might be significantly more variable across languages than the patterns we did test, which would, of course, not be predicted by our theory. We have no reason to believe that patterns sampled from English would be more cross-linguistically regular than other patterns, but this remains an open question.

The above discussion suggests that a more extensive survey will be needed to reach firm conclusions about the regularity of patterns and senses across languages. Importantly, such an effort could be valuable, as it could open a window onto conceptual structure. The finding that a pattern is relatively unattested across languages, for example, may raise questions about whether the relation it encodes is less salient than those encoded by other patterns. Consider, for example, the most absent pattern in our data, the set of locatum verbs describing the transfer of substances from sources (“John milked the cow”), which was absent in six languages. As noted in the results section, this level of absence was more than would have been expected by chance, suggesting that there may be good reasons for its absence. And indeed, recent work has suggested that children learning English have a surprising level of difficulty learning verbs like this, because they initially assume that they label the transfer of substance to goals rather than from substances (Srinivasan & Barner, 2013b). This results in striking errors, wherein children initially assume that “milking the cow” involves putting milk onto a cow! This bias likely stems from a more fundamental goal bias in language and thought, which has been documented extensively in the literature (e.g., Lakusta & Landau, 2005; Lakusta, Wagner, O’Hearn, & Landau, 2007; Papafragou, 2010; Regier, 1997).

4.3 Other predictions of our theory for lexical and conceptual development
To recapitulate, we have proposed that children’s hypotheses about the flexibility of word meaning are guided by cognitive biases through which they construe the world. These biases actively shape the structure of polysemy across languages, such that languages develop patterns of senses that are easily learnable for children, and thus facilitate their acquisition of a lexicon. In the present study, we tested our theory by exploring cross-linguistic regularity in polysemy. However, our theory also generates a number of predictions about lexical and conceptual development.

One of these predictions relates to the idea that children’s earliest hypotheses about word senses are guided by flexible, non-linguistic conceptual structures. In the introduction, we described some evidence that children’s early understanding of the world is guided by theories that allow them to construe and explain entities in different ways, by invoking explanatory factors like form, function, and origin. We also pointed out the parallels between this evidence and work showing that similar structures seem to constrain polysemy in adults. However, despite these parallels, we lack direct evidence that non-linguistic construals play a causal role in how young children learn new senses.

A second prediction of our theory concerns how children form higher-order generalizations about word senses, and how this affects their acquisition and representation of polysemous words. As described in the introduction, higher-order generalizations may be formed as young children learn multiple senses following a pattern (e.g., the animal/meat senses of chicken, lamb, etc.), and abstract their underlying alternation (e.g., that words for animals can label their meats). This proposal raises a number of testable questions. First, how quickly are children able to make these generalizations, and thus, to what extent can these generalizations constrain children’s early hypotheses about flexibility? On one hand, it is possible that the process of forming such generalizations is a slow one, such that these generalizations do not play an important role in the acquisition of polysemous words. However, it is also possible that such generalizations are quickly abstracted, and facilitate acquisition from early in life. Second, how might higher-order generalizations affect how children represent polysemous words? One possibility, for example, is that children initially represent senses using separate lexical items, but are able to derive these senses online once they have formed higher-order generalizations. This proposal is similar to previous accounts of how children acquire regular past-tense inflections for verbs (e.g., Marcus et al., 1992; O’Donnell et al., 2011; Pinker, 1991).

A final question raised by our theory is how children’s conceptual knowledge is related to their knowledge of polysemy. Throughout, we have assumed that children’s knowledge – e.g., their knowledge that a drinking glass is made of glass – supports their expectations about polysemy. However, it is also possible that learning word senses teaches children about the world – e.g., that because a drinking vessel and transparent material share the same label, they must be related. Indeed, there are anecdotes that support this idea: for example, parents have recounted that their children became vegetarians upon learning that it was no accident that the same word, chicken, labels an
animal and its meat (Foer, 2010; Gelman, 2003)!
Our theory does not rule out the possibility that polysemy teaches children about the world. However, we suspect that, at least early in life, it is likely that children’s knowledge about the world will precede their knowledge of how words relate to one another, given that young children have difficulty explicitly reflecting upon how words are used (e.g., Bialystok, 1986, Gombert, 1992). Still, it remains possible that, in some cases, learning to use words flexibly might invite children to form relations between concepts. If this is the case, polysemy may open a window onto how language shapes cognition, including whether speakers of different languages may come to think differently about the world, by virtue of learning language-specific forms of polysemy.
References


Boyeldieu, P. (2008) From semantic change to polysemy. The cases of “meat/animal” and “drink”. In M. Vanhove (Ed.) *From polysemy to semantic change. Towards a typology of lexical semantic associations* (pp. 303-315). Amsterdam/Philadelphia: John Benjamins.


Gelman, S. A. & Bloom, P. (2000). Young children are sensitive to how an object was created when deciding what to name it. Cognition, 76, 91-103.


Polysemy Across Languages


