Early Sensitivity to Language-Specific Spatial Categories in English and Korean

Soonja Choi
San Diego State University

Laraine McDonough
Brooklyn College

Melissa Bowerman
Max Planck Insitute for Psycholinguistics

Jean M. Mandler
University of California San Diego

This study investigates young children's comprehension of spatial terms in two languages that categorize space strikingly differently. English makes a distinction between actions resulting in containment (*put in*) versus support or surface attachment (*put on*), while Korean makes a cross-cutting distinction between tight-fit relations (*kkita*) versus loose-fit or other contact relations (various verbs). In particular, the Korean verb *kkita* refers to actions resulting in a tight-fit relation regardless of containment or support. In a preferential looking study we assessed the comprehension of *in* by 20 English learners and *kkita* by 10 Korean learners, all between 18 and 23 months. The children viewed pairs of scenes while listening to sentences with and without the target word. The target word led children to gaze at different and language-appropriate aspects of the scenes. We conclude that children are sensitive to language-specific spatial categories by 18–23 months.

241

Every event is characterized by a multitude of properties that in isolation or in various combinations could be a candidate for the meaning of a word (Gentner, 1982; Gleitman, 1990; Quine, 1960). Consider, for example, putting a book into a box-shaped cover designed to fit it. This action involves objects of a certain identity, shape, size, weight, and color, motion in a certain direction, agentive causation, an end state that features both *containment* (the book ends up in the box) and *tight fit* (the book and the box fit each other exactly), and many other characteristics. Which of these properties is picked out by an expression used to refer to the event? This depends to some extent on what the speaker wants to say about the event [cf. the difference between "I'll put it in" (agent included) and "It goes in easily" (agent omitted)]. However, it also depends critically on the language being learned.

In English, speakers typically talk about an event like this using the word *in*. This word picks out the *containment* aspect of the end state, but not the *tight fit* relation: (*put*) *in* can be used equally well for putting an object into a loose container. In Korean, speakers are likely to use the verb *kkita*, which means roughly "put something into a tight-fitting/interlocking relationship with something else." This word encodes *tight fit*, but not *containment*, since it can also be used for putting a top on a pen or two Lego pieces together (Choi & Bowerman, 1991). Korean does not have a word with a meaning comparable to that of *in*.

This example illustrates that different languages focus on different properties of events for purposes of categorizing and labeling them. Differences in the semantic categories associated with everyday vocabulary items raise important and difficult questions about semantic development. When and how do children discover which properties of an event are critical for assigning that event to a semantic category associated with a word in their local language? How does this process interact with language learners' nonlinguistic ways of perceiving and interpreting events? In this study, we explore the *when* question for events involving spatial relationships.

In the large literature on the acquisition of spatial expressions, it has been widely assumed that words like *in*, *on*, *under*, *up*, *down*, *in front*, and *behind* are mapped directly to a circumscribed set of universally shared, highly salient basic spatial concepts. Several lines of argumentation and evidence have contributed to this assumption. An early influence was the claim by investigators of adult language that spatial terms are highly constrained both by the way humans perceive space nonlinguistically and by universal environmental conditions like gravity (Bierwisch, 1967; H. Clark, 1973). To the extent that spatial words are constrained to converge on a set of privileged spatial concepts, their meanings could be expected to arise in children independently of language. Many studies have shown that infants indeed know a great deal about space before they begin to talk about it (Baillargeon, 1995; Gibson & Spelke, 1983; Levine & Carey, 1982; Needham & Baillargeon, 1993; Piaget & Inhelder, 1956), and other research has demonstrated that they draw on their nonlinguistic spatial knowledge in trying to interpret spatial words (E. Clark, 1973).

Further evidence suggesting a critical role in semantic development for a limited set of universal, nonlinguistic spatial notions is that locative terms emerge in a remarkably consistent order both within and across languages. In a crosslinguistic elicited production study, Johnston and Slobin (1979) found that children learning languages as diverse as English, Turkish, Hungarian, and Serbo-Croatian acquired words for notions of *in*, *on*, *under*, and *beside* before words for notions of *between* and *in front of*/*behind* a featured object, and they acquired these in turn before words for *in front of* or *behind* a nonfeatured object. Johnston and Slobin interpret this sequence as reflecting the order in which the relevant spatial notions mature in nonlinguistic cognition, e.g., topological concepts (*in*, *on*) before projective notions (*in front*, *behind*) (Piaget & Inhelder, 1956).

The developmental priority of topological notions has also been corroborated by longitudinal studies of spontaneous speech. Children learning English typically produce the words *in* and *on* before two years of age, and use them consistently and appropriately for containment and support/attachment (e.g., Bloom 1973; Bloom, Lightbown & Hood, 1975; R. Brown, 1973; Gopnik, 1980, 1982; Tomasello, 1987). In recent crosslinguistic work, Sinha, Thorseng, Hayashi, and Plunkett (1994) also find that words for containment and support are acquired very early, suggesting that these particular concepts play an especially privileged role.

A plausible hypothesis in view of these findings would be that when children view an action of putting a book tightly into its box-cover (for example), they conceptualize it as an event resulting in *containment*. On hearing a word like the English particle *in* or the Korean verb *kkita* applied to the event, they might then assume that the word encodes *containment*, and tend to generalize it to other events involving *containment*. For children learning English, reliance on *containment* would result in accurate learning, but for children learning Korean it would result in an incorrect mapping and subsequent errors of both overgeneralization (incorrect application of *kkita* to containment events that do not involve tight fit) and undergeneralization (failure to apply *kkita* to tight fit events that do not involve containment, like putting Lego pieces together). Critically, both sets of children would initially rely on the same conceptualization of the event, and the same mapping of words to events: Only later would they begin to diverge, with the Korean children (in this case) making the needed adjustments to their mapping. This course of development would conform to Slobin's (1985) hypothesized "Basic Child Grammar," according to which all children initially adhere to a universal and uniform starting organization of semantic space, and only later, after more linguistic experience, diverge in the direction of the semantic system of the input language.

Recent crosslinguistic research suggests that this view of the acquisition of spatial words overestimates the degree to which children's early spatial conceptualizations are tightly constrained to a limited set of highly salient notions. The evidence comes from studies of both adult grammars and language acquisition.

Studies of adult grammar show that—contrary to earlier emphasis on crosslinguistic uniformity—languages differ considerably in how they categorize spatial relations (e.g., Bowerman, 1989, 1996a, b; Bowerman & Pederson, 1992; Brown, 1994; Levinson, 1996a, b). For example, Tzeltal, a Mayan language spoken in Chiapas, Mexico, has no single word, comparable to English *in*, that picks out all kinds of containment relationships. Instead, it splits *containment* into a variety of categories, distinguishing, for example, among things in containers wider than they are tall, things in containers taller than they are wide, things contained in liquid, things that have been inserted through a narrow opening, and so on (Brown, 1994). Similarly, languages partition relationships of contact and support differently. For example, much of the domain of contact and support covered by the word *on* in English is divided up in Dutch between *op* (*on* when the located object is conceptualized as resting comfortably on an external surface, e.g., cup *on* table, sticker *on* cupboard, fly *on* ceiling), and *aan* (*on* when the object must be held in place against gravity or some other salient force working to separate it from its location, e.g., picture *on* wall, apple *on* twig, handle *on* suitcase) (Bowerman, 1996b).

To the extent that languages differ in their organization of space, there may be no one single way of construing spatial relations that is most obvious or basic for young children, so they may be open to acquiring words that pick out quite different aspects of spatial situations (Bowerman, 1989). Such flexibility in learning spatial categories would also be compatible with the view that infants may have a large set of spatial concepts available to them prelinguistically (Mandler, 1996).

Crosslinguistic studies of spatial language acquisition are consistent with the idea that learners are relatively flexible: They show that children tune in quickly to the spatial categories of their language. Several studies have shown that three- or four-year-old children learning different languages use spatial terms in a language-specific way (Bavin, 1990; Berman & Slobin, 1994; de León, 1997). In research on still earlier stages of development, Bowerman and Choi (Bowerman, 1996a; Bowerman & Choi, 1994; Choi, 1997) conducted an elicited production study with children age 2.0 to 3.5 and adult speakers of English, Korean, and Dutch. Participants were presented with a standardized set of familiar and unfamiliar objects, including Lego blocks, rings and poles, a doll and clothing items, and a suction hook that could be stuck on the wall. These stimuli were chosen because characteristic actions on them are categorized differently in the three languages. The objects were shown in either a joined or a separated state, and the participant's task was to instruct the experimenter what to do next: e.g., (*put it*) *on* for a ring being held over a pole, or (*take it*) *off* for removing the suction hook from the wall. Even the youngest children (2.0–2.5) proved sensitive to language-specific principles of categorization: through their choice of words they implicitly classified the actions significantly more similarly to adult speakers of their own language than to same-age children learning the other languages.

When does this sensitivity come in? Choi and Bowerman (1991) analyzed the spontaneous use of spatial words in learners of English and Korean from the sin-

gle-word period to about two years of age. Both English and Korean learners started to use spatial terms productively by about 16 to 20 months. Early spatial words in English included *in* and *on*, those in Korean included *kkita* ("fit tightly") and *nohta* ("put down on surface"). From this early age the children used the words appropriately or near-appropriately; for instance, English-speaking children said *in* for containment relations regardless of whether they were tight (e.g., putting a peg tightly into a hole) or loose (putting a toy into a basket), and Korean children said *kkita* for tight-fitting relations regardless of whether they involved containment (putting a piece in a puzzle) or surface contact (joining Lego pieces). Bowerman, de León, and Choi (1995) then compared these results with spontaneous speech data from children under two years old learning Tzotzil. They found that the children learning Tzotzil showed early skill with yet another language-specific spatial category: that of the Tzotzil verb *xoj*, which picks out causal actions in which an elongated object ends up encircled by a ring- or tube-shaped object (e.g., putting a ring on a pole or a pole through a ring; putting arms into sleeves, hanging a coil of rope on a peg).

Spontaneous speech data suggest, then, that children may learn language-specific spatial semantic categories before two years of age. However, it is difficult to make exact comparisons using such data because the specific contexts in which children spontaneously produce spatial words vary. The elicited production task used by Bowerman and Choi (1994) assesses children's knowledge of spatial words in a more controlled and systematic way, but the task is too demanding to use with children younger than two years.

In the present study, we explore very young children's sensitivity to language-specific spatial categories using a preferential-looking paradigm first introduced by Spelke (1976) for studying intermodal perception in infants, and later adapted for use with language stimuli by Golinkoff and colleagues (Golinkoff, Hirsh-Pasek, Cauley, & Gordon, 1987), with further modifications by Naigles (1990). This paradigm is a minimally demanding way to test language comprehension: The child need only look at scenes while listening to verbal input. Comprehension is measured by comparing looking time to the scenes that do and do not *match* the verbal input.

Our study systematically tests the comprehension by young English- and Korean-speaking children's comprehension of a key spatial word from their language: *in* in English and *kkita* ("interlock/fit tightly") in Korean. We chose these terms because they are among the first spatial words produced by learners of the two languages, and they are used by adult speakers for many of the same everyday referent situations (e.g., putting a cassette in its case, a piece in a puzzle, a block in a shape-sorting box, a peg in a hole); yet they differ significantly in meaning (Choi & Bowerman, 1991). Our goal is to determine whether by age 18 to 23 months—an age at which learners seem to understand some spatial words and may be just beginning to produce them—children show sensitivity to language-specific principles of semantic categorization.

THE CONTRAST BETWEEN "IN" IN ENGLISH
AND "KKITA" IN KOREAN

Our target words belong to different parts of speech: *in* is a particle/preposition, and *kkita* is a transitive verb, but they can still sensibly be compared. To see why, let us take a quick look at how languages characteristically express motion events (see Choi & Bowerman, 1991, for a fuller analysis).

Languages differ systematically in how they break up the elements of a motion event and assign them to different components of a clause (cf. Talmy, 1985, whose terminology we follow here). In English, along with most Indo-European languages, information about Path—the spatial relationship between a moving object (the Figure) and a reference point object (the Ground)—is characteristically expressed by particles or prepositions such as *in*, *out*, *on*, *off*, *up*, and *down.* The fact of motion itself is typically conflated (i.e., combined) with information about the manner or cause of the motion and expressed in the main verb. This division of labor is illustrated in Example 1:

(1) *Mary threw*   *the ball in   (to the box).*
     Agent [Motion + Cause] Figure   Path Ground

Korean lacks a system of particles and prepositions comparable to words like *in*, *on*, *up* and *down* in English. Along with Hebrew, Turkish, and Japanese, among other languages, it typically conflates Path with the fact of motion and expresses it in the verb. Languages of this type have a set of monomorphemic Path verbs with meanings roughly suggested by verb phrases in English like *go in*, *go out*, *go up*, *go down*, *go via*, *take out*, *put in*, *put up,* and *put down*. Information about the manner or cause of motion along one of these trajectories is expressed, if desired, with a co-verb such as *ttwui-* ("run") or *tenci-* ("throw") that precedes the Path verb. Korean is a subject–object–verb language, and the final verb takes the inflections for tense and modality. This is illustrated in Example 2:

(2)
Mary-ka   kong-ul   sangca-ey   tenci-e   neh-ess-ta.
Mary-SUBJ   ball-OBJ   box-GOAL   throw-CONNECTIVE   put.in-PAST-DECLAR
Agent   Figure   Ground   Manner   [Motion + Path]
"Mary threw the ball into the box."

In Example 2, *neh-* ("put in loosely") is the main verb. The suffix *-ey* on the Ground nominal, *sangca* ("box"), specifies that the box is the goal of the motion, but it does not encode the specific spatial configuration between the Figure and the Ground; this is the job of the main verb. A locative noun such as *an* ("interior") can be used to mark the spatial relation as well [e.g., *kong-ul sangca-an-ey tenci-e neh-ess-ta*, ball-OBJ, box-interior-GOAL throw-CONNECTIVE put.in-PAST-DECLAR ("threw the ball to the box's interior")], but this is optional and somewhat marked; the work of indicating the spatial relationship is largely left to the verb.

Learners of English and Korean start to express Path in their spontaneous single-word speech at around the same age, and they apply them to similar kinds of events (Choi & Bowerman, 1991). The forms they use for this purpose belong to the word class that in adult speech most characteristically encodes Path meanings in the language they are learning: particles/prepositions for English and verbs for Korean (see Bowerman et al., 1995, on the role of formal factors like stress and sentence position in determining which spatial words children learn first). It therefore makes sense to compare the meanings learners associate with these forms. Do they identify them with similar meanings, as would be predicted under the hypothesis that children start out by mapping spatial words to a small basic set of universal spatial concepts?

In adult English, *in* expresses a relationship of containment between Figure and Ground (Herskovits, 1986). What is important for present purposes is that this relationship is indifferent to whether the Figure fits into the Ground loosely (e.g., put the apple *in* the bowl) or tightly (put the cassette *in* the cassette case). (We restrict our examples here to actions in which an agent causes the Figure to change location, since it is in this use that we can compare the category of English *in* with the category of Korean *kkita*. For a more comprehensive comparison of the categories of English locative particles and Korean transitive and intransitive Path verbs, see Choi & Bowerman, 1991). An important semantic category contrasting with that of *in* is encoded by *on*: *on* is the preposition/particle of choice when the Figure is in contact with—typically supported by, attached to, or encircling—an external surface of the Ground (e.g., put the suitcase *on* the table, put the towel *on* the towel rack, put the ring *on* her finger). Thus, in English, a semantic distinction is made between *containment* on the one hand and contact with support, attachment, or encirclement on the other.

In adult Korean, the criterial semantic feature for the verb *kkita* is *tight fit*: The Figure ends up in a tight-fitting or interlocking relation with the Ground. The objects typically are somewhat rigid and have pre-existing complementary shapes that fit each other closely (e.g., Legos®, snaps, Pop-beads, pen and cap, cassette and case, ring and finger). As long as there is a tight-fitting relation, *kkita* is used regardless of whether the Figure ends up contained in the Ground (e.g., cassette in cassette case) or in contact with its external surface by attachment (e.g., one Lego piece on another) or encirclement (ring on finger).

The semantic category of Korean *kkita* thus crosscuts the categories of English *in* and *on*. This is illustrated further in Figure 1. For example, *kkita* is used both for putting a ring *on* a finger and a finger *into* a ring, for putting the top *on* a pen and a pen *into* its top, and for putting one Lego piece *onto* or *together* with another. The instances that fall outside the *kkita* category in Figure 1 all involve noninterlocking *on* and *in* relations, and they are expressed with various other verbs. For example: *neh-* ("put in/around loosely") is used for loose containment or encirclement, such as putting an apple in a bowl or a loosely-fitting ring on a pole; *noh-* ("deposit, put down") is used for placing objects on a roughly horizon-
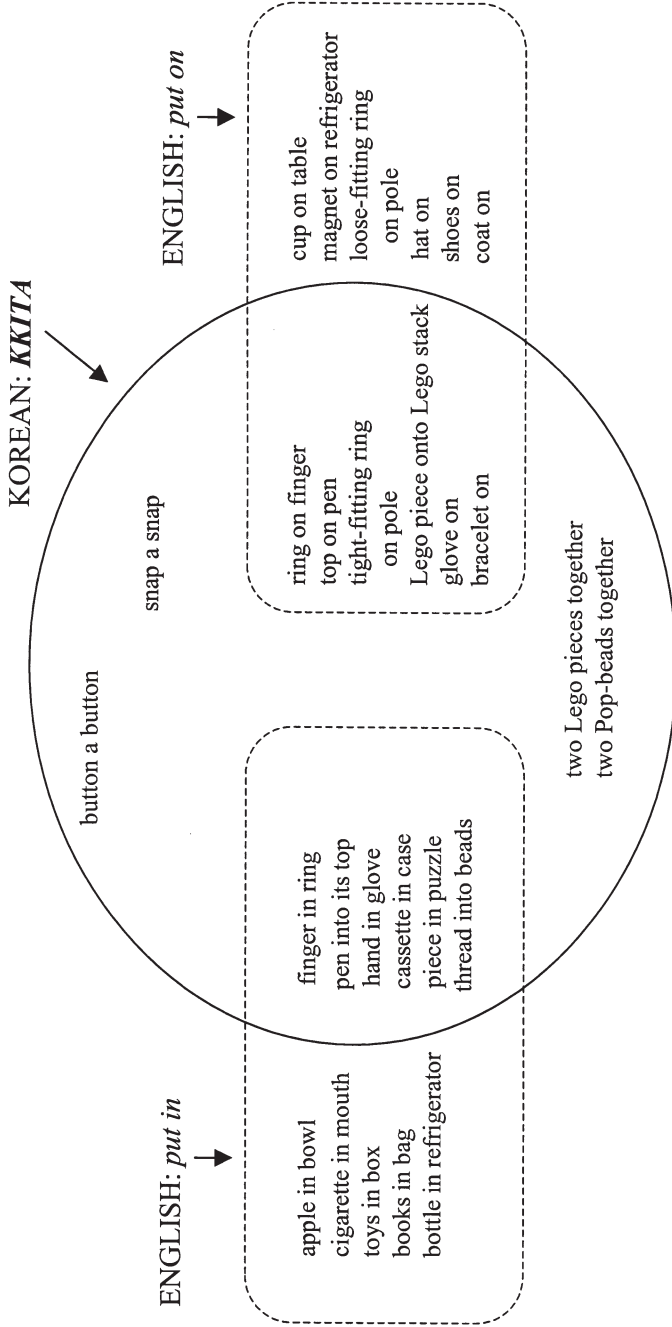
Figure 1. Cross-cutting catgeories in Korean and English: Korean *kkita* ("fit tightly/interlock") vs. English (*put*) *in* and (*put*) *on*.

KOREAN: *KKITA*

ENGLISH: *put on*

cup on table
magnet on refrigerator
loose-fitting ring
    on pole
hat on
shoes on
coat on

ring on finger
top on pen
tight-fitting ring
    on pole
Lego piece onto Lego stack
glove on
bracelet on

snap a snap

button a button

finger in ring
pen into its top
hand in glove
cassette in case
piece in puzzle
thread into beads

two Lego pieces together
two Pop-beads together

ENGLISH: *put in*

apple in bowl
cigarette in mouth
toys in box
books in bag
bottle in refrigerator

tal surface such as a cup on a table; *pwuchi-* is used for joining flat surfaces as in putting a magnet on a refrigerator or pushing two tables together, and *ssu-*, *sin-*, and *ip-* are used for putting clothing onto the head, feet, and trunk, respectively.

To summarize, in English there is a major semantic division between *containment* and contact with an external surface. In Korean, there is an equally important but cross-cutting distinction between actions that result in *tight fit* and actions that result in various other kinds of contact. How early do children begin to grasp the principle that is critical for the language they are learning?

METHOD

To assess the comprehension of the two terms in young learners of English and Korean, we used a preferential-looking paradigm that combines features of the paradigms devised by Golinkoff et al. (1987) and Naigles (1990). Like Golinkoff et al., we presented children with two video scenes shown simultaneously, and asked whether their attention was selectively drawn to the scene that *matched* a target word. Golinkoff et al. have shown that from as young as 16 months, children will look longer at the scene picked out by the target word, and this is true for both object words (e.g., *cookie*) and action words (e.g., *eat*).

In Golinkoff et al.'s (1987) paradigm, children demonstrate comprehension by looking longer at the matching screen than at the nonmatching screen. For this technique to work well, the *nonlinguistic* interest value or salience of the two scenes should be equivalent for the child. However, depending on the words to be tested, it can be difficult to insure this. In our domain, if one wants to know whether a child understands *in*, one needs to pair a scene showing, for example, an object being put into something with a scene showing the same object being put in some other spatial relationship. But we know that actions of putting an object into something are highly attractive to young children; e.g., E. Clark (1973) has shown that, given the opportunity, children from age 1.6 to 2.6 much prefer to put an object *into* another object than to put it on top or under. So more time spent looking at a scene of "putting in" would not in itself be sufficient to tell us whether the child understands *in*: this outcome could reflect a purely nonlinguistic preference. To circumvent the problem raised by nonlinguistic preferences, we adopted a modification in the experimental design used by Naigles (1990). Rather than relying on absolute looking times, we compared the children's looking behavior on a control trial, when the target word was absent, with their looking behavior on test trials, when the target word was present.

Participants

Thirty children between 18 and 23 months participated in the study (range: 18 months, 3 days to 23 months, 27 days). Twenty were learning English as their first language (mean age: 20 months, 4 days) and ten were learning Korean (mean age 20 months, 11 days). An additional six children were tested but are not

included in the analyses because of fussiness ($N = 3$), equipment failure ($N = 2$), or being rocked during the session so that reliable coding could not be obtained ($N = 1$). Since the session was short, most children found the scenes interesting.

The children were carefully screened to ensure as monolingual a background as possible. The children learning English were raised in homes where only English was spoken; if they spent time in a day care center, this facility also used only English. The children learning Korean lived in Korean families that had recently moved to Southern California. They were cared for solely by parents and other family members who spoke only Korean in the home. Korean immigrant families tend to form rather closed communities and only Korean is spoken in the stores and churches within the community.

After the test session, the parents were given a short list of words that children typically acquire in the one-word period, and were asked to check off the words they had heard their child produce at least once. The purpose of this survey was to assess whether the children had begun to produce the word to be tested, and to see whether the two language groups were similar in this regard. Only two of the ten learners of Korean and six of the twenty learners of English were reported to have produced the target word for their language at least once.

Apparatus

Our lab setup was adapted from Naigles (1990). The child sat on a parent's lap facing two 20″ video monitors, which were mounted side by side, 15″ apart, in a black panel. Between the two monitors there was a hole in the panel for a loud-speaker, through which audio input was presented during the video trials. Beneath the speaker was a ring of flickering colored lights, which was activated between trials to draw the child's gaze back to the center. In the center of the ring of lights there was a hole for the lens of a video camera, which recorded the child's gaze during the experiment; this videotape was coded and analyzed off-line. The video and audio presentations, lights, and video camera were controlled by a computer located behind the panel. The parent wore opaque glasses during the session so that he or she could not inadvertently cue the child to look at the screen that matched the target word.

Stimuli and Design

In previous studies of whether children look longer at (Golinkoff et al., 1987) or point to (Huttenlocher, Smiley, & Charney, 1983) the member of a filmed action pair that depicts a named action, each action word was represented by a single scene. This procedure shows whether children can make a global match of word to scene, but it does not reveal exactly what it is about the scene that they associate with the word: that is, it does not tell us which feature(s) of the scene they think are critical to the word's meaning.

In our study, we looked at only one word per language, but we tested these words across an array of events to explore how children understand them. Specifically, our video presentation was made up of four pairs of actions designed to test whether children understand that the relevant feature for (*put*) *in* is *containment* (English) and the relevant feature for *kkita* is *tight fit* (Korean). The scenes were professionally filmed and edited so that they ran simultaneously on two videotapes—one for each monitor—with trial duration and between-trial intervals perfectly synchronized. The action pairs are illustrated schematically in Figure 2, along with an indication of which member of each pair matches—i.e., is described by—the English and Korean target words, (*put*) *in* and *kkita.*
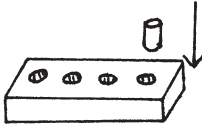
In the first and third pairs, Pegs and Books, the matching scene is the same for both languages: The Figure ends up both contained by the Ground and fitting tightly/interlocked with it, so the action is a good instance of both (*put*) *in* and *kkita.* We will term these pairs *conflated pairs,* since the properties of *containment* and *tight fit* are conflated, or combined, in the same scene. The matching scene of the Pegs pair shows pegs being put into holes in a wooden block; the nonmatching scene shows pegs being put on top of a solid block. The matching scene of the Books pair shows books being put into box-covers designed to fit them; the nonmatching scene shows books being stacked on top of each other.

In the second and fourth pairs, Legos and Rings, the properties of *containment* and *tight fit* are split up and assigned to different scenes, which means that the scene that matches the target word is *different* in the two languages. We term these *split pairs*. In the Legos pair, the matching scene for English shows Lego pieces being put into a large plastic container (*containment*), and the matching scene for Korean shows Lego pieces being added to the top of upright stacks of Lego pieces (*tight fit*). In the Rings pair, the matching scene for English shows plastic rings being put into a basket (*containment*), and the matching scene for Korean shows rings being put onto tapered plastic poles (*tight fit*).

If the English-speaking children understand that *in* picks out *containment*, the word should direct their attention to the scenes that instantiate containment regardless of whether the containment is tight (e.g., "putting books into fitted box-covers") or loose (e.g., "putting rings into a basket"). Similarly, if the Korean-speaking children understand that *kkita* picks out *tight fit*, the word should direct their attention to the scenes that instantiate tight fit regardless of whether the fit involves containment (e.g., "putting books into fitted covers") or surface attachment (e.g., "putting a Lego piece onto a stack of Legos"). On the conflated pairs, the two sets of children should have their attention directed to the same scene (e.g., "putting books into fitted box-covers") but for different reasons: the English-speaking children because the scene depicts *containment*, and the Korean-speaking children because it depicts *tight fit*. Which property children are attending to on the conflated pairs is revealed by their looking behavior on the split pairs.

Pair 1 (Conflated pair): **Pegs**
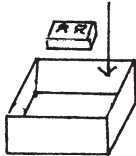
PUT PEGS IN TIGHT-FITTING HOLES   PUT PEGS ON SOLID BLOCK
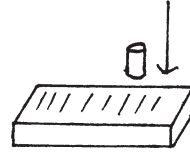
IN/KKITA[1]

Pair 2 (Split pair): **Legos**

PUT LEGOS IN BOX   PUT LEGOS ON ANOTHER LEGO

IN   **KKITA**
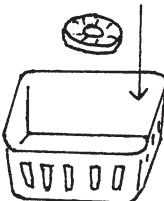
Pair 3 (Conflated pair): **Books**

PUT BOOKS IN FITTED BOX-COVERS   PUT BOOKS ON PILE OF BOOKS

IN/KKITA

Pair 4 (Split pair): **Rings**

PUT RINGS IN BIG BASKET   PUT RINGS TIGHTLY ON POLE.

IN   **KKITA**

Figure 2.   Four pairs of scenes with English (*put*) *in* and Korean *kkita* as target words. The target word is shown on the side of the scene that matches the word meaning.

Procedure

The audio and video presentations were structured in the same way for all four pairs of scenes. For each pair, three types of trial were administered: familiarization, control, and test. Each trial lasted 7 s, and was followed by a 3-s interval during which both screens were black. Table 1 gives an overview of the trials for the Books paire ("putting books into fitted box-covers" versus "putting books on top of other books").

Familiarization trials.   Children were first shown the two scenes of the pair separately. For the Books pair, "putting books into fitted box-covers" was shown (for example) on the left monitor while the right monitor remained black; then "putting books on top of other books" was shown on the right monitor while the left monitor remained black. During these trials children heard a prerecorded voice from the central speaker that encouraged them to look at the action (See Table 1 for illustrative text). We did not expect the children to answer questions posed in the audio, nor did they; the idea was simply to stimulate them to examine the scenes, and especially to draw their attention to the action. The audio for the familiarization trials and the subsequent trials started during the three-second black screen interval immediately preceding the trial, and was repeated until the trial ended. This "early start" was designed to keep children engaged in the experiment even when both screens were black and—when the target word was presented—to encourage them to examine the upcoming video pair from its onset for the action specified by the target word.

Control trial.   Following the two familiarization trials, both scenes of the pair were presented simultaneously, each on the same monitor as during the familiarization trials. The accompanying audio was neutral, in that it encouraged the child to look at both scenes (see Table 1). Recall that the purpose of the control trial is to allow us to control for possible nonlinguistic preferences for one of the two scenes in a pair. Within each pair, the control trial was always presented before the test trials so that the child's control trial looking behavior could not be influenced by prior exposure to the target word.

Test trials.   Immediately following the control trial, two identical test trials were administered. In these trials both scenes were again shown simultaneously, but this time with the addition of the target word embedded in sentences, as illustrated in Table 1.

This procedure—two familiarization trials, a control trial, and then two test trials—was followed for all four pairs. The pairs were ordered in the same way for all subjects, with an alternation between conflated and split pairs (Pegs, Legos, Books, Rings). There were two presentation conditions: For half the children in each language the location of the scene matching the target word was Left, Right, Right, Left, and for the other half it was Right, Left, Left, Right. The entire experiment lasted 3 min, 20 s.

Table 1. Sequence of Video and Audio Presentations for Each Pair of Scenes[a]

Scene type	Left screen	Audio (English version)[b]	Right screen
Familiarization trials			
[interval]	[black]	*Wow! Look at the books.*	[black]
Trial 1	PUTTING BOOKS IN	*What is she doing with the books?*	black
[interval]	[black]	*Now look. Some more books.*	[black]
Trial 2	black	*What is she doing with these books?*	PUTTING BOOKS ON
Control trial			
[interval]	[black]	*You see both of them now.*	[black]
Trial	PUTTING BOOKS IN	*Look! You see them both.*	PUTTING BOOKS ON
Test trials			
[interval]	[black]	*Where's she putting it IN?*	[black]
Trial 1	PUTTING BOOKS IN	*Find it! Where's she putting it IN? Where's she putting it in?*	PUTTING BOOKS ON
[interval]	[black]	*Look some more. Where's she putting it IN?*	[black]
Trial 2	PUTTING BOOKS IN	*That's right. She's putting it IN. She's putting it IN.*	PUTTING BOOKS ON

[a] Each 7-second trial is preceded by a 3-second interval during which the two screens are black, and the center light (located between the two screens) comes on and stays on until the next trial begins.

[b] The Korean audio for the experiment is as follows (translations are approximate):

Familiarization trials:

[Interval]	*eme! ike pwa. chayk iss-ta*	"Wow! Look at this. (There) are books."
Trial 1	*yay-ka chayk kaciko mwue hay?*	"What is she doing with the books?"
[Interval]	*eh! chayk-i tto iss-ney*	"Oh! (There) are more books."
Trial 2	*yay-nun i chayk kaciko mwue hay?*	"What is she doing with these books?"

Control trial:

[Interval]	*icey pwa. twul-i ta poi-ney.*	"Look now. Both are seen."
Trial	*twul-i ta naoko issci?*	"Both are coming on, right?"

Test trials:

[Interval]	*icey pwa. eti-ey KKI-e?*	"Look now. Where's (she) 'tight-fitting' (it)?"
Trial 1	*chaca pwa. eti-ey KKI-e? eti-ey KKI-ci?*	"Find (it). Where's (she) 'tight-fitting' (it)? Where's (she) 'tight-fitting' (it)?"
[Interval]	*com te pwa. eti-ey KKI-e?*	"Look some more. Where's (she) 'tight-fitting' (it)?"
Trial 2	*kulay maca. yay-ka KKI-ci? yay-ka KKI-e.*	"That's right. She is 'tight-fitting' (it). She is 'tight-fitting' (it)."

-ta, *-ney-*, *-ci*, *-e* are verb-final markers that express the speaker's construal of the epistemological status of the information being conveyed. *-ta* expresses new and noteworthy information, *-ney* indicates that the information is perceptually present, *-ci* indicates that the information is shared between speaker and listener, and *-e* is an unmarked form (Choi, 1995). For the English words in parentheses, there are no Korean counterparts in the audio.

Coding and Reliability

The video of each child's gaze behavior was coded off-line using a button-press box (*looks left*, *looks right*, *looks center*, *looks away*). Coders were blind to the condition in which the child was tested; that is, they did not know on which side the matching scenes were shown. All the videos were coded by at least two independent coders. Agreement was assessed on the basis of Difference scores (child's looking time to matching scene minus looking time to nonmatching scene), which were calculated independently for the control trial and the two test trials for each pair (see Results). Reliability was considered acceptable if both coders agreed on which scene was looked at longer on every trial (100% agreement) and discrepancies for Difference scores did not exceed .5 s. Discrepancies larger than this were resolved by calling in a third or fourth coder until there was agreement between at least two of the coders.

<div align="center">

RESULTS

</div>

Our central question is whether the children attributed a language-specific meaning to the target word for their language. Did the word *in* systematically direct the English group's attention to the scenes depicting *containment*, while the word *kkita* directed the Korean group's attention to the scenes showing *tight fit*?

Mean Looking Times at Matching vs. Nonmatching Scenes

We first examine how long children looked at each scene at a given trial. Table 2 shows the mean looking time (in seconds) broken down by language, scene, trial, and pair. The test trial scores for each child are calculated by averaging the time spent looking at each scene on the two presentations of the test trial for each pair (the maximum possible looking time per trial is 7 s). We will return shortly to the Difference Scores that are also shown in the table.

First, a mixed design analysis of variance was conducted with Language (English, Korean) as the between-groups factor, and Pair (Pegs, Legos, Books, Rings), Trial (Control, Test), and Scene (Matching, Nonmatching) as within-subject factors. Note that Scene refers to the matching or nonmatching scene relative to the language of the participant. On the conflated pairs (Pegs, Books) the matching scenes were the same for the two languages, while on the split pairs (Legos, Rings) they were different. The dependent measure was the time spent looking at each scene.

What do we predict? If children understand the target word for their language, they should spend more time looking at the matching scene, relative to the nonmatching scene, during the test trials than during the control trials. Thus, there should be an interaction between the variable Trial (Control, Test) and the variable Scene (Matching, Nonmatching), and this interaction should be consistent

across languages and pairs. This predicted interaction between trial type and scene was indeed obtained [$F(1, 28) = 5.02$, p $< .05$][1]; it is displayed in Figure 3.

To determine the locus of the interaction, we conducted one-way ANOVAs in accordance with the procedure suggested by Keppel (1982). During the control trials (target word absent), the children showed no consistent overall preference for either the matching or the nonmatching scene [mean looking time at Matching = 2.86 s; at Nonmatching = 2.54 s; $F(1, 29) = 0.65$, $p = .43$]. In contrast, during the test trials (target word present), they preferred the matching scenes [2.96 vs. 2.16; $F(1, 29) = 14.26$, p $< .01$]. Thus, our prediction is met.

We then compared the test and control trials directly to get a more detailed picture of how the target word directed the children's attention. The time spent looking at the matching scenes was not significantly different for the two sets of trials [test 2.96, control 2.86; $F(1, 29) = 0.41$, p $= .53$]. But the time spent looking at the nonmatching scenes was significantly lower for test than control trials [test 2.16, control 2.54; $F(1, 29) = 11.19$, p $< .01$]. This means that as the experiment progressed from control to test trials within a pair, the children maintained their attention to the matching scene at the same level, but—under the influence of the target word—significantly decreased their attention to the nonmatching scene. These findings are consistent across languages and pairs.

In devoting the same amount of attention to the matching scene on the test trial as on the preceding control trial, while allowing attention to the nonmatching scene to decline, the children showed an overall decrease in looking time between the control and the test trials within a pair. This is reflected in a main effect for Trial. During the control trials, the children examined the two scenes for an average of 5.40 s out of a possible 7, whereas during the test trials they examined them for only 5.12 s [$F(1, 28) = 8.63$, p $< .01$]. A decrease in overall looking time is not surprising, given that the same scenes were shown repeatedly on the

---

[1]In addition to this interaction and a main effect for Trial, which will be discussed shortly, there were significant main effects for both Language and Scene and an interaction between Language and Scene. These effects are not easy to interpret, but they do not bear on our experimental questions so we do not consider them further here. The main effect for Language shows that overall—i.e., collapsing over both Scene (Matching, Nonmatching) and Trial type (Control, Test)—the English learners looked longer at the scenes than the Korean learners did. The main effect for Scene shows that overall—collapsing over both Language and Trials (Control, Test)—the children looked longer at the matching scenes than the nonmatching scenes. The interaction between Language and Scene indicates that the main effect for Scene was not the same for the two languages: the preference for the matching scene, collapsing over Trials type (Control, Test), held only for the Korean learners; the English learners looked equally long at scenes of the two types. Analyses that collapse over the levels of Scene, Trial, or both are not useful for our purposes because they tell us nothing about whether the children were sensitive to the meaning of their target word (for which the contrast between matching and nonmatching scenes during the Test trials is essential) or whether they behaved differently when they heard the target word than when they did not (for which attention to matching vs. nonmatching scene on Test vs. Control trials is critical).

Table 2.   Amount of Time Looking (in Seconds) at Matching and Nonmatching Scenes by Pair and Language.

| Language | Pair | Control Trials | | | Test Trials | | |
| | | Scene | | Difference | Scene | | Difference |
		M	NM	M − NM	M	NM	M − NM
English	Pegs	3.33	2.34	0.99	3.23	2.17	1.06
(N = 20)	Legos ®	2.66	3.21	−0.56	2.64	2.62	0.02
	Books	3.43	2.08	1.35	3.59	1.62	1.97
	Rings	1.80	4.28	−2.48	1.92	3.75	−1.83
Korean	Pegs	3.41	1.23	2.18	2.73	1.87	0.86
(N = 10)	Legos ®	1.60	3.61	−2.01	2.70	2.35	0.35
	Books	3.16	1.87	1.29	3.20	1.68	1.52
	Rings	3.52	1.72	1.80	3.66	1.25	2.41
Total		22.91	20.34	—	23.67	17.31	—
Average		2.86	2.54	0.32	2.96	2.16	0.80

Difference scores were calculated by subtracting (the scores for time spent looking at) nonmatching from matching. M = Matching scene; NM = Nonmatching scene.

two screens. But despite a slight waning of attention the children spent about 73% of the duration of the test trials examining the scenes, so they were still quite engaged in the task.

Difference Scores: Test vs. Control trials

Table 2 also shows the mean difference scores, which were calculated trial-by-trial for each child by subtracting the time spent looking at the nonmatching scene from the time spent looking at the matching scene. The difference scores give us a more direct way of evaluating how the target word influenced children's looking behaviors, since they reveal the direction of looking preference on a particular trial: A positive score means a preference for the matching scene, whereas a negative score means a preference for the nonmatching scene. With difference scores, we can make a more straightforward comparison of performance on test vs. control trials, and evaluate performance on the different pairs.

Consistent with the preceding analysis of mean looking times, a three-factor ANOVA (Language × Pair × Trial) using difference scores as the dependent variable showed a main effect for Trial [$F(1, 28) = 5.02$, p < .05]: The difference scores for the test trials ($M = 0.80$) were significantly more positive than the difference scores for the control trials ($M = 0.32$). This means that both English and Korean learners showed a greater preference for the matching scene when they heard the target word than when they did not.

We illustrate this finding with two pairs of scenes that show the crosslinguistic differences most clearly: the conflated pair Books, and the split pair Rings. Figure 4 shows the difference scores for the two pairs by trial type and language.
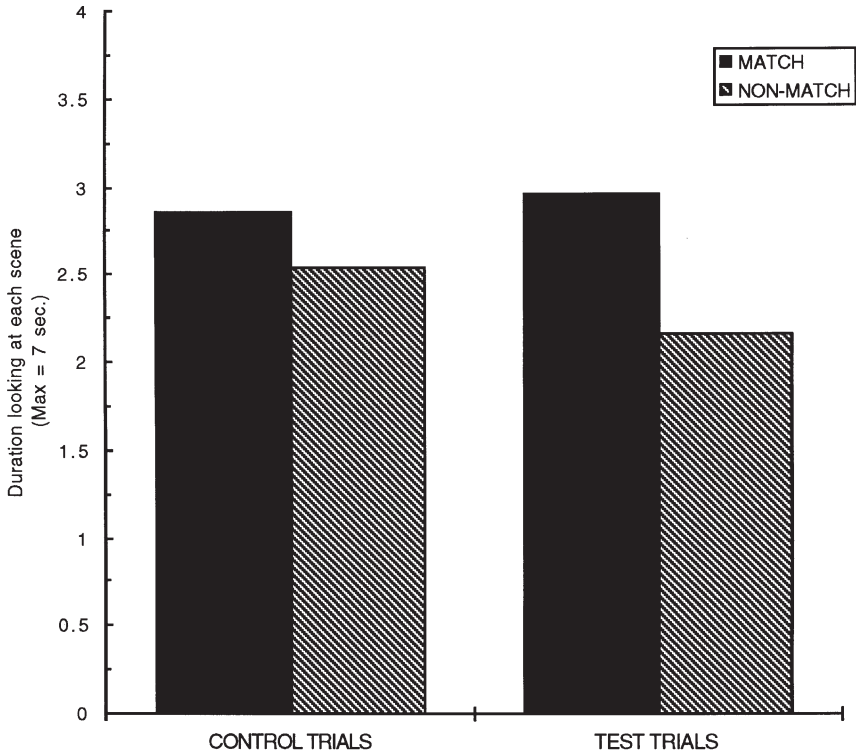
Figure 3. **Looking time in seconds at matching and nonmatching scenes during control and test trials (English and Korean learners combined).**

Since Books is a conflated pair, the matching scene is the same for both languages: "putting books into fitted box-covers." (The competing scene is "putting books on top of each other.") The difference scores for the control trial are positive for both languages, which means that both groups of children had a baseline preference for the matching scene. This is consistent with evidence mentioned earlier that children have a nonlinguistic preference for putting things in over putting things on top of other objects. But, critically, the difference scores are even more positive on the test trial. This means that both groups of children showed an even greater preference for the matching scene when they heard the target word (test trials) than when they did not (control trial).

But what was it about "putting books into fitted box-covers" that the children associated with the target words? The Rings pair offers a critical clue. Since Rings is a split pair, the matching scene for the two languages is different: "putting rings into a basket" for English; "putting rings on poles" for Korean. For this

pair, the difference scores are negative for English and positive for Korean for both test and control trials (see Fig. 4). This means that both sets of children again had an overall preference for the same scene: "putting rings on poles" (nonmatching for English, matching for Korean). But notice that although the English group's difference score on the test trial is negative, indicating a preference for the nonmatching screen, it is not as negative as on the control trial. This means that hearing the target word counteracted the English group's strong non-linguistic preference for the nonmatching scene: when these children heard their target word, *in*, their preference for "putting rings on poles," while still evident, diminished in favor of "putting rings into baskets." In striking contrast, the effect on the Korean group of hearing their target word, *kkita*, was to strengthen their pre-existing preference for "putting rings on poles." The difference score for this
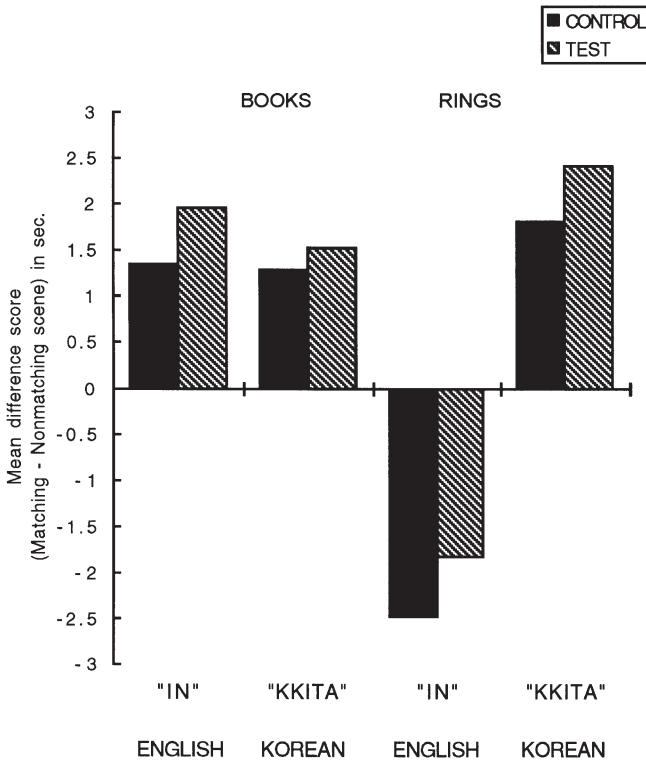
Figure 4. Difference scores (looking time at matching minus nonmatching scenes) for the Books and Rings pairs by trial type (control, test) and language (English, Korean).

pair is even more positive for the test trial than for the control trial. Thus, while both groups of children had the same baseline preference, the target word acted against this preference for the English group, while it reinforced the preference for the Korean group. The data from the Rings pairs suggest that although the two sets of children attended to the same action in the Books pair (i.e., "putting books into fitted box-covers") under the influence of their target word, they did so for different reasons: the English group because the action exemplified *containment*, and the Korean group because it instantiated *tight fit*.

Although neither pair nor language entered into a significant interaction in our ANOVAs, some differences among the pairs are evident in Table 2. The difference scores for the Legos, Books, and Rings pairs show the predicted pattern in both languages, but those for Pegs do not: for the English learners the difference scores on the control and test trials are almost the same, while for the Korean learners the scores actually decreased from control to test trials. The poor performance on the Pegs task may have been due to lack of perceptual saliency of the critical aspects of the stimuli.[2] Both the block and the pegs were made of unpainted wood, and the children may not have noticed the holes in the block. The holes were also relatively shallow, so that half of each peg still stuck out after it was inserted. The children may have interpreted both scenes as involving support, but with shorter pegs in one scene than the other.

Proportion of Children with Predicted Pattern

In our final analysis, we calculated the proportion of children learning each language who conformed to the predicted looking pattern (i.e., more attention to matching than nonmatching scene during test than control trials). The results are shown in Table 3. Most of the Korean group showed the predicted pattern for Legos, Books, and Rings, while most of the English group showed the predicted pattern for Books and Rings. These distributions depart significantly from chance ($p < .05$, binomial analysis) except for Books in Korean ($p = .11$), which may fail to reach significance due to the small Korean sample size. As just discussed, neither group did well on Pegs. On Legos, the English group looked relatively longer on average at the matching scene during test than control trials (cf. Table 2), but only slightly more than half of the children showed the predicted pattern. We will consider a possible reason for the difference between the Korean and English groups in the Discussion.

---

[2] It is also possible that the poor performance on pegs was an order effect, since Pegs was the first pair presented to all the children. To explore this and other issues concerning stimulus design, we have since had new tapes made using a different ordering of the pairs (Books, Rings, Pegs, Legos) and tested a small number of English-speaking children. The results of the new ordering replicate the results reported here, which suggests that the problem with the Pegs pair is something substantive, not the order of presentation.

Summary

The English and Korean target words used in this preferential looking study classify spatial manipulations according to cross-cutting principles of categorization. The study shows that children learning the two languages have begun to acquire sensitivity to these language-specific categories by the age of 18 to 23 months: They devote a greater proportion of their total looking time to the language-appropriate matching scene when they hear their target word for their language than when they do not. The majority of children in both languages showed the predicted looking pattern for three of the four action pairs.

DISCUSSION

In this paper, we have asked how early children show sensitivity to language-specific semantic categories of space. In an earlier study using an elicited production technique, Bowerman and Choi (Bowerman, 1996a; Bowerman & Choi, 1994; Choi, 1997) showed that learners of English, Korean, and Dutch categorize spatial events differently for purposes of talking about them by at least 2.0–2.5 years of age. Spontaneous speech data from children learning English and Korean had suggested that learners might be sensitive to language-specific principles of spatial categorization at an even earlier age—perhaps by as early as 1.6 to 2.0 years (Choi & Bowerman, 1991)—but controlled experimental evidence was lacking for these younger children. This left open the question whether learners everywhere might still associate their first spatial words with a small set of universally shared concepts derived directly from their perceptual and cognitive understanding of space.

Children's initial reliance on such a set would be predicted by Slobin's (1985) "Basic Child Grammar" hypothesis, as well as by the Piaget-inspired research tradition according to which spatial words are mapped directly to a limited set of spatial notions built up on a nonlinguistic basis during the first years of life (e.g., Johnston & Slobin, 1979). Prime candidates for membership in such a set have

Table 3. **Percent of Children in Each Language Group Who Looked Longer at the Matching Scene during Test than Control Trials.**

	Pair			
Language	Pegs	Legos	Books	Rings
English *in* (N = 20)	50%	55%	70%	70%
Koren *kkita* (N = 10)	40%	80%	70%	80%

been the concepts of containment and contact/support. But recent research on adult grammars has shown that spatial classification is more varied across languages than had previously been recognized. For instance, although containment and contact/support are the critical underlying concepts for spatial uses of *in* and *on* in English, these notions are often broken down or crosscut in other languages by concepts emphasizing other properties of spatial configurations. Crosslinguistic variation in spatial classification suggests a corresponding flexibility in human spatial cognition. Perhaps, then, infants do not approach the language-learning task with strong pre-established biases toward just a few basic spatial concepts, but instead are able to appreciate a wide variety of similarities and differences among spatial situations. In this case they might be able to home in on language-specific principles of spatial categorization from very early on in the acquisition process.

The present study explored the timing of sensitivity to language-specific semantic categories with a crosslinguistic design that compared the comprehension of a target spatial word in children 18 to 23 months old who were learning English or Korean. For English the target word was *in*, and for Korean it was *kkita* ("to fit tightly, interlock"). Although these words can be applied to many of the same referent situations, their meanings are language-specific: *in* specifies a *containment* relation that is indifferent to whether the Figure is contained tightly or loosely by the Ground, while *kkita* specifies an interlocking, *tight-fit* relation that is indifferent to whether the Figure ends up contained by the Ground, or attached to or closely encircling its outside surface.

We used a preferential looking technique in which children heard simple sentences containing the target word while viewing pairs of video-taped actions only one of which instantiated ("matched") the meaning of the word. In two pairs the scene matching the target word was the same for both sets of children (these were termed "conflated pairs," since the matching scene combined the criterial features of *containment* and *tight fit*). In two other pairs the matching scene was different (these were termed "split pairs," since the criterial features were separated and assigned to different scenes). Evaluating children's looking preferences across pairs of both kinds allowed us to determine whether learners associated their target word with the language-appropriate feature.

When they heard the target word, English and Korean learners indeed attended systematically to different and language-appropriate aspects of the spatial relations shown on the videos. For both conflated and split pairs, the English learners spent more of their total looking time on the scenes that depicted a containment relation when they heard *in* than when they did not. Most strikingly, they favored "putting books tightly into box-covers" for Books (a conflated pair), and "putting rings loosely into a big basket" for Rings (a split pair), thus focusing on containment at the expense of tight fit. Conversely, the Korean learners spent more of their total looking time on the scene that showed a tight-fit relation when

they heard *kkita* than when they did not. Thus, they preferred "putting books tightly into box-covers" on Books, and "putting rings tightly onto a pole" for Rings, thus concentrating on tight fit and disregarding containment. The crosslinguistic difference in looking behavior on the split pairs suggests that when the target word drew the attention of the two groups of children to the same scene in the conflated pairs, it did so for different reasons: For the English learners it was because the word drew their attention to the *containment* relation, while for the Korean learners it was because the word pointed to the *tight fit* relation.

The looking behavior of the two sets of children differed only on the test trials when the target word was presented. During the control trials, when the target word was absent, their looking behavior was the same. This means that it was the children's understanding of the target word that guided their attention to particular scenes during the test trials, not simply a nonlinguistic preference for these scenes. We can conclude, then, that children associate *in* and *kkita* with language-specific semantic categories from as early as 18 to 23 months.

Some pairs of scenes worked better than others. Only a few children in either group showed the predicted looking pattern on Pegs (a conflated pair), and only the Korean children performed well on Legos (a split pair). We suggested that the poor performance on Pegs may have been caused by the low perceptual salience of the critical property of the Ground object: that it contained holes into which the pegs were put tightly. Why the two sets of children differed on the Legos pair is unclear, but the crosslinguistic elicited production task mentioned earlier (Bowerman, 1996a; Bowerman & Choi, 1994; Choi, 1997) offers some clues. In that study, both the English- and the Korean-speaking children in the youngest age group (2.0–2.5) could appropriately label actions of putting objects, including Lego® pieces, into loose containers [*in* in English, *nehta* ("put loosely into/around") in Korean]. But only the Korean children could appropriately label actions of joining or stacking Lego® pieces [*kkita* ("fit tightly, interlock"; see Choi, 1997, for discussion]. These data suggest that the two actions comprising the Lego® pair in our comprehension study "putting Legos into a container" and "putting Lego pieces onto stacks of Legos"—were not semantically equally distinct for the two sets of children: for two-year-old Korean children, each of the actions is associated with an appropriate label, but for two-year-old English-speaking children, only "putting Lego pieces into a loose container" is easy to categorize. Learners of English may have known that "putting Legos into a container" was an instance of (*put*) *in*, but when this action was paired with an action for which they did not yet have a name they were puzzled and looked longer at the problematic action.

To our knowledge, our study is the first to demonstrate that children's comprehension of spatial terms is guided by language-specific principles of semantic categorization by as early as 18–23 months. But the finding is compatible with the results of several recent studies showing that there is a close relationship between caregivers' linguistic input and more global properties of children's early

lexical development. For example, the caregivers of children learning Korean and Mandarin Chinese provide more verbs and talk more about activities than the caregivers of children learning English, and, congruent with this, Korean- and Mandarin-speaking children begin to acquire verbs earlier than English-speaking children (Choi, 1998; Choi & Gopnik, 1995; Tardif, 1996; see also Goldfield, 1993; Hoff-Ginsberg, 1985; Hoff-Ginsberg & Shatz, 1982; Huttenlocher, Haight, Bryk, Seltzer, & Lyons, 1991 for further relevant findings).

Exactly how children's experience with a particular language results in sensitivity to language-specific ways of classifying space remains unclear. One possibility is that, just as in the traditional view, children do map spatial words directly to pre-established spatial concepts—it is just that this set of notions is larger and more varied than has typically been supposed; for example, in addition to "containment" and "support," children might also have a prelinguistic notion of "tight fit" (Mandler, 1996). Early semantic development would then consist of figuring out which spatial concepts from the available stock are the ones needed for language being learned. Another possibility is that children do not necessarily develop all the needed concepts nonlinguistically, but can construct them on the basis of the linguistic input (Bowerman, 1996a; Bowerman & Choi, in press; Landau & Jackendoff, 1993; Regier, 1995; see also Brown, 1958, Chapter 6, on how a word could serve as a "lure to cognition").[3] That is, hearing the same word across many different situations may lead children to identify properties that are shared by those situations, and that distinguish them from situations to which the word is not applied. This construction process presumably draws on children's perceptual ability to recognize many different kinds of similarities and differences across spatial situations, and on their ability to conceptualize what the situations picked out by a word have in common (Mandler, 1992).[4]

Much research will be needed to decide among competing accounts and to work out the details of the learning process. One important step will be to examine the contours of children's prelinguistic spatial categories in more detail than has been done so far, as we have begun to do in preferential looking studies designed to assess prelinguistic knowledge of tight- and loose-fitting relations (McDonough, Choi, Mandler, & Bowerman, in press). Linguistically informed work on this problem will be crucial for determining how language-specific input in-

---

[3] It is a separate issue whether language-specific meanings for spatial words, once constructed, influence nonlinguistic perception, memory, or inferencing about space. Choi and Bowerman take no stance on this Whorfian issue; their proposals are directed to the question of how children learn to talk about space in a language-specifically appropriate manner. However, see Levinson (1996a, b) for new evidence on this perennially intriguing debate.

[4] Whether some properties of spatial situations are especially privileged in the context of learning spatial language remains a matter of controversy; compare, for example, Landau & Jackendoff (1993) and Slobin (in press).

teracts with children's nonlinguistic understanding of space to provide knowledge of the spatial semantic categories of a particular language.

REFERENCES

Baillargeon, R. (1995). A model of physical reasoning in infancy. In C. Rovee-Collier & L. P. Lipsitt (Eds.), *Advances in Infancy Research* (Vol. 9, pp. 305–371). Stamford, CT: Ablex.

Bavin, E. (1990). Locative terms and Warlpiri acquisition. *Journal of Child Language, 17,* 43–66.

Berman, R., & Slobin, D. I. (1994). *Relating events in narrative: A crosslinguistic developmental study.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Bierwisch, M. (1967). Some semantics universals of German adjectives. *Foundations of Language, 3,* 1–36.

Bloom, L. (1973). *One word at a time.* The Hague: Mouton.

Bloom, L., Lightbown. P., & Hood, L. (1975). Structure and variation in child language. *Monographs of the Society for Research in Child Development, 40* (Serial No. 160).

Bowerman, M. (1989). Learning a semantic system: What role do cognitive predispositions play? In M. L. Rice & R. L. Schiefelbusch (Eds.), *The teachability of language* (pp. 133–169). Baltimore: Paul H. Brooks.

Bowerman, M. (1996a). Learning how to structure space for language: A crosslinguistic perspective. In P. Bloom, M. Peterson, L. Nadel, & M. Garrett (Eds.), *Language and space* (pp. 385–486). Cambridge, MA: MIT Press.

Bowerman, M. (1996b). 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, UK: Cambridge University Press.

Bowerman, M., & Choi, S. (1994). *Linguistic and nonlinguistic determinants of spatial semantic development: A crosslinguistic study of English, Korean, and Dutch.* Paper presented at Boston University Conference on Language Development, Boston, MA.

Bowerman, M., & Choi, S. (in press). Shaping meanings for languages: Universal and language specific in the acquisition of spatial semantic categories. In M. Bowerman & S. C. Levinson (Eds.), *Language acquisition and conceptual development.* Cambridge: Cambridge University Press.

Bowerman, M., de León, L., & Choi, S. (1995). Verbs, particles, and spatial semantics: Learning to talk about spatial actions in typologically different languages. In E. V. Clark (Ed.), *The Proceedings of the 27th Annual Child Language Research Forum* (pp. 101–110). Stanford, CA: Center for the Study of Language and Information.

Bowerman, M., & Pederson, E. (1992). *Cross-linguistic perspectives on topological spatial relationships.* Paper presented at the annual meeting of the American Anthropological Association, San Francisco, CA.

Brown, P. (1994). The INs and ONs of Tzeltal locative expressions: The semantics of static description of location. *Linguistics, 32,* 743–790.

Brown, R. (1958). *Words and things.* New York: Free Press.

Brown, R. (1973). *A first language: The early stages.* Cambridge, MA: Harvard University Press.

Choi, S. (1995). The development of epistemic sentence-ending modal forms and functions in Korean children. In J. Bybee & S. Fleischman (Eds.), *Modality in grammar and discourse* (pp. 165–204). Amsterdam: John Benjamins.

Choi, S. (1997). Language-specific input and early semantic development: Evidence from children learning Korean. In D. I. Slobin (Ed.), *The crosslinguistic study of language acquisition, Vol. 5: Expanding the context* (pp. 41–133). Mahwah, NJ: Lawrence Erlbaum.

Choi, S. (1998). *Caregiver input in English and Korean: Use of nouns and verbs in book-reading and toy-play contexts.* Paper presented at XVth ISSBD Biennial Meeting, Bern, Switzerland.

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., & Gopnik, A. (1995). Early acquisition of verbs in Korean: A cross-linguistic study. *Journal of Child Language, 22,* 497–530.

Clark, E. (1973). Non-linguistic strategies and the acquisition of word meanings. *Cognition, 2,* 161–182.

Clark, H. H. (1973). Space, time, semantics, and the child. In T. E. Moore (Ed.), *Cognitive development and the acquisition of language* (pp. 27–64). New York: Academic Press.

de León, L. (1997). Vertical path in Tzotzil (Mayan) early acquisition: Linguistic vs. cognitive determinants. In E. Clark (Ed.), *The Proceedings of the 28th Annual Child Language Research Forum* (pp. 183–197)*.* Stanford University: Center for the Study of Language and Information.

Gentner, D. (1982). Why nouns are acquired before verbs: Linguistic relativity versus natural partitioning. In S. A. Kuczaj II (Ed.), *Language development: Vol. 2. Language, thought, and culture* (pp. 301–334). Hillsdale, NJ: Lawrence Erlbaum.

Gibson, E. J., & Spelke, E. S. (1983). The development of perception. In J. H. Flavell & E. M. Markman (Eds.), *Handbook of child psychology: Vol. 3. Cognitive development* (4th ed., pp. 1–76). New York: Wiley & Sons.

Gleitman, L. (1990). The structural sources of verb meaning. *Language Acquisition, 1,* 3–55.

Goldfield, B. (1993). Noun bias in maternal speech to one-year-olds. *Journal of Child Language, 20,* 85–100.

Golinkoff, R., Hirsh-Pasek, K., Cauley, K., & Gordon, L. (1987). The eyes have it: Lexical and syntactic comprehension in a new paradigm. *Journal of Child Language, 14,* 23–45.

Gopnik, A. (1980). *The development of non-nominal expressions in 12–24 month old children.* Unpublished doctoral dissertation, University of Oxford, Oxford, England.

Gopnik, A. (1982). Words and plans: early language and the development of intelligent action. *Journal of Child Language, 9,* 303–318.

Herskovits, A. (1986). *Language and spatial cognition: An interdisciplinary study of the prepositions in English.* Cambridge, UK: Cambridge University Press.

Hoff-Ginsberg, E. (1985). Some contributions of mothers' speech to their children's syntactic growth. *Journal of Child Language, 12,* 367–385.

Hoff-Ginsberg, E., & Shatz, M. (1982). Linguistic input and the child's acquisition of language. *Psychological Bulletin, 92,* 3–26.

Huttenlocher, J., Haight, W., Bryk, A., Seltzer, M., & Lyons, T. (1991). Early vocabulary

growth: Relation to language input and gender. *Developmental Psychology, 27,* 236– 248.

Huttenlocher, J., Smiley, P., & Charney, R. (1983). Emergence of action categories in the child: Evidence from verb meanings. *Psychological Review, 90,* 72–93.

Johnston, J., & Slobin, D. I. (1979). The development of locative expressions in English, Italian, Serbo-Croatian, and Turkish. *Journal of Child Language, 16,* 531–547.

Keppel, G. (1982). *Design and analysis: A researcher's handbook* (2nd ed.). Englewood Cliff, NJ: Prentice Hall.

Landau, B., & Jackendoff, R. (1993). "What" and "where" in spatial language and spatial cognition. *Behavioral and Brain Sciences, 16,* 217–265.

Levine, S. C., & Carey, S. (1982). Up front: The acquisition of a concept and a word. *Journal of Child Language, 9,* 645–657.

Levinson, S. C. (1996a). Frames of reference and Molyneux's question. In P. Bloom, M. Peterson, L. Nadel, & M. Garrett (Eds.), *Language and space* (pp. 109–169). Cambridge, MA: MIT Press.

Levinson, S. C. (1996b). Language and space. *Annual Review of Anthropology, 25,* 353–382.

Mandler, J. M. (1992). How to build a baby: II. Conceptual primitives. *Psychological Review, 99,* 587–604.

Mandler, J. M. (1996). Preverbal representation and language. In P. Bloom, M. Peterson, L. Nadel, & M. Garrett (Eds.), *Language and space* (pp. 365–384) Cambridge, MA: MIT Press.

McDonough, L., Choi, S., Mandler, J. M., & Bowerman, M. (in press). Comprehension of spatial terms in English and Korean: The use of preferential looking as a measure of semantic development. In E. L. Bavin & D. Burnham (Eds.), *Advances in Infancy Research* (Vol. 12, pp. 336–354). Norwood, NJ: Ablex.

Naigles, L. (1990). Children use syntax to learn verb meanings. *Journal of Child Language, 17,* 357–374.

Needham, A., & Baillargeon, R. (1993). Intuitions about support in 4.5-month-old infants. *Cognition, 47,* 121–148.

Piaget, J., & Inhelder, B. (1956). *The child's conception of space.* (F. J. Langdon & J. L. Lunzer, Trans.). London: Routledge & Kegan Paul (original work published 1948).

Quine, W. V. (1960). *Word and object.* Cambridge, MA: MIT Press.

Regier, T. (1995). A model of the human capacity for categorizing spatial relations. *Cognitive Linguistics, 6,* 63–88.

Sinha, C., Thorseng, L. A., Hayashi, M., and Plunkett, K. (1994). Comparative spatial semantics and language acquisition: Evidence from Danish, English, and Japanese. *Journal of Semantics, 11,* 253–287.

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.

Slobin, D. I. (in press). Form/function relations: How do children find out what they are? In M. Bowerman & S. C. Levinson (Eds.), *Language acquisition and conceptual development.* Cambridge, UK: Cambridge University Press.

Spelke, E. (1976). Infants' intermodal perception of events. *Cognitive Psychology, 8,* 553–560.

Talmy, L. (1985). Lexicalization patterns: Semantic structure in lexical forms. In T. Shopen (Ed.), *Language typology and syntactic description: Vol. 3. Grammatical categories and the lexicon* (pp. 57–149). Cambridge, UK: Cambridge University Press.

Tardif, T. (1996). Nouns are not always learned before verbs: Evidence from Mandarin speakers' early vocabularies. *Developmental Psychology, 32,* 492–504.

Tomasello, M. (1987). Learning to use prepositions: A case study. *Journal of Child Language, 14,* 79–98.