Supplemental Materials for

On the plasticity of semantic generalizations:

Children and adults modify their verb lexicalization biases in response to changing input

Carissa L. Shafto[1], Catherine Havasi[2], & Jesse Snedeker[1]

[1]Harvard University

[2]Massachusetts Institute of Technology

Supplemental Materials for

"On the plasticity of semantic generalizations"

Languages differ in how they package the components of an event into words to form sentences. For example, while some languages typically encode the manner of motion in the verb (e.g., *running*), others more often use verbs that encode the path (e.g., *ascending*). Prior research has demonstrated that children and adults have biases to assume that novel motion verbs will reflect the dominant pattern of their own language (*lexicalization biases*). The experiments in our paper explore the plasticity of these lexicalization biases in children and adults (Author, Author, & Author, under review). In addition to exploring the plasticity of lexicalization biases, this work also examined the degree to which these biases vary across syntactic contexts. In our experiments with adults, we manipulated the syntactic context in which the verb appeared, using either the prepositional frame ("He is torging around the chair") or the direct-object frame ("He is torging the chair"). Prior research has documented that English speakers have a strong bias to interpret verbs in the prepositional frame as encoding manner of motion, but have no strong preference for the direct-object frame (Naigles & Terrazas, 1998). We replicate this pattern (Experiments 1 & 2). We also find that, while biases in both frames can be shaped by new learning, response patterns for prepositional frames continue to show a manner bias after training while those in the direct-object condition come to very closely match the experimental input.

We suggest that this pattern reflects how these frames are used in English. Specifically, based on the prior studies and our intuitions, we argue that: 1) In English, prepositional frames are commonly used to describe motion events and thus participants will expect that a novel verb used with an appropriate prepositional phrase is likely to encode motion with respect to a goal. 2) In English, prepositional frames can be used with both manner and path verbs ("He is

skipping around the chair" and "He is circling around the chair"), but are overwhelmingly biased toward manner verbs. These two facts in concert explain both the initial bias and participants' unwillingness to completely give it up when faced with probabilistic input.   3) In contrast, the direct-object frame is only infrequently used to describe self-generated motion, but is frequently used to encode other types of events and states ("He is examining/buying/touching the chair"). Thus participants who hear a novel verb in that frame may initially interpret it as referring to something other than motion. 4) When motion verbs are used with a direct-object frame, they are equally likely to encode manner and path ("He entered the room" and "He ran the race"). As a result, there should be no strong lexicalization bias in this frame and, since the data for developing a bias in the input is sparse, participants should be willing to adopt whatever bias is available in the training materials.

The prior research on the use of manner and path verbs has not focused on these specific questions. However, it did directly inform the assumptions that we made and it provides strong circumstantial evidence that these assumptions are likely to be true.

Prior work on the distribution of manner and path verbs across languages has employed three methods: 1) The traditional linguistic method of systematically generating examples of the relevant kinds and obtaining grammaticality judgments from native speakers (Aske, 1989; Jackendoff, 1990; Talmy, 1985). While this method is critical for generating hypotheses and identifying acceptable and unacceptable utterances, it cannot tell us about the relative frequency of different forms. It also may fail to describe the true diversity of what language users actually do (as opposed to what they ought to do). 2) Narration tasks in which naïve speakers of different languages are asked to narrate or describe a picture book or video that features a variety of motion events (Allen, Ozyurek, Kita, Brown, Furman, Ishizuka & Fujii, 2007; Berman & Slobin,

1994; Naigles, Eisenberg, Kako, Highter & McGraw, 1998). This method does provide

quantitative data about how language is used by actual speakers and allows experimenters to

manipulate semantic features of the events. However, because the experimenter controls the

situation that is being described, the language that is produced is unlikely to be representative of

language use in the wild. 3) The analysis of children's spontaneous speech (Choi & Bowerman,

1991). These methods have the potential to capture a more representative sample of language

use, but to date they have only been employed with very small numbers of speakers.

In this supplement, we present data using three different methods (analyses of written

corpora, sentence completion studies with naïve adults, and acceptability judgments from naïve

adults). The corpus analyses supplement the existing work by providing information about the

distribution of forms in a broad sample of language. The completion and judgment studies allow

us to directly look at the interpretation and acceptability of frames and verbs that seem to rarely

appear in elicited production tasks. These methods also have obvious limitations. For example,

we cannot know precisely what events participants are writing about, and if we wish to work

with large amounts of data (e.g., millions of webpages or books), we cannot even code the

apparent meanings of the utterances. These methods also force us to work with adults and thus

provide no insight into how these patterns develop over time.

Prior studies have provided a rich picture of how motion verbs are used in a variety of

languages, which informed our hypotheses and research design. Specifically, this work

demonstrates that English-speaking adults overwhelmingly prefer to describe motion events

using verbs that encode manner of motion rather than path of motion, this pattern reverses for

adults speaking path languages such as Greek, Korean, Japanese & Turkish, and both of these

language-specific patterns are present in the speech of children as young as three (see e.g., Allen

et al., 2007; Naigles et al., 1998; Papafragou & Selmis, 2010; Berman & Slobin, 1994).

Researchers working on these questions have been acutely aware of how the syntactic resources

of a language can influence these lexicalization patterns, and they often suggest that the use of

manner verbs in manner languages is related to the widespread use of the prepositional frame to

encode directed motion across a wide range of event types (see e.g., Aske, 1989; Slobin &

Hoiting, 1994). The elicited production studies have documented that adult and child speakers of

manner languages often encode both manner and path in a single clause containing just one verb

(e.g., Berman & Slobin, 1994; Allen et al., 1997) and the examples that are given suggest that

most of these utterances involve a prepositional phrase or prepositional particle (e.g., *pop out*).

Taken together these facts *suggest* that the prepositional frame is commonly used to

encode motion events and is typically used with a manner of motion verb, while the direct-object

frame may not be used to encode motion very often. However, we can find no analyses that

actually compare the use of prepositional and direct-object frames in English speakers. Studies 1

and 2 directly explore the syntactic contexts in which manner and path verbs are used in two

large corpora of written English. Study 3 examines how adult English speakers use known

words in sentence frames like those that are used in our experiments, in order to determine what

speakers have learned about the distributional properties of their language. The results of these

three studies support our hypotheses about the statistical properties of the language.

Study 4 asks adults to judge the acceptability of manner and path verbs in direct-object

and prepositional frames. We find that the prepositional frame is seen as compatible with both

manner and path verbs, while the direct-object frame is often seen as unacceptable with a manner

verb. These judgments, in concert with the distributional analyses, justify our decision to

conduct our developmental studies of the acquisition of lexicalization biases (Author et al., under review) using the prepositional frame.

Study 1: Motion Verb Use in Primarily Informal Text

To explore the syntactic contexts in which motion verbs are used we conducted a google page search for a selected set of motion verbs, comparing their use in prepositional frames ("She passed by the playground") with their use in direct-object frames ("She passed the playground"). Most of the material indexed in google is not edited. In many cases the writer may view it as temporary and casual (facebook posts, twitter feeds, blogs). Thus we consider this to be a sample of informal, unreflective language use.

Analysis 1

Method. Fifteen verbs were selected for investigation: six path verbs, six manner verbs and three verbs whose status is disputed.

The path verbs were: *ascend*, *descend*, *pass*, *circle*, *enter* and *exit*. They were selected based on three criteria: 1) They are clearly path verbs. Each of these verbs lexicalizes a distinct path with respect to a reference object. None of these verbs contain information about the manner of motion (unlike *climb*) and none has a deictic element (unlike *leave*). 2) They can be used with a variety of reference objects (unlike *traverse*). 3) They are fairly frequent in English. There is only one other verb that meets these criteria (*cross*), which we excluded accidentally.

The manner verbs were: *walk*, *run*, *skate*, *jump*, *slide* and *swim*. There are many manner verbs in English. These verbs were selected from the Macarthur-Bates CDI-2 to ensure that we focused on common verbs that are learned early. We excluded manner verbs that encode an instrument (e.g., *drive*) because this instrument will often surface as a direct object, complicating our analyses.

In our initial analysis, we also looked at three verbs whose status is less clear. The first, *climb*, sometimes appears to encode the manner of motion ("He carefully climbed down from the telephone pole") and sometimes appear to encode the path ("He climbed up the mountain, running as he went"). The other two verbs, *come* and *go*, do not encode either manner or path but instead encode the motion and deixis. However, because they are among the most frequent verbs in the language, their pattern of use may have an effect on the degree to which particular syntactic frames are associated with motion verbs more generally.

In order to compare the prepositional frame and the direct-object frame we had to generate search strings that would pick out sentences of the relevant kind. To do this, we generated 2 to 5 prepositions (or sets of prepositions) for each of the path verbs ($M = 3.66$). These prepositions were selected because the paths that they encoded were consistent with the path lexicalized in the verb. For example, for ascend we used *over*, *to*, *up* and *up to*. We selected a matched number of prepositions for the other classes of verbs, choosing the prepositions that seemed most likely to be used with that verb. This method of sampling prepositions served to reduce the search space. It may lead us to underestimate the number of prepositional uses of a verb, but critically it cannot inflate the number of prepositional uses or affect the number of direct-object uses.

Each search string consisted of the verb in its past tense form, the preposition (for the prepositional frames), and a definite article. For example, the search strings for circle are shown in (1) - (3).

1. "circled around the"

2. "circled over the"

3. "circled the"

Each string was entered as a google search (using quotes to ensure exact matches) and the number of pages (in millions) was entered into a spreadsheet.

Results. For each verb, we tallied the total number of uses in the prepositional frame (1 and 2) and the total number of uses in the direct-object frame (3). Then we calculated the proportion of direct object uses [3 / (1+2+3)]. The results appear in Figure 1.

Manner verbs were almost always used in the prepositional frame. Less than 4% of the uses were in the direct-object frame ($M = 3.8\%$ when averaged by verb type, 3.3% of all tokens). This was consistent across the six verbs (range 7% to 2%). Path verbs were more likely than manner verbs to appear in the direct-object frame ($U = 35$, $z = 2.34$, $p = .008$) but only a minority of uses were in this frame ($M = 20.1\%$ by verb type, 9.9% of tokens). The proportion of direct-object uses varied considerably across the six verbs (from 7% for *ascend* to 46% for *exit*). The three other motion verbs were rarely used in the direct-object frame (4% for climb, 0.3% for go, and 0.9% for come) patterning with the manner verbs, rather than the path verbs.

The high proportion of prepositional uses of path verbs may seem surprising, given that many usage guides caution against redundant prepositions. To understand how these frames were being used, we examined the first pages that were retrieved in these searches. Many of them were clearly uses of a path verb describing a physical motion event with a prepositional phrase that specified the path being traversed (e.g., "He descended down from the sky" and "She entered into the house"), but many others were metaphorical motion events ("I exited out of the site by mistake") or cases in which the prepositional phrase did not mark an agent ("the fiscal deal passed by the Senate"). Similarly, some of the direct-object frames that were used with manner verbs clearly encoded the path of an event ("We swam the Grand Canyon") but others were idiomatic ("jumped the shark") or involved the causative meaning of the verb ("If I ran the

zoo"). Our second analysis explored whether the same patterns are present in a small set of contexts in which the verbs are more likely to refer to a motion event and the post-verbal noun is likely to be the ground.
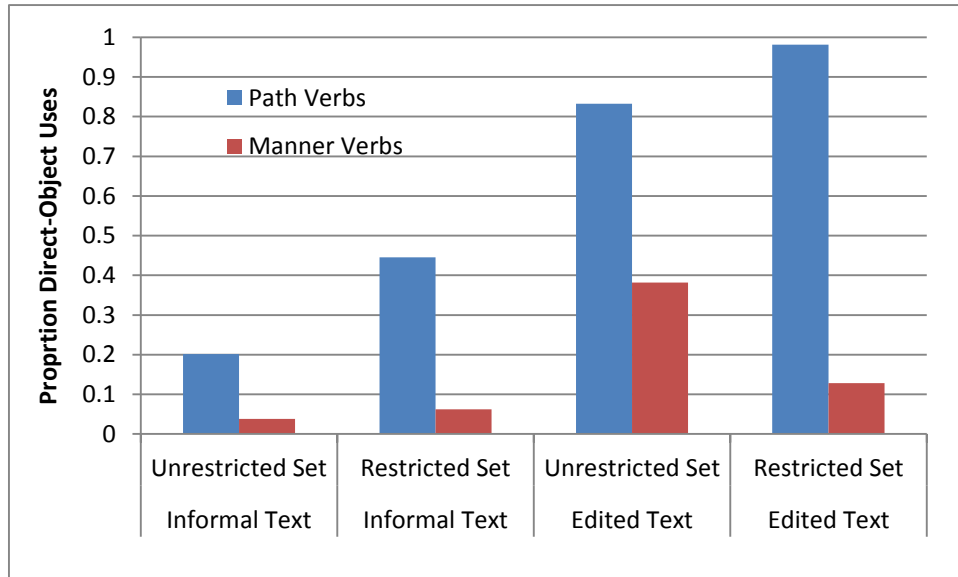
Figure 1: The proportion of direct-object uses for verbs of each type in two corpora. The analysis of informal text is based on google page counts and the analysis of edited text is based on the google Ngram Viewer corpus. The unrestricted set is designed to find most uses with a definite determiner, while the restricted set searches for specific noun phrases (NP's) that are likely to be the ground object in a motion event.

Analysis 2

   Method. For each of the 6 manner and path verbs used for analysis 1, we selected a direct object noun phrase that would make a good ground object for the motion encoded in that verb. We adjusted the prepositions used in these frames so that they were plausible given the reference objects but we kept the number of prepositions constant (e.g., "jump up to the" became "jumped off the fence"). These ground objects and prepositions are listed in Table 1. The google page search was run following the procedure above – we tallied the number of results for each search and calculated the proportion of direct object uses.

Table 1: Verbs, prepositions and noun phrases used in analysis 2

	Verb	Prepositions	Noun phrase
Manner verbs			
	jumped	over/from/across/off	the fence
	skated	across/over	the ice
	ran	to/from/across/into	the room
	swam	across/to/around	the pool
	walked	to/from/across/around/into	the store
	slid	down/under/across/to	the slide
Path verbs			
	ascended	to/up/up to/over	the ladder
	circled	around/over	the block
		down/from/down from/under/from	
	descended	under	the stairs
	entered	into/in/to	my house*
	exited	from/off/off of/out/out of	the store
	passed	by/into/over/across	the playground

* "the house" was used initially but produced a sample dominated by biblical quotes

Results. The manner verb stimuli were primarily used in the prepositional frame (See Figure 1; $M = 6\%$ direct-object frames by verb, 1% direct-object tokens). One of the manner verb searches ("jumped the fence") produced a large number of direct-object uses (22%) and in most cases the complement appeared to clearly encode the path of motion. The remainder of the manner verb searches had few direct-object uses and examination of the pages suggested that these were often cases with intervening punctuation or where the direct object did not encode a path ("out-skated the Ice Tigers" and "slid the slide too far back").

The path verb stimuli were used in the direct-object frame on a sizeable minority of the trials ($M = 44\%$ by verb type, 12% of tokens). The proportion of direct-object uses varied considerably across the six items. The direct object form was vastly preferred for some path stimuli ("ascended the ladder", 91%) while the prepositional form was clearly preferred for others ("passed by the playground", 0.7%). As a result of this variability the difference between

the path and manner searches was only marginal ($U = 29$, $z = 1.68$, $p = .09$) despite the apparent difference in the means.

Discussion

The two analyses produced very similar findings. In informal English, manner verbs are used in prepositional frames about 20 times as often as they are used with direct objects. The direct-object uses of manner verbs are semantically diverse. There are a few cases where the object indicates a path of motion ("swam the English channel", "jumped the fence") but in other cases the transitive form reflects a causative use of the verb and the direct object is the thing undergoing motion ("slid the glass down to the other end of the counter") with the path still appearing in a prepositional phrase.

In informal English, path verbs are used in both prepositional and direct-object frames. Prepositional uses appear to be somewhat more frequent than direct-object uses, but which form is preferred varies considerably depending on both the verb and the specific reference object that defines the path.   What is clear is that prepositional uses of path verbs are common in many contexts. The examples below illustrate this. For each of these verbs, the variant with the prepositional phrase is far more common than the variant without the preposition.

4. On the way to the market, I passed by the playground.

5. Vernon bellowed as he descended from the stairs.

6. It didn't hit me until I entered into my house.

In all of these uses, the prepositional phrase clearly encodes the path of motion. In many cases, the preposition redundantly marks the path encoded in the verb (4 & 6).

While this analysis has many limitations, these limitations are unlikely to account for the differences that we observed. First, in both analyses we did not search for all possible

prepositions. As a result we are, to some degree, overestimating the proportion of direct-object uses. This overestimation should be greater for the manner verbs (which are compatible with all paths) than it is for the path verbs (which are typically incompatible with all but a few prepositions). Thus this limitation cannot account for our critical results: 1) the direct-object frame is used less than the prepositional frame and 2) this is particularly true for manner verbs.

Second, our search string also included information about the NP complement, which could affect our results. In the first analysis, the only constraint on the NP complement was the determiner ("the"), which allows for a wide range of nouns. In this analysis, some of the sentences that were retrieved described metaphoric motion ("entered into the argument") or caused motion ("walked the dog"). To determine whether the same patterns held up in a sample with fewer of these uses, we conducted the second analysis, selecting specific NP complements that were likely to be used as reference objects for the verb in question but unlikely to appear in metaphoric motion descriptions or causatives. Because the noun phrases were selected arbitrarily, our sample will not necessarily be representative of all uses of the verb as a motion verb. Both analyses have their limitations, but these limitations are different ones. Nevertheless, the same basic patterns emerged in both analyses. Both manner verbs and path verbs are commonly used in prepositional frames. Manner verbs are only rarely used with direct object complements that encode the path. The frequency with which the direct-object form is used with a path verb varies considerably across verbs and reference objects.

Study 2: Motion Verb Use in Edited Text

Perhaps the most surprising finding of the first study is that path verbs are often used with prepositional complements. This use defies the stylistic conventions that many writers are taught in school. Thus we might expect that the use of prepositional complements for motion verbs

would be rare in published books, which, presumably, have been edited with these standards in mind. To explore this we conducted the same searches on the google Ngram Viewer, which calculates the relative frequency of strings of words based on a corpus of over 300 billion words from over 3 million books (Michel et al., 2011).

Analysis 1

A search was conducted using the path and manner search strings from Study 1, Analysis 1. We used the English corpus, generating graphs from 1800-2000, with a smoothing function of 3 years. We tabulated the estimated percentage of use for the year 2000. From this we calculated the proportion of direct-object strings (relative to all strings that were searched), just as we had in Study 1 (see Figure 1).

The path verbs were used in the direct-object frame considerably more often than they were used in the prepositional frame ($M = 83\%$ direct object uses by verb, 81% by tokens). In contrast, the manner verbs were used in the direct-object frame less than half of the time ($M = 38\%$ by verb, 40% direct-object tokens). In both cases, the pattern was consistent across verbs (manner verbs range 29% - 49%, path verbs range 71% - 92%) and the difference between the path and manner verbs was reliable ($U = 36$, $z = 2.8$, $p = .005$).

As we predicted, the N-gram analysis resulted in a higher proportion of direct-object uses for the path verbs than the google page count analysis ($W = 21$, $p = .05$), consistent with the hypothesis that prepositional frame is suppressed for path verbs in edited text. However, the N-gram analysis also resulted in a higher proportion of direct-object uses for the manner verbs ($W = 21$, $p = .05$).

This difference could reflect a difference in the topics that are discussed in books as opposed to webpages. Books might typically take on abstract topics resulting in more uses of

motion verbs in metaphoric or causative sentences, and these meanings may often give rise to direct-object frames. A quick perusal of some of the contexts that were provided for manner verbs in direct-object frames supports this conjecture (e.g., "If my Nevada mama ran the world" and "walk the talk"). To explore these differences more systematically, we conducted an analysis of search strings that targeted the use of each verb to describe non-causative concrete motion events. We expected that under these conditions, path verbs would be used almost exclusively in direct-object contexts (reflecting prescriptive pressure) while the manner verbs would primarily be used in prepositional contexts.

Analysis 2

This analysis used the search strings generated for Study 1, Analysis 2 and the procedure employed in Study 2, Analysis 1. Thus the search strings included a noun in the complement which was selected to be a good reference object for a path but a poor theme or complement for a metaphoric motion event.

We found that the path verb search strings were almost exclusively used in direct-object frames (see Figure 1, $M = 98\%$ by verb, 99% direct-object tokens). In contrast, the manner verb strings were typically used in prepositional frames ($M = 13\%$ direct-object uses by verb, 10% by tokens), resulting in a reliable difference between the two classes of verbs ($U = 36$, $z = 2.8$, $p = .005$). Most of the uses of manner verbs in direct-object frames came from the stem "jump the fence" where the direct-object form occurred 55% of the time.

We compared this analysis to the parallel analysis on google page counts (Study 1, Analysis 2). There was a higher proportion of direct-object uses for the path verbs in the N-grams corpus ($W = 21$, $p = .05$), suggesting that the prepositional frame may be suppressed for path verbs in edited text. However, there was no reliable difference between the corpora for the

manner verbs ($W = 13$, $p > .1$). Thus, when we focus on cases where manner verbs are used to describe non-causal, physical motion events, we find no evidence that their use is affected by the formality of the text or the degree to which it is edited. In both contexts, manner verbs are predominantly used with prepositional complements.

Study 3: Sentence Completion for Prepositional and Direct-Object Frames

In designing the experiments in the paper (Author et al., under review) we made four assumptions about the expectations that English speakers will have about verbs that appear in different syntactic frames:

1) Intransitive sentences with prepositional phrases will often be interpreted as encoding motion along a path, because this is a common function of the prepositional frame in English. When the preposition encodes a path (e.g., *to* or *around* rather than *with* or *of*) and the postverbal noun is a plausible ground object (e.g., *tree* or *stairs* rather than *idea* or *honey*), this will be the dominant interpretation of these sentences.

2) Transitive sentences without prepositional phrases will generally be interpreted as encoding other kinds of events and states, even when the direct object is a plausible ground object for defining a path. This will occur because the direct-object frame is used less often with motion verbs (see Study 1) and can be used to express a variety of other events and states (caused motion, caused change of state, contact, psychological states, etc).

3) When prepositional frames are interpreted as encoding motion along a path, there will be a strong expectation that the verb will encode manner of motion. This expectation reflects the linguistic experience of the participants: verbs encoding manner are more common in this frame than verbs encoding path and participants have substantial

experience with this frame in both formal and non-formal language use (see Studies 1

& 2).

4) When direct-object frames are taken to encode motion participants will not have

strong expectations about whether the verb will encode the manner or the path. This

lack of a bias would reflect participants' limited experience with directed motion

verbs in transitive sentences and the conflicting input that they get about the

meanings of verbs in these fames (compare Study 1 and Study 2).

Study 3 tests these hypotheses directly, by presenting participants with sentences similar to those

used our experiments and asking them to fill in the verb with an actual word. This method is

commonly used in psycholinguistics to tap participants' implicit knowledge of the distributional

patterns in their language.

Method

Participants. Forty-eight native English speakers were tested via Amazon Mechanical

Turk. Twenty-four participated in the prepositional frame condition, and 24 in the direct-object

frame condition. No further demographic information was requested or obtained.

Stimuli. Thirty-six stimulus sentences were created on the basis of the sentences that

accompanied the initial ambiguous scene and the five training scenes in Experiments 3 and 4 of

the paper (Author et al., under review). Specifically, we took the critical sentences from the path

version of the experiment, we put them in the present progressive tense (e.g., "She is bamping up

the steps") and we removed the verb and replaced it with a blank. We identified places where

the same reference object was used twice for a given verb and changed the noun to another noun

with a similar meaning (e.g., "hill" became "mountain"). The resulting sentences were the

stimuli for the prepositional frame condition (7) and the sentences for the direct-object condition

(8) were constructed by removing the preposition.

7. Prepositional Condition: She is _____ up the steps.

8. Direct Object Condition: She is _____ the steps.

Participants received either 36 direct object sentences or 36 prepositional sentences, interspersed

with 20 filler sentences. Six different lists were created each with different subject nouns (*she*,

he, *the teacher*, *the actor*, *the man*, and *the woman*). A single randomized order was used.

Procedure. Participants were instructed to type in one word to complete the sentence.

They were told that their responses must be grammatical but were given no feedback.

Results and Discussion

The participants' responses were coded for the following four features: 1) What part of

speech is the word (verb, preposition, adjective, adverb or noun)? 2) If the word is a verb, does

the sentence describe a motion event? An event was coded as a motion event if the verb entailed

a change of location of one or more participants. 3) If the verb encodes a motion event, does the

post-verbal noun phrase encode the ground object for the path of motion (coded as "directed

motion") or does it encode a theme in a caused motion event (coded as "caused motion"). 4) If

the verb encodes motion, does it specify the manner of motion, the path of motion, or

both/neither (e.g., *climb*, *go*, *come*). Based on these codes we divided the responses into six

categories (see Figure 2).

"Non-Verbs" included all responses that were not verbs. These responses were far more

frequent for the direct-object frames than they were for the prepositional frames ($M = 41.3\%$, M

$= 1.4\%$ respectively; $U = 552.5$, $z = 5.45$, $p < .001$). In the direct-object condition, all but one of

the non-verb responses was a preposition ("She is <u>at</u> the church"). In the prepositional frame

condition many were adverbs ("The man is <u>almost</u> into the office"). All of the non-verb

sentences described states.

"Non-Motion Verbs" included all responses that were verbs but did not encode motion.

These responses occurred at roughly the same rate in the direct-object condition and

prepositional condition ($M = 31.7\%$, $M = 26.9\%$ respectively; $U = 310$, $Z < 1$, $p > .5$). In the

direct-object condition, the most common non-motion verbs were *cleaning*, *painting*, *reading*

and *watching*. In the prepositional condition the most common non-motion verbs were *standing*,

sitting, *looking* and *hiding*.
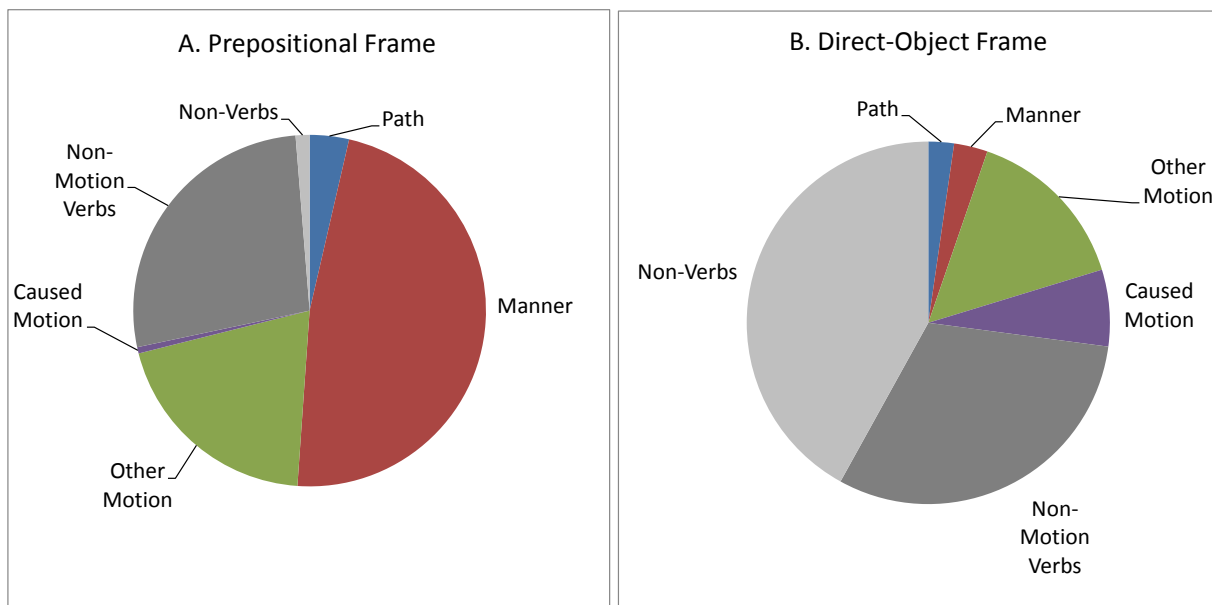
Figure 2: Responses in the sentence completion task. Participants were given a sentence with
the verb replaced by a blank and asked to provide the missing word.

"Caused-motion Verbs" included all responses where the verb encoded motion but the

noun phrase in the complement was a theme (an object that moved) rather than the ground that

defined the path (e.g., "The man is rolling the log"). Caused-motion verbs were more frequent in

the direct-object condition ($M = 6.6\%$, $M = .6\%$ respectively; $U = 500$, $z = 4.37$, $p < .001$). Some

of these responses also encoded the manner of motion (e.g., *spin*), a few encoded the path (e.g.,

lift), but most encoded neither (e.g., *move*).

"Manner Verbs" included all responses where the sentence described motion with respect

to a ground object and the verb specified the manner of motion. These responses were far less

frequent in the direct-object frame than in the prepositional frame condition ($M = 2.9\%$, $M =$

47.2% respectively; $U = 576$, $z = 5.93$, $p < .001$).

"Path Verbs" included all responses where the sentence described motion with respect to

a ground object and the verb specified the path of motion. These responses occurred with similar

frequency in the direct-object and prepositional conditions ($M = 2.2\%$, $M = 3.6\%$ respectively; U

$= 363.5$, $z = 1.56$, $p = .12$).

Finally, "Other Motion Verbs" included all responses where the sentence described

motion with respect to a ground object but the verb did not specify either the manner or the path

(e.g., go and come) or can specify both of these components (e.g., climb). These responses were

numerically less frequent in the direct-object frame than in the prepositional frame, but this

difference was not reliable ($M = 15.2\%$, $M = 20.2\%$ respectively; $U = 380$, $z = 1.90$, $p = .059$).

To address our four hypotheses, we conducted two analyses, which focused on the verb

responses to ensure that any differences that we observed were not solely due to the greater use

of prepositions in the direct-object condition.

First, we examined the degree to which participants used these frames to express motion

with respect to a ground object. We calculated the proportion of all verb responses that were

Path Verbs, Manner Verbs or Other Motion Verbs. These verb classes represent the concepts

that would be relevant in our experimental task in which the video clearly shows the agent

moving relative to a still ground object. In the prepositional condition, 72% of the verb

responses encoded motion with respect to a ground object, while in the direct-object condition only 35% of the verb responses did so. This resulted in a reliable difference between conditions ($U = 576$, $z = 5.94$, $p < .001$), a reliable preference for directed motion verbs in the prepositional frames condition ($W = 300$, $z = 4.28$ $p < .001$), and a reliable preference for other kinds of verbs in the direct-object condition ($W = 243$, $z = 3.94$ $p = .001$). These findings confirm our first two hypotheses: Sentences with prepositional frames are often interpreted as descriptions of directed motion events. In contrast, sentences with direct-object frames are generally interpreted as encoding other kinds of events and states.

Next we examined the degree to which each frame was biased toward a manner or path interpretation, given that a directed motion verb was used. Path Verbs were coded as "1" and Manner Verbs were coded as "0", to conform to the coding used in the word learning experiments (Author et al., under review). The Other Motion Verbs were coded as 0.5, because the verbs in this set either contain no information about manner or path (like "move") or can encode either manner or path (like "climb"), and thus should not contribute to a systematic bias toward either dimension. The bias score was calculated by averaging the scores for the relevant verbs for each participant. One participant from the direct-object condition produced no responses of the relevant kinds and was excluded from the analysis.

Participants in the prepositional frame condition showed a robust preference for manner verbs, resulting in a bias score that was reliably below chance ($M = .19$; $W = 300$, $z = 4.28$ $p < .001$). In contrast, participants in the direct-object condition showed no systematic preference for path or manner ($M = .49$; $W = 33$, $z < 1$, $p > .3$). As a result, the bias scores were reliably different in the two conditions ($U = 549$, $z = 5.49$, $p < .001$). The strong bias in the prepositional frame for encoding manner was even more marked when the Other Motion Verbs were excluded

from the analysis ($M = .07$). Only 14 of the 24 participants in the direct-object condition

produced any Manner or Path Verbs, but those who did showed no systematic preference for

either verb type ($M = .43$). These findings confirm our remaining two hypotheses: When

prepositional frames are used to describe directed motion events, adults systematically prefer to

use verbs that encode manner. However, when direct-object frames are used to describe directed

motion events, adults do not have a strong preference for encoding either manner or path.

<div align="center">

Study 4: Grammaticality Judgments
</div>

Our experiments with children (Experiments 3 & 4; Author et al., under review) used the

prepositional frame rather than the direct-object frame. This design choice reflected two

assumptions. First, we believed that, although the prepositional frame is highly biased toward

manner verbs, it is acceptable to use it with path verbs in English despite the inevitable

redundancy. In contrast, we assumed that manner verbs would typically be unacceptable in

direct-object frames. These two assumptions receive some support from our corpus analyses

(Study 1 & Study 2). However, to test them directly, we collected acceptability judgments from

adults.

Method

Participants. Thirty-two native English speakers were tested on Amazon Mechanical

Turk, eight in each list. No additional demographic information was requested or collected. All

participants correctly answered 80% or more of the control questions ($M = 95\%$, responses of 1-3

counted as correct for unacceptable sentences and responses of 5-7 counted as correct for

acceptable sentences).

Stimuli. The critical sentences that were used in Study 3 were modified for this

experiment. For each sentence, we created two versions, one with a path verb and one with a

manner verb. The path verbs were matched to the preposition (and path) that had been used in the original video, to ensure that the sentences were plausible and that the verb and preposition (if any) were compatible. This resulted in five verbs, one of which was used with two sets of events (*ascend-up, descend-down, pass-by, circle-around, enter-into, pass-over*). The manner verbs consisted of six manners of motion which could be done by a human agent along a variety of paths. They were assigned to the 6 sets of 6 frames such that the same manner verb never appeared with the same preposition. This process resulted in four versions of each sentence (see 9 - 12).

9. Direct Object Path Condition: The teacher is descending the ramp.

10. Prepositional Path Condition: The teacher is descending down the ramp.

11. Direct Object Manner Condition: The teacher is skipping the ramp.

12. Prepositional Manner Condition: The teacher is skipping down the ramp.

We used a within subjects design with Latin Square counterbalancing. Thus each participant saw nine sentences of each kind, but only one version of each base sentence. Twenty filler sentences were created that were highly plausible and acceptable ("The doctor gave her patient some medicine") or were highly implausible ("The king rode his house through the town"). A single randomized order was used.

Procedure. Participants were instructed to read each sentence and rate how acceptable it was on a scale of 1 to 7, where 1 was defined as "absolutely unacceptable", 7 as "completely fine", and 4 as "somewhat acceptable".

Results and Discussion

The ratings for the high and low acceptability fillers indicated that participants were using the scale appropriately and understood the task ($M = 6.81$, $M = 1.72$, respectively). The

acceptability ratings for the critical items appear in Figure 3. In the prepositional frame, both manner and path verbs were rated as highly acceptable ($M = 5.71$, $M = 5.79$, respectively), as were path verbs in the direct-object frame ($M = 5.99$). As expected, manner verbs in the direct-object frame were generally found to be unacceptable ($M = 2.51$).
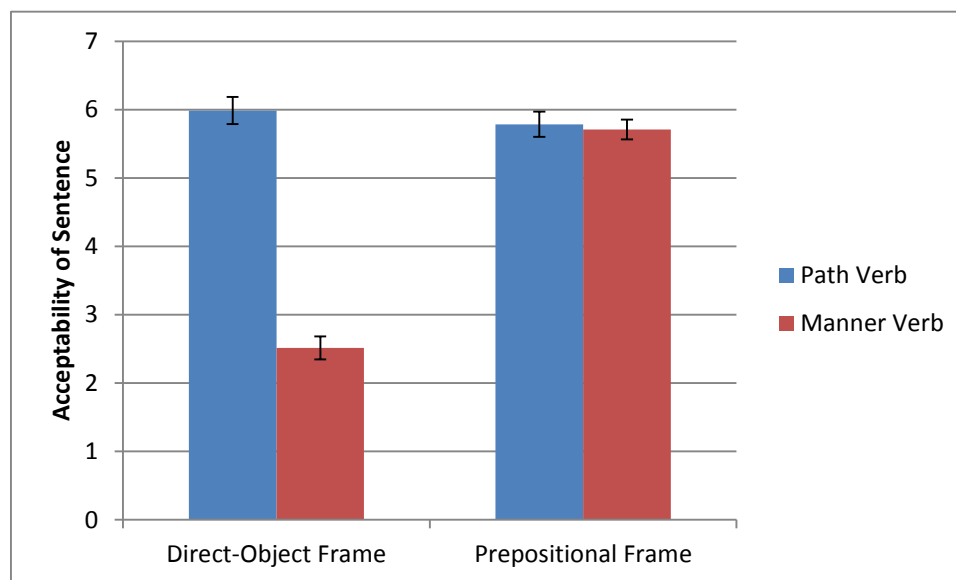
Figure 3: Acceptability of critical sentences on a 7 point scale. 7 = completely fine; 1 = absolutely unacceptable.

Responses were analyzed with a 2x2 ANOVA with both frame and verb type as within-participant variables. There was a reliable effect of frame and a reliable effect of verb type, which were superseded by a robust interaction (all F's > 80, all p's < .0001). This pattern of effects was driven by the low ratings given to the manner verbs in the direct-object frames. There was no difference between the two verb types in the prepositional frame ($W = 37$, $z < 1$, $p > .5$), but there was a robust difference for the direct-object frame ($W = 526$, $z = 4.91$, $p < .0001$). Path verbs were equally acceptable in either frame ($W = 78$, $z < 1$, $p > .3$), but manner verbs were considerably less acceptable in the direct-object frame ($W = 528$, $z = 4.93$, $p < .0001$). These results support our working assumption that both verb types are grammatical in the prepositional frame, but manner verbs are not grammatical in the direct-object frame.

General Discussion

Taken together these four studies strongly support the assumptions that guided the design of our experiments on the flexibility of verb lexicalization biases and our interpretation of the findings (Author et al., under review). In these studies, we found that participants who heard verbs in the prepositional frame had an initial bias to interpret them as manner verbs, while those who heard verbs in the direct-object frame did not, consistent with prior research (Naigles & Terrazas, 1998). We also found that participants in the prepositional frame condition continued to show effects of this manner bias when the input in the experiment was probabilistic, while those in the direct-object condition quickly adapted to whatever input they were given. We proposed that these differences between the frames could be seen as a rational response to distributional properties of the English. Specifically, we suggested that the prepositional frame is frequently used to encode motion events in English and is often used with manner of motion verbs. Thus participants should have strong bias to interpret verbs in this frame as encoding motion. In contrast, we argued that the direct-object is not frequently used to encode directed motion in English and when it is used it appears with verbs of both types. Thus participants should initially have no strong bias for verbs in this frame. These assumptions about the distributional properties of English were grounded in previous studies using elicited production tasks (e.g., Slobin & Berman, 1994; Allen et al., 2007), but they go beyond what that work can support by making claims about the distribution of the forms in the language as a whole (as opposed to language produced in contexts that are designed to elicit motion verbs) and by focusing on the relationship between particular frames and particular verb types. The studies in this supplement validate these assumptions.

Specifically, we find that manner verbs are used with prepositional frames considerably more often than they are used with direct-object frames. This is true in both informal text (Study 1) and formal text (Study 2). Many of the violations of the pattern were caused by motion verbs with theme objects ("run a store", "slide the knob"), but there were also cases where the direct object clearly defined a path ("run the race", "jump the fence"). In contrast, the use of path verbs varied across contexts. In informal, unedited text path verbs were more likely to occur in prepositional frames than in direct-object frames (Study 1). However, in published books, path verbs were used far more frequently in direct-object frames (Study 2).

Evidence for the differences in the use of these frames was also apparent in the completion task. Participants who were given a prepositional frame often completed the sentence with a verb of motion, and this verb was typically one that encoded manner. Participants who were given a direct-object frame instead produced static prepositions or other verbs. When they did produce a verb of motion it was as likely to encode manner as it was to encode path. These studies support our working assumption that prepositional frames will have a strong manner bias that may persist to some degree even after training, while direct-object frames will have no initial bias and will quickly come to reflect the experimental input (see Author et al., under review, Experiments 1 & 2).

Finally, these results justify our decision to use the prepositional frame for our initial experiments with children. Study 3 demonstrates that the prepositional frame is strongly associated with motion, while the direct-object frame is not. This is critical because, if our participants are primarily considering other meanings for the novel verbs (e.g., *seeing*, *liking*, *touching*), they may have difficulty identifying the correct meaning in either the manner or path conditions, and thus may fail to learn the verbs. If they do not learn the verbs, then we cannot

expect the training manipulation to have a strong effect on their biases, and thus the experiment would not be a valid test of the hypothesis. Furthermore, Study 4 suggests that, for adults at least, both manner verbs and path verbs are grammatical in the prepositional frame. Thus, while the frame may be highly biased toward the manner interpretation, prepositional frames with path verbs are consistent with the adult grammar. Children were able to learn both kinds of verbs in this context (see Experiments 3 & 4) suggesting that both hypotheses were accessible for them.

In contrast, adults found manner verbs in direct-object frames to be unacceptable. Nevertheless, adults were able to learn both kinds of verbs in the direct-object frame and develop both manner and path biases (see Experiment 2). Thus they must be capable of generating motion verb meanings in this frame when given sufficient contextual support. We assume that, with sufficient support and time, young children would also be able to learn these mappings, since they are able to learn other mappings that are absent or ungrammatical in their language (Casenhiser & Goldberg, 2005). However, we have reason to believe that they cannot reliably learn path or manner verbs in the transitive construction, given the limited evidence available in our studies. Specifically, after conducting the first study with children (Experiment 3; Author et al., under review), we conducted a pilot study with a small sample of children using the direct-object frame. We found that they had difficulty learning both manner and path verbs and often insisted that both of the test scenes were instances of the verb (as would be expected if they had interpreted the verb as *see* or *like*). Using the verbs with path prepositions may help children focus on the relevant semantic field.

References

Allen, S., Özyürek, A., Kita, S., Brown, A., Furman, R., Ishizuka, T., & Fujii, M. (2007).

    Language specific and universal influence in children's syntactic packaging of Manner

    and Path: A comparison of English, Japanese, and Turkish. *Cognition, 102*, 16-48. doi:

    10.1016/j.cognition.2005.12.006

Aske, J. (1989). Path predicates in English and Spanish: A closer look *Proceedings of the 15th

    Annual Meeting of the Berkeley Linguistics Society* (pp. 1-14). Berkeley, CA: BLS.

Berman, R., & Slobin, D. I. (1994). *Relating events in narrative: A cross-linguistic

    developmental study*. Hillsdale, NJ: Erlbaum Associates, Inc.

Casenhiser, D., & Goldberg, A. (2005). Fast mapping of a phrasal form and meaning.

    Developmental Science, 8, 500-508. doi: 10.1111/j.1467-7687.2005.00441.x

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.

Jackendoff, R. S. (1990). *Semantic structures*. Cambridge, MA: MIT Press.

Michel, J.-B., Shen, Y. K., Aiden, A. P., Veres, A., Gray, M. K., The Google Books Team, . . .

    Lieberman-Aiden, E. (2011). Quantitative analysis of culture using millions of digitized

    books. *Science, 331*, 176-182. doi: 10.1126/science.1199644

Naigles, L. R., Eisenberg, A., Kako, E., Highter, M., & McGraw, N. (1998). Speaking of motion:

    Verb use by English and Spanish speakers. *Language and Cognitive Processes, 13*, 521-

    549. doi: 10.1080/016909698386429

Naigles, L. R., & Terrazas, P. (1998). Motion-verb generalizations in English and Spanish:

    Influences of language and syntax. *Psychological Science, 9*, 363-369. doi:

    10.1111/1467-9280.00069

Papafragou, A., & Selimis, S. (2010). Lexical and structural biases in the acquisition of motion

      verbs. *Language Learning and Development, 6*, 87-115. doi:

      10.1080/15475440903352781

Author, Author, & Author. (under review). On the plasticity of semantic generalizations:

      Children and adults modify their verb lexicalization biases in response to changing input.

Slobin, D. I., & Hoiting, N. (1994). Reference to movement in spoken and signed languages:

      Typological considerations. *Proceedings of the Twentieth Annual Meeting of the Berkeley*

      Linguistics Society (pp. 487–505). Berkeley, CA: Berkeley Linguistics Society.

Talmy, L. (1985). Lexicalization patterns: Semantic structure in lexical forms. In T. Shopen

      (Ed.), *Language typology and syntactic description: Grammatical categories and the*

      lexicon (pp. 57-149). New York: Cambridge University Press.