

Slobin, D. I. (in press). The child learns to think for speaking: Puzzles of crosslinguistic diversity in form-meaning mappings. *Studies in Language Sciences* 7. Tokyo: Kurosio Publishers.

The child learns to think for speaking Puzzles of crosslinguistic diversity in form-meaning mappings

DAN I. SLOBIN
University of California, Berkeley

Abstract

The basic problem of language acquisition is to arrive at the mappings between meaning and form in the exposure language. Neither the meanings nor the forms can be given to the child in advance, because there is great diversity between languages. It is proposed that the child is equipped with various procedures for extracting both meanings and forms. The paper focuses on problems of extracting appropriate semantic categories in various languages. The procedures available to the child ("bootstrapping") require attention to fine-grained semantic distinctions, syntactic constructions, and the typological tendencies of the exposure language.

1. Cognitive prerequisites: What are the starting points?

For the past half-century or more, child language researchers have been arguing about whether emerging linguistic forms reflect concepts that are already available to the child, or whether basic semantic concepts are, themselves, shaped in the process of learning to express them in language. My own starting point, 35 years ago (Slobin, 1971, 1973) was to propose "cognitive prerequisites" in the form of universal concepts that all children would attempt to map onto linguistic expressions. I proposed: "The appearance in child speech of a new formal device serves only to code a function which the child has already understood and expressed implicitly" (Slobin, 1971, p. 330). This sort of cognitive prerequisite had long been familiar in the field of developmental psychology: "New forms first express old functions, and new functions are first expressed by old forms" (Slobin, 1973, p. 184).

Twenty-five years later I found myself radically revising this position: "Children use linguistic cues to discover the collections of semantic elements that are packaged in the lexical and grammatical items of the language" (Slobin, 1997, p. 318). This revision was based on accumulating evidence of crosslinguistic diversity in the earliest meanings of children's utterance. Melissa Bowerman presented a critical formulation as early as 1985: "The initial organization of semantic space is not fixed but flexible... One important factor that can influence the meanings children adopt is the *semantic structure of the input language*..." (Bowerman, 1985, p.).

The goal of this paper is to lay out some of the evidence for this position of linguistic determinism, with a focus on three sorts of **relational meanings** and their encoding in early child language across various types of languages: (1) spatial locations (as expressed by verbs, prepositions, postpositions), (2) event participants (agent, actors, and patients, associated with verbs), and (3) motion events (verbs, verb particles).

2. Spatial relations

The cognitive prerequisites position was well supported by early crosslinguistic work on children's learning of terms expressing locative relations (prepositions or postpositions, depending on the type of language). In an early study (Johnston & Slobin, 1979), we compared acquisition of locative prepositions in English, Italian, and Serbo-Croatian, along with locative postpositions in Turkish. The four languages differ considerably in typology (e.g., the three languages that use prepositions use subject-verb-object order, while the one language that uses postpositions favors subject-object verb order), nevertheless, in all four languages, the basic order of acquisition of locative relational terms was the same, as shown in Table 1.

Table 1. Order of acquisition of locative relational terms in four languages

	English (prepositions)	Italian (prepositions)	Serbo-Croatian (prepositions)	Turkish (postpositions)
2-year-olds	<i>in, on, under, beside</i>	<i>in, on, under, beside</i>	<i>in, on, under, beside</i>	<i>in, on, under, beside</i>
3-year-olds	<i>back, front, between</i>	<i>back, front, between</i>	<i>back, front, between</i>	<i>back, front, between</i>

The order of acquisition matches the Piagetian development from simple topological relations (only requiring attention to features of two objects) to projective relations (requiring additional attention to perspective). To be sure, cognitive development is necessary to acquire the linguistic expressions for more complex relations. However, these four languages also happen to make use of the same relational concepts in the meanings of adpositions (prepositions and postpositions). More recent research has shown that there is considerable crosslinguistic diversity in form–meaning mappings (Bowerman, 1996; Bowerman & Choi, 2001; Levinson, 2003; Levinson et al., 2003).

One of the clearest examples comes from research by Bowerman and Choi comparing English with Korean (Bowerman & Choi, 2001; Choi & Bowerman, 1991). Developmental psychologists have assumed that infants begin with sensitivity to basic relations such as **containment** and **support**, as expressed by the English prepositions *in* and *on*. However, Korean makes a different distinction: what is important in Korean is not whether one thing is supported by another or is contained by another, but rather whether the relation between the two things is one of **tight fit** or **loose fit**. Consider, for example, the scenes shown in Figures 1 and 2. These figures show part of a larger set of contrasts between English and Korean

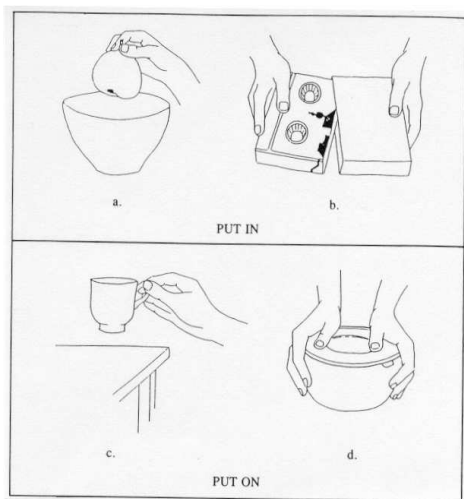


Figure 1. Classification of four actions as instances of containment (a, b) versus support (c, d). (Bowerman, 1996, p. 152)

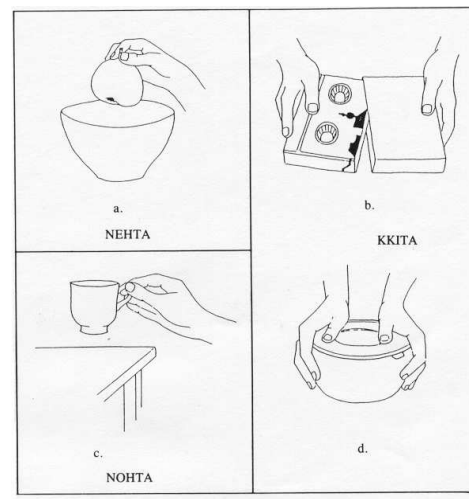


Figure 2. Classification of four actions as instances of loose fit (a) versus tight fit (b, d) versus loose surface contact (c). (Bowerman, 1996, p. 153)

English distinguishes **containment**—using *put in* regardless of tightness of fit, from **support**—using *put on* regardless of tightness of fit. Korean uses *nehta* for **loose fit**, *kkita* for **tight fit**—whether containment or support, and *nohta* for putting something **loosely on a horizontal surface**. In a preferential looking experiment with American and Korean infants aged 18–23 months, Choi et al. (1999) found that American babies, when looking at pairs of videos and hearing *put in*, preferred to look at scenes depicting containment, whether the fit was loose or tight. Korean babies in the same task, when hearing *kkita*, preferred to look at scenes depicting tight fit, whether the fit was one of containment or support. That is, the two groups oriented to language-specific categories in comprehension, early in the one-word period. In one- and two-word speech in the two languages, there were comparable differences in the semantic categories encoded by early words. For example, Figure 3 schematizes part of the domain of spatial relations expressed by children of 16–20 months of age in the two languages (Choi & Bowerman, 1991). The core notions that receive early expression do not line up between the two languages; however, either system is easily mastered. Bowerman concludes (1996, pp. 169-170): “[I]t is striking how quickly and easily children adopted language-specific principles of semantic categorization. There was little evidence that they had strong prelinguistic biases for classifying space differently from the way introduced by their language.”

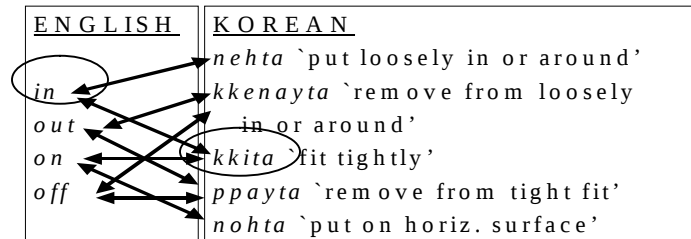


Figure 3. Early semantic categories in English and Korean child speech, 16–20 months.
(data from Choi & Bowerman, 1991)

With regard to the issue of cognitive prerequisites, current data and theory lead to a clear conclusion: ***The child can easily learn a wide range of semantic categories, depending on the exposure language.***

Given this conclusion, we can now ask: How does the child figure it out? Various sorts of operating principles (Slobin, 1973, 1985) have been proposed to account for the child’s acquisition of language. Such principles can be considered in terms of various proposals for “bootstrapping.”¹ Here I will examine three types of bootstrapping proposals: semantic bootstrapping (Pinker, 1984), syntactic bootstrapping (Gleitman, 1990), and typological bootstrapping (Slobin, 1997).

3. Semantic Bootstrapping

Pinker (1984, p. 40) has proposed “that the child initially uses semantic notions as evidence for the presence of grammatical entities in the input.” That is, semantics is used to bootstrap into syntax. Semantic bootstrapping requires that both the semantic and syntactic categories are available to the child, and that the child also knows the relevant mappings between the two types of categories. For example, if a

¹ “Bootstrapping” comes from the strange English expression, “to pull oneself up by one’s own bootstraps.” The humorous image is that of a person wearing boots and holding onto the straps at the top of each boot in an attempt to move upwards. *The Oxford English Dictionary* offers the following definition: “To make use of existing resources or capabilities to raise oneself to a new situation or state; to modify or improve by making use of what is already present.” The term was taken over by computer science to refer to loading and initializing an operating system (“boot up a computer”); in business, it refers to starting up a company with limited capital. In general, bootstrapping refers to procedures in which some built-in or already existing processes or capacities make it possible to advance in operating and building up a system.

child can (1) identify the agent of an action in a situation that is being talked out, and can (2) identify the word that refers to the agent, then the child also knows (3) that the word is a noun, and (4) it is the subject of the sentence. That is, the child is equipped with such knowledge as the fact that nouns refer to persons or things, and that a noun referring to an agent can be the subject of a sentence. As Pinker puts it: “The categorization of words can be inferred from their semantic properties, and their grammatical relations can be inferred from the semantic relations in the event witnessed” (1984, p. 40).

Semantic bootstrapping assumes that the child has a table of correspondences, or a set of linking rules between form and function, such as the following:

Table 2. Correspondences between form and function (excerpted from Pinker, 1984, Table 2.1, p. 41)

Grammatical element	Semantic inductive basis
SYNTACTIC CATEGORIES	
Noun	Name of person or thing
Verb	Action or change of state
Preposition	Spatial relation, path, or direction
...	...
GRAMMATICAL FUNCTIONS	
Subject	Agent of action...
Oblique	Source, goal, location, instrument
...	...
CASES	
Nominative or Ergative	Agent of transitive action
Accusative or Absolutive	Patient of transitive action
Dative	Goal of beneficiary
...	...

Semantic bootstrapping is an attractive hypothesis. Surely, children begin with some basic, prelinguistic notions (Slobin, 1985). But the child can use these notions as a reliable guide to the discovery of grammar ***only if meanings and grammatical forms line up neatly***, as suggested in Table 2. However, linguistic diversity such as that studied by Bowerman and Choi poses serious problems to a reliance on universal semantic starting points. The problem is that both form and function vary. For example, we have seen that spatial relational concepts vary; crosslinguistic comparison also shows that spatial relations can be mapped onto either prepositions, postpositions, case inflections, nouns, or verbs. Figure 4 shows part of the crosslinguistic diversity in means of expressing the locative notion: ‘Put (something) in a box’.

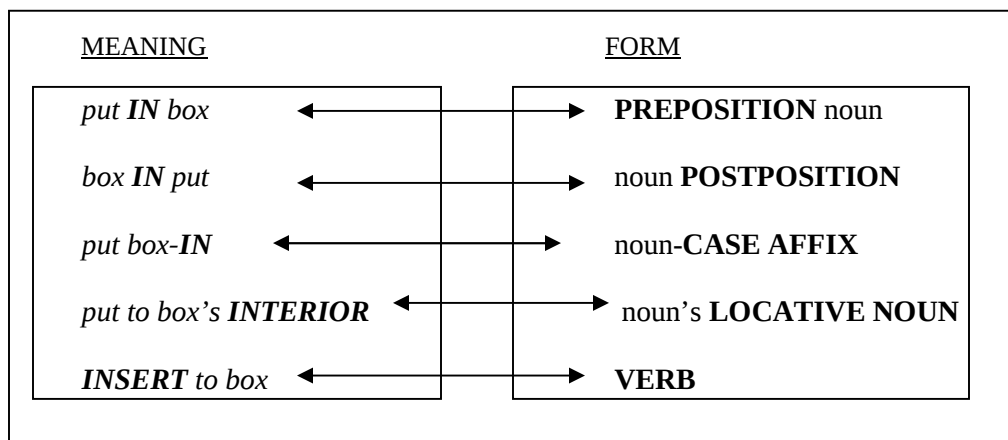


Figure 4. Diversity in grammatical expression of a domain: Various ways of encoding ‘put (something) in a box’

There clearly are things that very young children notice and want to talk about, and these semantic domains must be the starting point for acquiring forms of linguistic expression. For example, very young children always talk about transitive actions, especially those sorts of highly salient actions that I have called “the Manipulative Activity Scene” (Slobin, 1985)—namely, situations in which (a) an agent effects a change of state or location of a patient, (b) by means of direct body contact or instrument. This is an obvious starting point for acquiring language-specific means of encoding agent–patient relations. And children are quick to pick up the frequent linguistic patterns for expressing the Manipulative Activity Scene, such as English word-order patterns or accusative case inflections in Russian or Turkish. But, having established a link between a concept and a linguistic form, is the job done? Not yet: because core concepts are elaborated differently across grammars of the world’s languages. There is diversity of both form and function. Consider, for example, languages that use an explicit grammatical marker to indicate the patient argument of transitive action verbs (direct object markers). Table 3 sets forth a small part of the considerable diversity presented by the world’s languages.

Table 3. Crosslinguistic diversity in grammatical marking of the patient (direct object)

- | |
|---|
| <ul style="list-style-type: none"> • Turkish case inflection: <ul style="list-style-type: none"> ○ definite patient only • Mandarin Chinese object marker: <ul style="list-style-type: none"> ○ definite patient only ○ only patient of direct physical action • Finnish case inflection: <ul style="list-style-type: none"> ○ whole vs. partial patient ○ completed vs. non-completed action • Russian case inflection: <ul style="list-style-type: none"> ○ whole vs. partial patient ○ singular vs. plural patient ○ masculine animate vs. other ○ affirmative vs. negative clause • Georgian case inflection: <ul style="list-style-type: none"> ○ patient marking in present tense only |
|---|

It is apparent that there is a long way from a child’s attention to the Manipulative Activity Scene to identifying and acquiring just which particular features of such events receive grammatical expression in the exposure language.

In brief, there is no unique conceptual starting point for even such basic notions as location, agency, and object manipulation. It is no longer evident that there are universal starting points of both form and function that get the child started in language acquisition. With regard to the issue of semantic bootstrapping, therefore, current data and theory lead to a relativistic conclusion: ***There is no pre-established link between particular concepts and particular linguistic forms.***²

² Pinker acknowledged this problem, but without abandoning his nativist framework: “The discussions of the acquisition of languages radically departing from English has at time been perfunctory, and many technical problems and contingencies among rule systems within a typology were not even considered. There are also far more significant facts about the development of non-English-like languages than I was able to examine (see Slobin, 1985), and no doubt still more that have not yet been documented. A particularly fundamental problem for future development of the theory is how to state the semantic correlates of grammatical categories and relations in a way that does full justice to linguistic diversity, and that enables the child to acquire the languages that counterexemplify the universals that the use of these semantic correlates presupposes” (Pinker, 1984, p. 352).

4. Syntactic Bootstrapping

Another approach available to the child might be to start with syntax and take syntactic cues as guides to meaning. This procedure has been advocated by Gleitman:

“Children’s sophisticated perceptual and conceptual capacities yield a good many possibilities for interpreting any scene, but the syntax acts as a kind of mental zoom lens for fixing on just the interpretation, among these possible ones, that the speaker is expressing. ... This examination of structure as a basis for deducing the meaning is the procedure we have called ***syntactic bootstrapping***” (Gleitman, 1990, p. 3).

However, can syntax be a reliable guide to word meanings? In one broad sense, the answer is yes. A half-century ago, Roger Brown had already established that children can identify the part-of-speech of a nonsense word on the basis of the syntactic frame in which it is presented. For example, the syllable *sib* was presented along with a picture of two hands manipulating a substance in a container. If the accompanying sentence was, “This is a picture of some *sib*” children identified the ***substance*** as the referent of the unfamiliar word, but chose the ***container*** if the sentence was, “This is a picture of a *sib*.” In an article appropriately titled “Linguistic Determinism and the Part of Speech,” Brown concluded: “It was shown experimentally that young English-speaking children take the part-of-speech membership of a new word as a clue to the meaning of the word” (R. Brown, 1957, p. 5).

However, Gleitman’s proposal goes further. She suggests, for example, that a child could determine the meaning of the verb *bring* by hearing it in sentence frames such as “Bring me the doll.” Clearly, the presence of two arguments (*me* and *doll*) along with the implicit third argument of the imperative (*you*), shows the child that *bring* is a verb that can take three arguments. But the syntactic frame cannot fully determine the meaning of an individual verb. Consider a situation in which a child is sitting on the floor and holding a doll; the mother is seated at some distance from the child and says, “Bring me the doll.” The child can determine, using syntactic bootstrapping, that the unfamiliar verb *bring* means that the child should do something with the doll that affects the mother. But the syntactic frame alone is compatible with a range of possible verb meanings, such as ‘transfer to speaker’, ‘show to speaker’, or even ‘act on the doll on my behalf’ (e.g., pat the doll for me). Note, too, that many languages make frequent use of ellipsis, thereby presenting the child with expressions such as ‘bring doll’ or simply ‘bring’ in contexts in which an English-speaking mother might say *Bring me the doll*.

Syntactic frames help; but, again, we are faced with crosslinguistic diversity in the semantic elements that are linguistically relevant to encoding particular types of scenes. For example, in Tzeltal Mayan, different verbs are used depending on what body part is involved (in arms, on back, on head, etc.). In Navajo it is necessary to distinguish the shape and substance of the object that is being moved, requiring the child to attend to the following set of object characteristics in determining the meanings of verbs of object manipulation and transfer:

- a single roundish bulky object
- a large number of small objects
- a burden, pack, or load
- non-compact matter
- matter in an open container
- one slender flexible object
- one slender stiff object
- mushy matter
- a single animate object

In sum, neither semantic nor syntactic bootstrapping can provide a full account of child language acquisition.

5. The problem of the starting points

Regardless of the language-specific semantic patterns that the child must eventually learn, there must be natural starting points to the process. That is, the child must begin with *some* conceptions of referent situations and events that will be relevant for form–function mappings. Much research is needed in order to establish the appropriate level of *granularity* for describing the child’s beginning construction of semantic space. The initial level of analysis will furnish the child with basic object and event dimensions that can then be shaped by the exposure language into the particular “packages” of features that receive linguistic expression. The research briefly summarized above suggests several potential starting points in various domains. In the spatial domain, for example, starting points might be found in contrasts between CONTAINMENT and SUPPORT, and between TIGHT and LOOSE FIT. In the domain of transitive events, possible starting points are combinations of FIRST-PERSON AGENCY, CONTROL ACT, MANIPULATED OBJECT, WHOLE VERSUS PARTIAL OBJECT, and CHANGE OF STATE.

Several types of research methods are available for exploring levels of granularity of event analysis. Here we briefly consider three: infant comprehension, computational modeling, and analysis of spontaneous speech data.

5.1. Infant comprehension

Infant comprehension can be studied using the method of preferential looking. In this method, the infant watches scenes on two TV monitors while hearing a verbal stimulus. If the infant reliably looks at the monitor that presents a situation that is being verbally labeled, it is concluded that the infant comprehends the label. This is a useful method, because it makes it possible to study comprehension in quite young infants, often before they have begun to speak at all.

A critical study was carried out by McDonough, Choi, and Mandler (2001). The question was whether American infants would attend to the distinction between tight and loose fit, even though it is not systematically marked in English as it is in Korean. The participants were preverbal infants at ages 9, 11, and 14 months. In this study there was no verbal labeling at all; the infants simply watched activities on the two monitors. They were tested for recognition of tightness of fit within the containment relationship. In the first phase of the experiment, infants were simply familiarized with either loose or tight fit. For example, if an infant was being familiarized with situations of loose fit, each of the two screens would present a different instance of loose fit—for example, placement of wooden shapes in a basket on one screen and placement of blocks in a cloth bag on the other. For tight fit, the infant would see situations such as putting a book in a case and putting a cork in a bottle. (Note that these are all examples of containment in English.) There were six pairs of situations, and each infant was familiarized on either loose or tight fit. In the comprehension test, one screen showed loose fit and the other tight fit—using the same moveable objects on both screens—for example, placement of sticks in tight-fitting holes on one screen, versus placing the same sticks in plastic containers on the other. If an infant is attentive to the type of relation that occurred in the familiarization trials, then the monitor with a similar situation should attract more attention. That is, the infant should spend more time looking at the familiar situation—be it loose fit or tight fit. It was found that American infants of all three ages looked significantly longer at the familiar situation on the test trials. That is, infants hearing English are apparently attending to the distinction between loose and tight fit, though it will not end up playing a systematic role in semantic or grammatical structure. When the same test was carried out with Korean and American adults, the Koreans—like the American infants—looked longer at the familiar situation. But American adults did not distinguish between the two situations. Thus we can conclude that American babies—at least up to 14 months of age—are attending to a distinction that they will eventually learn to ignore. (Korean babies, presumably, will perform the same as American babies in that age range.) English speakers thus stop attending to the distinction between loose and tight fit as they grow up, having realized that containment/support is the relevant dimension. Korean speakers, by contrast, continue attending to the loose/tight fit distinction as they grow up. This experiment matches a much richer research tradition in the acquisition of phonology. It has been well established, for example, that American and Japanese 6-month-olds can distinguish *l* from *r*, but that this distinction is no longer available to older Japanese infants or adults. The explanation is that the distinction does not play a role in Japanese phonology, whereas it is critical in English phonology. In

both semantics and phonology, then, infants begin with the ability to make a large range of distinctions, eventually zeroing in on those distinctions that are linguistically relevant in the exposure language.

5.2. Computational modeling

Another way to establish plausible levels of granularity for acquisition is the method of simulation. An excellent example is Bailey's (1997) computational model of the role of motor control in the acquisition of action verbs. The domain chosen was verbs of object manipulation, such as English *push* and *shove*. Bailey's task was to train the model with pairs of actions and verbs and then test it for both production (appropriate labeling of novel action) and comprehension (appropriate obeying of verbal commands in novel situations). The simulation was an arm-and-hand that had to grasp and move objects of various sizes and shapes. The task was explicitly conceived as a simulation of child language acquisition, in which a child encounters verbs in the presence of activities. The model was trained on a set of English verbs, such as *push*, *slide*, *lift*, *turn*, *pull*, *press*, *shove*, *touch*, etc. It was successful in learning these verbs and applying them to new situations, with a 78% recognition rate and an 81% compliance rate. For example, when presented with the verb *push*, the simulated agent chose SLIDE if the object was small and moveable, and DEPRESS if the object was a button. What made the simulation successful was the fact that the agent was provided with motor information about the human arm and hand—at a particular level of granularity. That is, to learn the meanings of verbs of object manipulation, it is necessary to have representations of the human motor and perceptual systems. The following motor features ("embodied") had to be specified:

- hand posture: grasp, palm, index finger
- direction: toward, away, up, down, etc.
- elbow joint motion: flex, extend, fixed
- force: low, medium, high
- aspect: single/repeated action
- object size: small, medium, large
- object property: depressability, etc.

Note that these sorts of features generally don't figure in the linguistic descriptions of lexical items and grammatical morphemes—yet they may well be features at the level of analysis that is most natural to a human child, with a human body that moves and acts on a perceivable world of objects.

The model was further tested on languages that make divisions of semantic space that differ from English. It was able to learn for example, the Spanish distinction between *pulsar* 'press with index finger' and *presionar* 'press with palm', and the Tamil verb *pudi* 'use high force to clutch, hold, restrain, catch'. The level of granularity built into the model was thus applicable across languages, reinforcing the plausibility of the model as revealing a level of analysis that might underlie child language acquisition. Bailey's successful simulation suggests that the child might begin with a level of detail that is more granular than the categories of verb semantics in any exposure language, and is thereby applicable to the acquisition of any particular exposure language. In this instance, the detail is provided by representations of the body and acts of motor control; in other domains we will, no doubt, find other bases for levels of specificity of starting points.

5.3. Spontaneous speech data

Our richest source of information about child language remains naturalistic data, presented in the form of audio and video recordings and associated corpora of transcriptions. Close analysis of early acquisition can reveal the extent to which children's levels of granularity are based on universal predispositions, or are shaped by the exposure language.

Consider verbs of ingestion. It seems natural to an English speaker that the child must know, in advance, that there are three ways of taking things into the mouth, depending on the substance: if it's a solid, you can *eat* it; if it's a liquid, you can *drink* it; and if you look at some adults, they take in something that's not

solid or liquid when they *smoke*. Again, however, crosslinguistic comparison shows that this level of granularity is not universal. In Turkish, the verb *içmek* applies to both drinking and smoking—that is, taking in any non-solid through the mouth. In German, there are two verbs for eating: *essen* for humans and *fressen* for animals, along with *trinken* ‘drink’ and *rauchen* ‘smoke’. In Kalam, a language of New Guinea, there is a single verb root, *nñ-* for all types of ingestion—solid, liquid, and smoke. And in Tzeltal Mayan there is no generic verb for ingestion at all, but rather a set of quite specific verbs, such as the following: *bik* (things that are swallowed whole), *k’ux* (crunchy solids, beans), *lo’* (soft solids, fruits), *ti’* (meat), *we’* (tortillas, bread), *uch’* (liquids), *nuk’ sigarro* (smoke), and more. If there were one preferred starting point, we might expect English-learning children to have no difficulty with the basic distinction *eat/drink*, and might expect children learning the other languages to make various errors in trying to find the appropriate verbs for these two core concepts. However, P. Brown (2001) reports that Tzeltal-learning children start with correct use of the various specific verbs of eating, and do not make generalization errors that would indicate a search for a general eating verb. The level of granularity varies from language to language, and children seem prepared to accept various sorts of semantic categories, without preferring to start at a particular level of generality. With regard to Tzeltal verbs of eating we lack something like Bailey’s simulation study, so we can only speculate about the possible level of specificity of the starting points.

6. Typological Bootstrapping

As the child learns more about general characteristics of the exposure language, operating principles for **types** of form–meaning mappings develop. That is, frequency of occurrence of language-specific patterns reinforces the child’s sense of the **typology** of the exposure language. The shorthand, “typological bootstrapping,” refers to the general phenomenon that repeated use of particular learning strategies strengthens those strategies. In terms of Operating Principles (Slobin, 1985, p. 1193), the use of such principles “becomes ‘specialized’ to the particular language, in accordance with repeated attempts to solve linguistic problems in particular ways. ... As a consequence, children come to develop an overall sense of the general form and consistent patterns of the parental language...”

One example of a developing typological sense of the exposure language comes from crosslinguistic differences in the domain of motion (Talmy, 1985, 1991, 2000). The languages of the world tend to encode the path of motion in one of two ways: either in a verb (‘enter’, ‘exit’, etc.) or in an associated particle or “satellite” (‘in’, ‘out’). A simple example is provided by English and French:

- (1) a. The dog went **into** the house.
- b. *Le chien est **entré** dans la maison.*
 ‘The dog **entered** the house.’³

English “frames” path by means of a satellite (*in*); French “frames” path by means of a verb (*entrer*). English is a “satellite-framed” language (S-language); French is a “verb-framed” language (V-language). Path is highly codable in both languages. However, the languages differ in codability with regard to another dimension of motion events—**manner** of motion:

- (2) a. The dog **ran** into the house.
- b. *Le chien est entré dans la maison **en courant**.*
 ‘The dog entered the house **by running**.’

Manner is highly codable in English, because it is carried by the main verb. Every clause requires a verb, and it is just as easy to say *go in* as *run in*. In a sense, English-speakers get manner “for free,” and make widespread communicative and cognitive use of this dimension. In French, by contrast, manner is an adjunct—an optional addition to a clause that is already complete. French-speakers indicate manner when it is at issue, but otherwise do not mention it. Satellite-framed languages include the Germanic, Slavic, and Finno-Ugric languages; verb-framed languages include the Romance, Semitic, and Turkic languages, as

³ The Latinate form of 1b is available in English, but is not the everyday expression. Typological claims here are based on the most frequent and habitual forms of expression provided by the exposure language.

well as Japanese, Korean, and Basque. Various sorts of evidence indicates that learners and users of verb-framed languages pay less attention to manner of motion, in comparison with users of satellite-framed languages (Slobin, 1996a, 1996b, 2000, 2003, 2004, 2005). In terms of typological bootstrapping, a child learning English must repeatedly pay attention to manner in learning new lexical items, and presumably comes to expect that this dimension of experience will be linguistically relevant. That is, the child becomes sensitive to the satellite-framed typology of the exposure language in the course of repeated attempts at the acquisition of lexical items and their associated grammatical constructions. Children learning satellite-framed languages use a large manner verb lexicon in the preschool period. For example, British, American, and Australian preschoolers (age 2–5) in the available CHILDES corpora for English use the following 34 types of verbs of manner of self-movement: *bump, chase, climb, crawl, creep, dance, float, flop, fly, hike, hop, jog, jump, march, paddle, pounce, race, roll, run, rush, scoot, skip, slide, slip, sneak, step, swim, tread, trip, trot, walk, wiggle*. By contrast, Spanish, French, and Italian preschoolers in CHILDES corpora use a limited set of such verbs, almost all of them relatively “non-expressive” in relation to English—mainly the equivalents of *climb, dance, fly, jump, run, swim, walk* (Slobin, 1997a). That is, while children exposed to a satellite-framed language are learning to distinguish expressive nuances of manner—such as *hop* versus *jump*, or *hike, jog, race, run, trot*—children exposed to a verb-framed language are learning broad categories of basic types of motor patterns, such as *run* versus *walk*. As a consequence, it seems reasonable to conclude that S-language children have been guided by the typology of the exposure language to pay attention to manner of motion and to construct a set of systematic semantic categories in this domain.

In related research, Naigles and co-workers have found evidence for typological preferences in the learning of new words in experimental contexts. English- and Spanish-speaking adults were presented with novel motion verbs in situations in which the verb could refer to either path or manner of motion. Naigles and Terrazas (1998) found that English speakers were more likely to attribute manner meanings to novel verbs, while Spanish speakers were more likely to attribute path meanings. Hohenstein (2001) replicated these findings for monolingual English- and Spanish-speaking 7-year-olds (but not for 3-year-olds). These findings suggest that, in learning a language, the child develops expectations about the dominant lexicalization patterns of the language, and uses these expectations as the basis of acquiring the meanings of new lexical items. Naigles et al. (1998, p. 547) suggest that language-specific lexicalization patterns should enable children “to *fast-map*, or quickly and accurately associate a new verb with its meaning.”

Another example comes from children’s acquisition of Tzeltal Mayan verbs. The language favors a high level of granularity, what P. Brown (2001) calls “verb specificity.” We have already seen this with regard to verbs of ingestion, and it applies to many other domains. Consider verbs of holding/carrying: the child must learn to make distinctions such as *pet* (in both arms), *kuch* (weight on head/back), *k’ech* (weight across shoulders), *lik* (in hand, supported from top), *tuch’* (vertically extending from hand), *lut* (in mouth), and several more. In the set of motion verbs, the child is faced with a number of distinctions, including *tal* (toward self), *ba* (away from self, or neutral direction), *och* (inwards), *lok’* (outwards), *jelaw* (crossways), *jajch* (getting up from sitting/lying position), and more. P. Brown reports that Tzeltal Mayan children quickly learn verbs at the levels of granularity indicated by these examples. She proposes that semantic specificity in the verb lexicon is a typological feature of Mayan languages and that children easily become sensitive to this characteristic in acquiring new verbs. Her “verb specificity hypothesis” is a clear claim for typological bootstrapping:

The proposal developed here is that the language Tzeltal children hear provides many different verb labels for a given domain of activity, and thereby affects the hypotheses they bring to bear on what new verbs can mean” (P. Brown, 2001, p. 536).

The child thus comes to expect that certain types of information will be encoded in the exposure language: (a) at a particular level of granularity, and (b) in a particular sort of lexical item or construction.

7. Learning to think for speaking

In the process of acquiring a number of specific form–function mappings, the child begins to operate in terms of general patterns in the exposure language. Because these patterns vary from language to language,

linguistically-relevant cognition is shaped in the course of acquisition. This process was already evident to Roger Brown, based on his experiment that showed that English-speaking children know the semantic categories associated with parts of speech, such as count nouns and mass nouns. He concluded (1957, p. 5): “It seems likely that speakers of languages other than English will also have detected the semantic characters of their parts of speech. There is a sense, then, in which this grammatical feature of a language affects the cognition of those who speak the languages.” He formulated the learning situation in terms of “the Original Word Game” (R. Brown, 1958a, 1958b) in which: “We take a new word as a lure for cognition... For the player of the Original Word Game a speech invariance is a signal to form some hypothesis about the corresponding invariance of referent” (R. Brown, 1958b, p. 210). In the current framework, “speech invariance” includes not only lexical items, but also grammatical morphemes and syntactic constructions.

More broadly, in the process of learning any particular language, one learns how to think in terms of that language’s grammatical and lexical categories—that is, one learns to “think for speaking” (Slobin, 1987, 1991, 1996a). In acquiring a language, the child’s attention is directed to the characteristics of objects and events that are regularly encoded in the grammar and lexicon. Accordingly, learning a language is learning how to think in the terms of that language in order to speak and to understand. Frequent occurrences of verbs that specify ways of moving or ingesting act as “lures to cognition,” sensitizing the learner to the dimensions and features of moving or ingesting.

8. Conclusions

As the child acquires lexical items, grammatical morphemes, and linguistic constructions, the necessary semantic meanings are assembled from prelinguistic concepts. We are only beginning to understand the levels of specificity and content of such concepts. From the start, the exposure language plays a role in directing the child’s attention. Consequently, the child determines appropriate mappings between form and meaning on the basis of patterns in the exposure language. This requires attention to the situations and activities that co-occur with speech, along with embodied representations of motor behavior. As the child acquires more form–meaning mappings, the language comes to organize itself along lines of underlying typology. The result is an individual who has learned to analyze events to and attend to dimensions of experience in a language-specific way—someone who both speaks and thinks in terms of a particular language.

As Roger Brown (1958b) put it in his formulation of the Original Word Game:

The bounds of verbal categories are set by human beings. They are not built into the apparatus but are a cultural acquisition. ... Speech, therefore, is the principal instrument of cognitive socialization.

In conclusion, children learn to **think for speaking** and to **listen for understanding** in terms of the language (or languages) to which they are being exposed. Many different sorts of “bootstrapping” capacities must function together to make this achievement possible.

REFERENCES

[Items marked with * are available for download at <http://ihd.berkeley.edu/slobin.htm>]

- Bailey, D. R. (1997). *When push comes to shove: A computational model of the role of motor control in the acquisition of action verbs*. ICSI Technical Report tr-97-041. International Computer Science Institute, Berkeley, CA.
- Berman, R. A., & Slobin, D. I. (1994). *Relating events in narrative: a crosslinguistic developmental study*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Bowerman, M. (1985). What shapes children’s grammars? In D. I. Slobin (Ed.), *The crosslinguistic study of language acquisition. Vol. 2: Theoretical issues* (1257-1319).
- Bowerman, M. (1996). The origins of children’s spatial semantic categories: cognitive versus linguistic determinants. In J. J. Gumperz & S. C. Levinson (Eds.), *Rethinking linguistic relativity* (pp. 145-176). Cambridge: Cambridge University Press.

- Bowerman, M., & Choi, S. (2001). Shaping meanings for language: universal and language-specific in the acquisition of spatial semantic categories. In M. Bowerman & S. C. Levinson (Eds.), *Language acquisition and conceptual development* (pp. 475-511). Cambridge: Cambridge University Press.
- Brown, P. (2001). Learning to talk about motion UP and DOWN in Tzeltal: is there a language-specific basis for verb learning? In M. Bowerman & S. C. Levinson (Eds.), *Language acquisition and conceptual development* (pp. 512-543). Cambridge, MA: MIT Press.
- Brown, R. (1957). Linguistic determinism and the part of speech. *Journal of Abnormal and Social Psychology*, 55, 1-5.
- Brown, R. (1958a). How shall a thing be called? *Psychological Review*, 65, 14-21.
- Brown, R. (1958b). *Words and things*. Glencoe, IL: Free Press.
- Choi, S., & Bowerman, M. (1991). Learning to express motion events in English and Korean: The influence of language-specific lexicalization patterns. *Cognition*, 41, 83-121.
- Choi, S., McDonough, L., Bowerman, M., & Mandler, J. M. (1999). Early sensitivity to language-specific spatial categories in English and Korean. *Cognitive Development*, 14, 241-268.
- Johnston, J. R., & Slobin, D. I. (1979). The development of locative expressions in English, Italian, Serbo-Croatian and Turkish. *Journal of Child Language*, 6, 531-547.
- Gleitman, L. R. (1990). The structural sources of verb meanings. *Language Acquisition*, 1, 3-55.
- Hohenstein, J. M. (2001). *Motion event similarities in English- and Spanish-speaking children*. Unpublished doctoral dissertation, Yale University.
- Levinson, S. C. (2003). *Spatial language and cognition*. Cambridge: Cambridge University Press.
- Levinson, S.C., Meira, S., & The Language & Cognition Group (2003). 'Natural concepts' in the spatial typological domain – Adpositional meanings in cross-linguistic perspective: An exercise in semantic typology. *Language*, 79, 485-516.
- Maratsos, M. (1998) The acquisition of grammar. In D. Kuhn & R. Siegler (Eds.), *Cognition, perception and language (Vol. II)*, in W. Damon (Series Ed.), *Handbook of child psychology* (pp. 421-466). Wiley.
- McDonough, L., Choi, S., & Mandler, J. M. (2003). Understanding spatial relations: flexible infants, lexical adults. *Cognitive Psychology*, 46, 229-259.
- Naigles, L. R., & Terrazas, P. (1998). Motion-verb generalizations in English and Spanish: Influences of language and syntax. *Psychological Science*, 9, 363-369.
- Özçalışkan, S., & Slobin, D. I. (1999). Learning "how to search for the frog": Expression of manner of motion in English, Spanish, and Turkish. In A. Greenhill, H. Littlefield, & C. Tano (Eds.), *Proceedings of the 23rd Annual Boston University Conference on Language Development* (pp. 541-552). Somerville, MA: Cascadia Press.
- Özçalışkan, S., & Slobin, D. I. (2000). Expression of manner of movement in monolingual and bilingual adult narratives: Turkish vs. English. In A. Göksel, & C. Kerslake, (Eds.), *Studies on Turkish and Turkic languages*. Wiesbaden: Harrasowitz.
- Pinker, S. (1984). *Language learnability and language development*. Cambridge, MA: Harvard University Press.
- Slobin, D. I. (1971). Developmental psycholinguistics. In W. O. Dingwall (Ed.), *A survey of linguistic science* (pp. 298-411). College Park, MD: University of Maryland Linguistics Program.
- Slobin, D. I. (1973). Cognitive prerequisites for the development of grammar. In C. A. Ferguson & D. I. Slobin, (Eds.), *Studies of child language development* (pp. 175-208). New York: Holt, Rinehart & Winston.
- Slobin, D. I. (1985). Crosslinguistic evidence for the Language-Making Capacity. In D. I. Slobin (Ed.), *The crosslinguistic study of language acquisition: Vol. 2. Theoretical issues* (pp. 1157-1256). Hillsdale, NJ: Lawrence Erlbaum Associates.
- * Slobin, D. I. (1987). Thinking for speaking. *Proceedings of the Thirteenth Annual Meeting of the Berkeley Linguistics Society*, 435-444. [Expanded version: Slobin, D. I. (1996). From "thought and language" to "thinking to speaking". In J. J. Gumperz & S. C. Levinson (Eds.), *Rethinking linguistic relativity* (pp. 70-96). Cambridge: Cambridge University Press.
- Slobin, D. I. (1996a). From "thought and language" to "thinking for speaking." In J. J. Gumperz & S. C. Levinson (Eds.), *Rethinking linguistic relativity* (pp. 70-96). Cambridge: Cambridge University Press.
- Slobin, D. I. (1996b). Two ways to travel: Verbs of motion in English and Spanish. In M. Shibatani & S. A. Thompson (Eds.), *Grammatical constructions: Their form and meaning* (pp. 195-217). Oxford: Oxford University Press.
- * Slobin, D. I. (1997a). Mind, code, and text. In J. Bybee, J. Haiman, & S. A. Thompson (Eds.), *Essays on language function and language type: Dedicated to T. Givón* (pp. 437-467). Amsterdam/Philadelphia: John Benjamins.
- Slobin, D. I. (1997b). The origins of grammaticizable notions: Beyond the individual mind. In D. I. Slobin (Ed.), *The crosslinguistic study of language acquisition: Vol. 5. Expanding the contexts* (pp. 265-323). Mahwah, NJ: Lawrence Erlbaum Associates.
- Slobin, D. I. (2000). Verbalized events: A dynamic approach to linguistic relativity and determinism. In S. Niemeier & R. Dirven (Eds.), *Evidence for linguistic relativity* (pp. 107-138). Amsterdam/Philadelphia: John Benjamins.

- * Slobin, D. I. (2001). Form function relations: how do children find out what they are? In M. Bowerman & S. C. Levinson (Eds.), *Language acquisition and conceptual development* (pp. 406-449).
- * Slobin, D. I. (2003). Language and thought online: Cognitive consequences of linguistic relativity. In D. Gentner & S. Goldin-Meadow (Eds.), *Language in mind: Advances in the investigation of language and thought* (pp. 157-191). Cambridge, MA: MIT Press.
- * Slobin, D. I. (2004). The many ways to search for a frog: Linguistic typology and the expression of motion events. In S. Strömqvist & L. Verhoeven (Eds.), *Relating events in narrative: Typological and contextual perspectives*. Mahwah, NJ: Lawrence Erlbaum Associates.
- * Slobin, D. I. (2005). Relating events in translation. In D. Ravid & H. B. Shyldkrot (Eds.), *Perspectives on language and language development: Essays in honor of Ruth A. Berman* (pp.115-129). Dordrecht: Kluwer.
- Talmy, L. (1985). Lexicalization patterns: Semantic structure in lexical forms. In T. Shopen (Ed.), *Language typology and lexical description: Vol. 3. Grammatical categories and the lexicon* (pp. 36-149). Cambridge: Cambridge University Press.
- Talmy, L. (1991). Path to realization: A typology of event conflation. *Proceedings of the Seventeenth Annual Meeting of the Berkeley Linguistics Society*, 480-519.
- Talmy, L. (2000). *Toward a cognitive semantics*. Cambridge, MA: MIT Press.