Evidence for a Broad Notion of Source in Child Language
Laura Lakusta, Malathi Thothathiri, Deanna Mendez, and Marija Marinkovic

1. Introduction

In language, source paths encode the starting point of a figure’s motion (The bird flew from the lake) and goal paths the endpoint (the bird flew to the lake). Linguistically, these paths show up in different conceptual domains (transfer, motion, change of state), suggesting a broad notion of source and goal in semantic structure (Jackendoff, 1983), and supporting theories positing broad thematic roles in language (e.g., Grimshaw, 1981; Jackendoff, 1990; Pinker, 1989). Whether such thematic role representations are available early in language development has been a question of much debate, with some arguing that children represent only verb-specific templates rather than abstract linguistic categories (e.g., Tomasello, 1992). Such theories argue, for example, that in a ‘giving’ event (the girl gave a present to the boy), the child linguistically represents the ‘girl’ as a ‘giver’ rather than as a ‘source’ or ‘agent’. The aim of the current study is shed light on this debate by testing whether children, early in language development, represent sources broadly and abstractly across two different event domains, thus lending support to theories arguing for broad thematic roles in language development.

The observation that the same linguistic structures show up in different conceptual domains cross-linguistically, was first described by Gruber (1965) and later formalized in Jackendoff’s Thematic Relations Hypothesis (Jackendoff, 1983). For example, the motion event of a ‘dog running out of his doghouse and into the pool’ and the change of state event of ‘a girl turning from happy to sad’ both include a source path (‘out of the doghouse’, ‘from happy’) and a goal path (‘into the pool’, ‘to sad’). Despite the deep conceptual difference that the former event involves motion being traversed over space and the latter event involves a state change over time, language treats the paths in these events similarly. In semantic structure all these events are encoded in terms of a source path ([FROM X]) and goal path ([TO X]). In English syntactic structure, source and goal paths are often encoded in prepositional phrases and marked with the prepositions, from, off, and out, for source paths, and to, onto, and into, for goal...
paths. These generalizations hold even for events in which the entities involved may take on more than one role. For example, consider the event, ‘the dolphin was splashing the girl’/’the girl was being splashed by the dolphin’. In this event, ‘the dolphin’ is the spatial starting point – the source of the action, as well as the agent. Such an analysis is consistent with Jackendoff’s theory positing that semantic roles should be analyzed as falling into two tiers: a thematic tier that encodes objects in terms of their motion and spatial locations, and an action tier that encodes those same objects in terms of their intentions (Jackendoff, 1990).

Given the broad and abstract notion of semantic roles in linguistic theory, the question arises whether children’s linguistic representations are also broad and abstract. Studies in language acquisition shed some light on this question. For example, recent studies suggest that there is a language bias to encode goal paths over source paths in prepositional phrases (Lakusta & Landau, 2005; 2012) and, notably, this goal bias extends to events across different domains. Lakusta and Landau (2005) reported that when 3-4 year old children watched simple motion events (e.g., a plane flies out of a bowl and into a pot), and were then asked to describe them, they mentioned goal paths more often than source paths (e.g., “the plane flew into the pot”). This pattern was also observed for motion events involving inanimate/non-intentional motion, such as a tissue blowing off a coaster into a candle (Lakusta & Landau, 2012), attachment/detachment (e.g., hook/unhook), change of possession events (e.g., give/get), and change of state events (e.g., change, turn) (Lakusta & Landau, 2005).

Further, studies by Clark and Carpenter (1989a; 1989b) suggest that children go through a period in development where they mark different types of starting points all with the locative marker ‘from’, and sometimes this marking is used non-conventionally. For example, in an examination of children’s spontaneous speech, a two-year-old child uttered, “These fall down from me”, to describe an event where he pushed pieces of sandwich off of his plate (Clark & Carpenter, 1989a, p. 350). In this example, although the child is the agent, he is marking himself with the locative source marker “from”. Clark and Carpenter (1989a) explored this further in an empirical study where children (2;5 – 6;1 years of age) were presented with oblique agent phrases in passive sentences (e.g., “Dan got chased by a big snake”) and were asked to imitate, and if needed, repair, sentences. The findings revealed that in their imitations, 2-year-olds produced the locative source marker “from” rather than “by” (e.g., rather than repeating, “Dan got chased by a big snake” a child may say “Dan got chased from a big snake”) (Clark & Carpenter, 1989a; 1989b). Clark (2001) explains that these errors are evidence for an ‘emergent category’ of source; “certain phenomena in early language acquisition suggest that some conceptual categories may surface in children’s speech even when they are not supported by the ambient language. These phenomena offer evidence for a set of general conceptual categories underlying language” (page 379).
The current study seeks further evidence for such broad source representations early in language development. We do so by testing whether exposing children to source paths in one semantic domain (locative motion) can prime source paths in another semantic domain (agentive action). If children represent sources broadly, as is suggested by several linguistic theories, then such priming may be possible.

Structural priming refers to the facilitation of a linguistic structure after prior exposure to the same structure. This phenomenon can be used to investigate the nature of linguistic representations in adults and children (Chang, Bock, & Goldberg, 2003; Huttenlocher, Vasilyeva, & Shimpi, 2004; Thothathiri & Snedeker, 2008). The logic of this paradigm is as follows: If a linguistic structure primes another structure in the absence of lexical overlap, then the two structures must involve common structural representations. The kind of structural overlap between prime and target sentences can be manipulated in different ways to precisely identify the nature of the linguistic representations that led to priming. In the present case, we used this logic to investigate whether a prime sentence like “The dog ran from the doghouse” (hereafter called the source prime) facilitated processing of a target sentence like “The dolphin was splashed by the child” more so than another prime sentence like “The dog ran to the doghouse” (hereafter called the goal prime). The source and goal primes share the same syntactic structure (Noun-phrase Verb Prepositional-phrase). Thus, the facilitation of syntactic structure alone could not lead to differential effects after one kind of prime versus another. The critical difference between the two kinds of primes lies is in the thematic role indicated by the prepositional phrase (“from the doghouse” = source, “to the doghouse” = goal). Thus, greater facilitation of a passive target by the source prime than a goal prime would indicate priming at the level of thematic roles—specifically that mapping the prepositional phrase in the source prime to the “source” thematic role facilitated mapping the prepositional phrase in the passive sentence to the “agent” thematic role. This, in turn, would allow us to infer that children have overlapping representations of the source and agent thematic roles.

2. Method
2.1. Participants

Seventy-six children between the ages of 2;5 and 4;5 ($M = 3;5$) participated in the current study. Participants were recruited from the surrounding area in northern New Jersey through mailings and flyers posted at parents’ groups and daycares.

2.2. Materials and Design

Thirty-two pairs of colored illustrations were constructed to portray motion events. Specifically, one picture in each pair portrayed a figure moving from ($n = 4$), out of ($n = 6$), or off of ($n = 6$) a starting point object (source path) and the
other picture in each pair portrayed the figure moving to \((n = 4)\), into \((n = 6)\), or onto \((n = 6)\) an endpoint object (goal path). Thus, there were 16 pictures portraying source paths and 16 portraying the corresponding goal paths (see Figure 1a). These pictures (and the corresponding phrases that accompanied them – see below) are referred to as the source and goal primes, respectively. Half the primes \((n = 16)\) had animate figures and half had inanimate figures. All primes had inanimate ground objects. Further, the 16 source primes were randomly made into eight pairs and the 16 goal primes were randomly made into eight pairs. For example, one pair of source primes was “the dog ran from the doghouse” and “the car drove from the gas station”. One pair of goal primes was “the baby crawled onto the blanket” and “the car drove to the gas station”. As described below, each pair of primes preceded a target.

In addition, 16 pairs of colored illustrations were constructed to portray agent/patient action events. Specifically, each picture pair portrayed two animate objects, where one animate object (henceforth, object A) affected another animate object (henceforth, object B). One picture in each pair displayed object A as the agent and object B as the patient and the other picture in the pair displayed the reverse (see Figure 1b). These pairs of pictures are referred to as the targets. Further, for half the targets, the child was presented with a partial passive phrase when viewing the target (e.g., “The dolphin was splashed by the child”) and for the other half the child was presented with a partial active phrase when viewing the target (e.g., “The child was splashing the dolphin”; see Figure 1b). The pairing of phrase type (passive and active) with targets was counterbalanced across participants.

The primes and targets were then randomly paired to construct 16 trials. Each trial consisted of two different primes of the same type (two source primes or two goal primes) and a target (i.e., the pair of pictures). Thus, the final set of test stimuli consisted of 16 test trials: eight targets described by partial passive phrases, four of which were preceded by source primes (henceforth, SourcePassive) and four of which were preceded with goal primes (henceforth, GoalPassive) and eight targets described by partial active phrases, four of which were preceded by source primes (henceforth, SourceActive) and four of which were preceded with goal primes (henceforth, GoalActive).

2.3. Procedure

Children viewed each prime picture while listening to an experimenter read a sentence describing the picture (e.g., “the dog ran from the doghouse”). Children were asked to repeat each sentence immediately after the experimenter. Following the two primes of the same type (source or goal), children then viewed the two target pictures simultaneously while the experimenter read aloud a carrier question and the target phrase in passive or active voice (carrier: Can you point to the picture where…; passive continuation: The dolphin was splashed by the child); active continuation: The dolphin was splashing the child; Fig. 1). The dependent measure was how accurate children were in pointing to
the correct picture. Children’s language and pointing actions were recorded for later transcription/analysis. All of the children’s responses were reviewed by a second coder and any discrepancies were resolved by a third coder.

Prior to the test phase described above, children received a practice phase to acquaint them with the structure of the task. This practice phase consisted of three practice trials. If the participant pointed to an incorrect picture during practice, the experimenter explained to the child why it was incorrect and redid the trial. Children who scored two out of the three practice trials correctly were able to proceed to the test phase.

**Example primes**

Locative source prime
“The dog ran from the doghouse.”

Locative goal prime
“The dog ran to the doghouse.”

*Figure 1a.* Example of source and goal primes. Note that children received both of these primes during the experiment, although they received them on different trials. As described in the text, for eight trials children received two different source primes followed by a target and for eight trials they received two different goal primes followed by a target.
Example target

Dolphin splashing child.      Child splashing dolphin.

Figure 1b. Example target. After being presented with two primes of the same type (i.e., two different source primes or two different goal primes), children viewed both target pictures simultaneously, and the experimenter said, “Point to the picture where….”. Children then heard one target phrase (either the active continuation, “The dolphin was splashing the child” or the passive continuation “The dolphin was splashed by the child”).

2.4. Hypotheses

Our hypotheses are as follows: 1) if children have overlapping thematic roles of locative source paths and agents, then when presented with a target paired with a passive phrase (e.g., “the dolphin was splashed by the child”), children should point more frequently to the correct picture when primed with a source path vs. a goal path and 2) as observed in previous research (e.g., Maratsos et al., 1985), children should point more frequently to the correct picture when the target is described by an active phrase vs. a passive phrase, regardless of prime type.

3. Results

Trials where the participant did not successfully repeat at least one prime phrase were excluded from the analysis (N = 3). Further, children who did not repeat at least one prime phrase for at least 9 trials were excluded from the analyses (N = 6).

Children’s data were split into two groups by performing a median split based on age (Median age = 3.48 years) and data for each age were analyzed separately to test the two hypotheses outlined above.

In order to test the hypothesis that children would perform better on the passive targets when primed with locative source phrases vs. locative goal phrases, one-tailed, paired t-tests examined the proportion of passive targets correct when primed with sources (SourcePassive) vs. when primed with goals
(GoalPassive) for each age group. As shown in Figure 2, for the younger age group, children scored marginally better on passive targets when primed with source paths vs. goal paths, $t(37) = 1.67, p = .05$. However, this was not the case for older children, with the proportion of passive targets correct not significantly differing between the two types of primes (source vs. a goal prime), $t(37) = .04, p = .97$ (Figure 2). Non-parametric tests further support these results, as shown in Table 1.

![Figure 2](image_url)

**Figure 2. Average proportions (and SEs) of passive targets correct for the younger and older age groups by prime type (source or goal)**

In order to test the hypothesis that children would perform better on the active targets vs. passive targets, regardless of prime type, one-tailed, paired $t$-tests examined the proportion of passive targets correct (SourcePassive + GoalPassive) vs. active targets correct (SourceActive + GoalActive) for each age group. As shown in Figure 3, for both age groups, children scored better on active targets vs. passive targets, $t(37) = 4.96, p < .001$, $t(37) = 2.89, p = .006$, for the younger and older age groups, respectively. Non-parametric tests further supported these results, as shown in Table 1.
Figure 3. Average proportions (and SEs) of passive vs. active targets correct for the younger and older age groups collapsed over prime type.

Table 1. Individual participants’ performance

<table>
<thead>
<tr>
<th>Age Group</th>
<th># of children performing better on passive targets when primed with source paths vs. goal paths</th>
</tr>
</thead>
<tbody>
<tr>
<td>Children &lt; 3.48 yrs.</td>
<td>19/38                                                                                     [ z = 1.97, p &lt; .05 ]</td>
</tr>
<tr>
<td>Children &gt; 3.48 yrs.</td>
<td>13/38                                                                                     [ z = .16, p &gt; .10 ]</td>
</tr>
</tbody>
</table>

Table 1b

<table>
<thead>
<tr>
<th>Age Group</th>
<th># of children performing better on active targets vs. passive targets (collapsed over prime type)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Children &lt; 3.48 yrs.</td>
<td>25/38                                                                                     [ z = 3.96, p &lt; .05 ]</td>
</tr>
<tr>
<td>Children &gt; 3.48 yrs.</td>
<td>21/38                                                                                     [ z = 2.95, p &lt; .05 ]</td>
</tr>
</tbody>
</table>

Note: The statistical tests performed were Wilcoxin-signed ranks tests.
4. Discussion

Using a structural priming task, the current study tested the hypothesis that children represent sources broadly in semantic structure. Our findings revealed that children in the younger age group performed better on the passive targets when primed with source phrases than goal phrases. Further, replicating past findings, both groups of children performed better on the active targets than the passive targets.

These findings suggest that a broad notion of source may indeed characterize the semantic structures of the younger children. Specifically, young children’s representations of locative source paths (the dog ran from the doghouse) and agents that are the starting points of actions (the dolphin was splashed by the child) may overlap. This lends support to theories positing broad thematic roles in language and language development (e.g., Grimshaw, 1981; Jackendoff, 1990; Pinker, 1989), as well as specific linguistic analyses positing that participants in events are represented both in terms of a thematic and action tier (Jackendoff, 1990). The current findings also extend previous findings suggesting a broad notion of source (Clark & Carpenter, 1989a) in early language development.

The lack of evidence for a broad notion of source in older children is compatible with at least two possibilities: 1) a broad notion of source may get differentiated into finer-grained concepts over development or 2) a broad notion of source persists at older ages but plays a weaker role during language processing. Considering the first possibility, it’s possible that over development, a broad notion of source may get differentiated into concepts such as ‘locative starting point’, ‘agentive starting point’, starting point in change of state event, etc. Such as change may be related to children’s developing language; for example, whereas young English-speaking children tend to use the same word, “from”, to mark starting points across different domains (Clark & Carpenter, 1989a), older children learn more conventional markings (in English, “by” rather than “from” to mark an agent). Thus, distinctions that emerge in language may bootstrap distinctions at the conceptual level, as has been observed in other cases of language learning and conceptual development (Shusterman et al., 2016). Another possibility is that a broad notion of source persists over development, but plays a weaker role in language processing; thus, the locative source prime was ineffective in priming the passive structure for older children not because there was no representational overlap between the two types of sources (locative and agent), but rather because the overlap was too weak for any effective priming to take place. If this was the case, methods should be able to be developed such that a broad notion of source is indeed revealed in older children, and perhaps even in adults.

In conclusion, the current findings suggest that the broad thematic role of ‘source’ does indeed characterize the semantic representations of young children, and sets the stage for future studies to explore whether and how these representations may change over development.
References


