

"This printout has been approved by me, the author. Any mistakes in this printout will not be fixed by the publisher. Here is my signature and the date \_\_\_\_\_".

# Drawing Inferences about Novel Raising and Control Verbs

Misha Becker and Bruno Estigarribia  
University of North Carolina at Chapel Hill

## 1. Introduction

The core question we address is how children come to distinguish the syntactic structures of sentences which on the surface appear to be identical, but whose underlying structures are quite divergent. This situation arises in the following case, where the verbs *seem* and *claim* both occur in the frame in (1).

- (1) John seems/claims to be clever.
  - a. [John<sub>i</sub> seems [<sub>t<sub>i</sub></sub> to be clever]
  - b. [John<sub>i</sub> claims [PRO<sub>i</sub> to be clever]

The overlapping distribution belies crucial differences in the verbs' thematic properties. The structural difference between (1a) and (1b) is that in the first case, the raising verb *seem* assigns no external  $\theta$ -role, and the subject of the infinitive clause (*John*) raises up to matrix subject position. In the second case, the control verb *claim* assigns its external  $\theta$ -role to the subject, *John*, which is base-generated in the main clause and controls the reference of PRO in the infinitive. The question for language learners is how these two structures are determined, since, presumably, at some point a learner will encounter these abstract verbs for the first time. One way to illustrate the problem is to consider how you would interpret a sentence like *John gops to be clever*.

Certainly, it is the case that raising and control verbs are distinguished in other types of sentence frames. Raising, but not control, verbs can occur in a *there*-construction or with weather-*it* subjects, and control, but not raising, verbs can occur in pseudoclefts. Some control verbs (though not all) can occur in transitive constructions.

- (2) a. There seems/\*claims to be a problem.
  - b. It seems/\*claims to be raining.
- (3) a. What John \*seems/claims is to be clever.
  - b. John \*seemed/claimed victory.

A common learning assumption is that children reason across occurrences of verbs (Naigles, 1996; Bunker & Lidz, 2004). If this is true, hearing raising and control verbs in these respective environments should help learners discriminate the two classes. However, three considerations make us think a learning problem remains.

First, expressions with expletive subjects (as in (2)) are far less frequent in speech to children than expressions containing raising verbs and a referential subject. Hirsch and Wexler (2007) found that 87% of adult sentences containing *seem* contained a lexical (non-expletive) subject. Second, the diagnostic environments for control verbs are not entirely clear. A number of common control verbs

---

\* We would like to thank the UNC Humanities Division's Associate Professor Support Grant (to Becker), UNC's Cognitive Science Program support (to Estigarribia), and the assistance of Hugh Meriwether, Steve Balen, Molly Jabeck and Taylor Shirley.

cannot occur in a transitive frame (*\*John hoped victory; \*John tried the winner*), showing that transitivity is not a necessary condition for control. Even the pseudocleft construction is not perfectly well-formed for all control verbs (*??What John tried was to be the winner*). Third, a set of verbs that are ambiguous between being raising and control, such as *begin*, can occur in all of the syntactic environments that unambiguous raising and control verbs do (Perlmutter, 1970), implying that there are no conditions that are logically sufficient for categorizing a verb as raising or control (Becker, 2006). Otherwise stated, the existence of ambiguous verbs means that distributional evidence alone is inconclusive.

Yet, children grow up to be speakers who distinguish these classes of verbs. Our question, then, is which cues can help children to distinguish raising from control verbs under these conditions. We will examine the possibility that sentence-internal cues, such as subject animacy, and sentence-external cues, such as frequency of exposure or context of exposure, can modulate speakers' categorization guesses.

## 2. Background

We believe that subject animacy can serve as an important sentence-internal cue about sentence structure. The typologically prevalent tendency of subjects to be animate (Keenan, 1976) is psycholinguistically robust. Previous experimental work underscored the importance of subject animacy as a cue in influencing speakers' assumptions about the thematic relationship between subjects and verbs. Clark and Begun (1971) showed that when an animate sentence subject was replaced with an inanimate subject, participants rated the resulting sentence as having a lower acceptability score than if it was replaced with another animate subject, regardless of predicate type. Moreover, when a second group of participants were asked to modify sentences with various types of subjects to make the sentences more acceptable, participants were more likely to change an inanimate subject to make it animate, than they were to change an animate subject at all.<sup>1</sup>

Becker (2005) conducted a fill-in-the-blank experiment with adults in which subject animacy was systematically manipulated. When adults read a sentence like *The salesman \_\_\_\_\_ to advertise an interesting new product*, participants were more likely to provide a control verb (*claim/try/want*) than a raising verb. When presented with *The banner \_\_\_\_\_ to advertise an interesting new product*, participants showed the opposite pattern, offering raising verbs more frequently.

Children may also be influenced by subject animacy. One of the experiments reported in Becker (2006) showed that when the subject of the sentence is *inanimate*, children permit verbs that are typically control verbs to occur there (e.g., they allow sentences like *The flower wants to be pink.*) This result can be understood in one of two ways. One possibility is that this sentence frame is uninformative for the child and any verb would be accepted as grammatical in it. The other possibility is that this frame is highly informative—so informative, in fact, that it coerces the verb to drop its typical agent/experiencer-assigning theta grid and take on a more auxiliary-like function. We have reason to believe it is the second option that is correct. A further experiment (Becker, 2009) demonstrated a similar pattern in that young children permitted control verbs to occur with weather-*it* subjects. Posthoc interviews with the children in that experiment revealed that they interpreted the control verbs in these expressions as having a modal-like meaning (e.g., *It wants to rain ~ it will probably rain/it might rain*). This is precisely the sort of meaning that would result from coercion of the verb in this sentence frame—in other words, these children were being “frame compliant” (in the sense of Naigles et al., 1989; Lidz et al., 2004).<sup>2</sup>

---

<sup>1</sup> Certainly, the acceptability of a given noun as subject depends entirely on the predicate it is paired with. However, Clark and Begun note that even in the substitution task a higher proportion of human subjects were paired with verbs requiring human NPs (80%) than were abstract nouns paired with verbs requiring abstract NPs (20%).

<sup>2</sup> This is analogous to Gentner and France's (1988) “Verb Mutability Effect.”

However, none of these studies on subject animacy were, strictly speaking, *learning* studies. Becker (2005) asked adults to fill in a blank with a verb already in their lexicon. Becker's two previous studies with children asked children to judge whether a verb already in their vocabulary (or at least present in their target language) sounded okay in an unusual sentence frame.

In the present study, we remedy this by conducting a novel verb learning study with adults, in which, not only is *subject animacy* manipulated as a variable, but several other cues are as well. Specifically, we ask whether knowing that a novel verb has a raising-like or control-like *meaning* biases speakers' categorization, and whether the *frequency* or the linguistic *context* (here: within a story or in individual sentences), with which a verb is presented, influences how conclusive concurrent cues are. Let us now turn to our experimental procedure.

### 3. Experiment

We conducted a novel verb learning experiment with 186 adult native English-speakers. Participants were recruited through a mass e-mail sent to the entire campus community of UNC Chapel Hill. Participants included students, employees, spouses of UNC employees, and other members of the community. Ages ranged from 18 to 71 years (mean age 29), and gender breakdown was 134 females and 52 males. Participants were paid \$5 for their participation, which lasted between 5 and 30 minutes, depending on the condition they were assigned to.

#### 3.1. Design and Procedure

Participants were exposed to nine novel verbs. Three of the nine verbs were fillers: they were presented in either a transitive or an intransitive sentence. Three of the remaining 6 verbs were *raising* verbs, and the remaining three were *control* verbs. We will return to what this means below (section 3.2).

After each presentation of a stimulus (or stimulus set; see below), participants were asked to make a forced-choice grammaticality judgment about the verb in a new linguistic environment. They had to determine in which of two new sentences (not part of the initial exposure) the novel verb *sounded better*. We were interested in comparing how the following three independent factors influenced speakers' judgment about the category of a novel verb, and thus drive their sentence judgment:

##### 1. Cue type: 3 levels

- (a) Syntactic cue (*frame*): multiple presentations of the verb occur in different sentence frames. Depending on verb type (see below), a verb appears at least once with an inanimate subject or at least once in a transitive frame (the rest of the time appearing with an animate subject and infinitive predicate, as in (1)). Participant is not told what the novel verb means.
- (b) Semantic cue (*definition*): participant is told a made-up definition of the novel verb, but each occurrence of the verb is in the same sentence frame that accepts both raising and control verbs (animate subject, infinitive predicate, as in (1)).
- (c) Both (*frame+definition*): participant receives both types of cues (sentence frame varies across multiple presentations of the verb, and participant is given a lexical definition).

##### 2. Frequency: 3 levels

- (a) 1 exemplar: each novel verb is presented only once, and after this single occurrence the participant has to perform their sentence judgment.<sup>3</sup>
- (b) 3 exemplars: each novel verb is presented three times.

---

<sup>3</sup> Obviously in this case there is no room for sentence frame variation, as described in (1a) above.

- (c) 5 exemplars: each novel verb is presented five times. In the 3- and 5-exemplar conditions the sentence judgment task is done after reading all occurrences of the verb.

3. Linguistic context: 2 levels

- (a) Story: the novel verb is used in the context of a (very) short story. Stories often had a fairytale quality to them or otherwise were appropriate for children. (To justify this, participants were told in advance that the stimuli would be later used with children.)
- (b) List: multiple presentations of each verb were in a list of unrelated sentences. In the single-presentation condition there was just a single sentence presented.

These three factors were fully crossed, for 18 different conditions. The design is shown in Table 1.

Table 1  
*Design of Experiment*

	1 exemplar			3 exemplars			5 exemplars		
	Def	Frame	Both	Def	Frame	Both	Def	Frame	Both
List									
Story									

Each of these factors was manipulated between participants, and each participant was assigned randomly to a condition. Our power analysis suggested a minimum of 10 participants in each cell, for a minimum of 180 participants.<sup>4</sup>

All stimuli were presented on a computer. After providing informed consent, participants were told to read a set of instructions on the screen. The instructions explained that the participant would read a series of sentences or stories containing a made-up word, and that, after reading the sentence(s) or story, they should choose between two new sentences containing this made-up word; their choice would be based on which of the two sentences sounded better or more natural to them. First, participants underwent a brief training session in which they had to give forced-choice grammaticality judgments for sentences containing only English words (e.g., *The lady slept for a long time*; *The lady slept the baby*). At least 8 correct answers were required to continue in the experiment. Only one participant failed to reach this criterion and that participant’s data was excluded from the analysis. Then the test session began. For the *raising* and *control* verbs, participants had to choose between a pseudocleft (e.g., *What John gorpded was to eat pie*) and a *there*-construction (e.g., *There gorps to be a book on the table*).

3.2. *Raising and Control verbs*

It might seem strange to label novel, made-up verbs as belonging to the class of raising or control verbs. What do we mean by this distinction?

The syntactic distribution of raising and control verbs is not unrelated to their lexical meaning. Raising verbs generally have meanings that relate to appearance (*seem, appear*), aspect (*used (to), tend, going (to)*), and states of affairs (*happen (to), turn out*). These sorts of lexical meanings do not easily (if at all) admit an agentive or experiencer type of subject. Instead, these predicates are predicated of other predicates, rather than arguments. This is consistent with the fact that, syntactically, raising verbs select a clausal complement and assign no external  $\theta$ -role.<sup>5</sup>

<sup>4</sup> The program presented the conditions in random order. We monitored which conditions had been presented so that when we reached 10 participants in a given condition we could remove that condition from the set. However, due to a problem in implementing the program certain filled conditions were not actually removed and hence included a few extra participants. This accounts for our actual number of 186 participants.

<sup>5</sup> A reviewer wonders why it is that precisely this set of verbs allows raising; that is, why one cannot say *Bill*

Control verbs, on the other hand, have meanings that relate to desire (*want*), effort (*try*), declaration (*claim*), emotions (*love, hate, like*), and other mental states or faculties (*decide, choose*). These predicates not only assign an external  $\theta$ -role, but they require an animate, sentient subject.<sup>6</sup> Therefore, we created definitions that were consistent with these parameters. Our made-up definitions were in fact modeled on actual definitions of English raising and control verbs, though we attempted to make the definitions slightly different (e.g., more restrictive), in the hopes that participants would not simply insert a known English verb into the stories or sentences. Note that, for example, *joop* is defined as ‘to look a certain way,’ and when it is used in a sentence, such as *The old man joops to be very tired*, the infinitive predicate specifies what that ‘certain way’ is. We acknowledge that one could also define a verb like *seem* as ‘to look/appear a certain way,’ and it is possible that speakers simply replaced *joop* with *seem* when they read the sentences, and we cannot rule out with certainty the possibility that our participants did this. We will return to this issue in Section 5. To give a slight preview of our findings, the inclusion of explicit definitions made no difference in participants’ responses.

We list the novel verbs we used and their definitions in Table 2.<sup>7</sup> Note that in the *frame-only* condition participants were not given the definition of any verbs. In this condition, it was the raising verbs that appeared at least once with an inanimate subject while the control verbs appeared at least once with a transitive frame, the remainder of the time appearing in ambiguous frames. Thus, the distinction between the two categories was maintained across conditions.

Table 2  
*Novel Verbs and Definitions*

<i>Raising Verb</i>	Definition	<i>Control Verb</i>	Definition	Filler	Definition
joop	to look a certain way	rickle	to really dislike being someplace	ballop	to swing your arms in circles
meb	to probably be a certain way	sart	to make a big effort to be some way	flim	to breathe fire and be scary
trollick	to be some way most of the time	zid	to really enjoy being someplace	frell	to touch something very gently

Below are some concrete examples of our stimuli. The following item is from the 3-exemplar, *definition+frame*, story condition, and contains a control verb.

(4) **sart: to make a big effort to be some way**

Once there was a very fearful dinosaur named Hugh. Hugh was afraid of everything! He was afraid of mammoths, saber-toothed tigers, volcanoes, lakes, and even of his own shadow. His classmates at the dino-school always made fun of him. So when the new dino-student, Melissa, came to school, he sartered to be courageous and talk to her. She was so pretty with her long neck and spikey tail. When he finally sartered a brave attitude and said hi, she looked away

---

*forced to go downtown* meaning “something forced Bill to go downtown.” The reason this sentence is ungrammatical is the same as the reason that *Bill seemed to go downtown* does not mean “Something seemed Bill to go downtown.” In other words, it is inextricably linked to the fact that raising verbs, by definition, fail to select an external argument.

<sup>6</sup> We know of only two control verbs that allow an inanimate subject, *serve* and *help*.

<sup>7</sup> Some of these novel verbs come from the Comprehensive Test of Phonological Processing (Wagner et al., 1999).

and blushed. Hugh realized she was shy but she liked him, so he invited her to have a glass of iced lava at the new dino-diner in town. The next time he saw her he still felt shy, but he started to be more outgoing and talk to her again. They became friends and now Hugh is a very brave dinosaur.

(5) Test sentences:

- a. What Hugh started was to be brave.
- b. There started to be a new kid at school.

Two buttons appeared on the computer screen corresponding to the first and the second sentence, and participants had to click on the button of their choice in order to proceed to the next item. The order of pseudocleft and *there*-construction was counterbalanced. The following example illustrates an item from the 5-exemplar, *frame-only*, list condition and contains a raising verb.

(6) **joop**

- a. The old man joops to be very tired.
- b. The book joops to be very long.
- c. It joops to be sunny outside.
- d. That mountain joops to be too steep to climb.
- e. It joops to be about to rain.

(7) Test sentences:

- a. What the fairy joops is to be small.
- b. There joops to be a computer on the desk.

In example (6), no definition is given corresponding to the *frame-only* condition, but the verb is presented with a variety of subject types: one animate subject, two inanimate referential subjects and two expletive subjects (compatible with raising). In (4), a definition is given before the story, and the verb is also presented in more than one sentence frame: two occurrences are with an animate subject and an infinitive predicate, and one occurrence has a transitive frame (compatible with control). In the *definition-only* condition participants would have been given a definition, as in (4), but every occurrence of the verb would have been with an animate subject and an infinitive predicate, regardless of whether the verb was *raising* or *control* (our categorization).

Because the factors of cue type, frequency, and context were manipulated between participants, all nine of the test items viewed by a given participant were of the same type (e.g., 3-exemplars, frame-only, story, or whatever). Because the factor of verb type was manipulated within participants, each participant was given 3 raising verbs, 3 control verbs, and 3 (transitive or intransitive) fillers.

### 3.3. Predictions

Participants' task was to choose whether the novel verb they had just encountered sounded better in a pseudocleft sentence or a *there*-construction. We assume that since pseudoclefts admit control verbs, but not raising verbs in English, if the participant selects the pseudocleft as sounding better, they have categorized the novel verb as a control verb. And since *there*-constructions admit raising, but not control verbs in English, if a participant selects the *there*-construction as sounding better, they have categorized the novel verb as a raising verb.

Given past work on subject animacy, we predicted that encountering a verb used with an inanimate subject would lead participants to choose the *there*-construction as sounding better (i.e., categorize the verb as a raising verb). On the other hand, encountering a verb used only with animate subjects should lead participants to choose the pseudocleft (i.e., categorize the verb as a control verb), since there is a strong underlying assumption that animate subjects are agents or experiencers. Note that this need not be the case, given that raising verbs freely admit animate subjects.

Before proceeding, let us flesh out this prediction a bit more. Generally speaking, verbs may take one, two or three arguments, and the category of the internal argument (the complement of the verb)

may be a Noun Phrase (NP) or a sentence (IP/CP).<sup>8</sup> In addition, certain frames require an external argument: transitive and ditransitive verbs require both an external argument and one or two internal arguments. Moreover, there is an assumption that external arguments tend to be animate and tend to be mapped to an Agent thematic role (supported by robust crosslinguistic patterns (Keenan, 1976); see also Jackendoff, 1972; Pinker, 1984).

However, when a verb is intransitive, taking only a single argument, or when a verb takes an infinitive complement, there may or may not be an external argument projected in the underlying representation. An intransitive verb may be either unergative (single argument is external) or unaccusative (single argument is internal) (Perlmutter, 1978; Burzio, 1986). Also, when a verb takes an infinitive complement, it may also project an external argument, as control verbs do, or it may not, as with raising verbs (i.e., in languages like English, a surface subject must raise from inside the infinitive complement). Therefore, in these cases children must determine whether the surface subject is also the external argument of the verb (i.e., the ‘deep’ subject).

Due to the assumption that subjects will be animate and will be Agents, the *violation* of this expectation can provide a powerful cue that the surface subject is not in fact the ‘deep’ subject. That is, in sentences in which the surface subject may or may not be the external argument of the verb (unaccusatives and raising verbs), an *inanimate* subject serves as a strong cue that the subject is not an external argument. Instead, it is derived from an internal position. This is why we predict that the occurrence of an inanimate subject will lead speakers to assume that the novel verb is a raising verb.<sup>9</sup>

Now let us proceed with our other predictions. Since the meaning of a verb is intimately connected to its status as either a raising or a control verb in English, we also predicted that knowing the lexical meaning of the novel verbs would influence participants’ categorization. Thus, in the *definition-only* condition, participants should choose the *there*-construction for verbs that had a raising-like definition (e.g., ‘to look a certain way’) and the pseudocleft for verbs that had a control-like definition (e.g., ‘to really enjoy being someplace’). Moreover, we hypothesized that these cues would have an additive effect, so that participants in the *definition+frame* condition would be even more likely to choose the *there*-construction, for raising verbs, but the pseudocleft, for control verbs.

As for the factors of frequency and linguistic context, we did not have a straightforward prediction of how these cues would affect participants’ responses. We supposed that the frequency cue could work in a couple of ways. One possibility is that multiple exposures to a given verb would provide extra evidence for the verb’s categorization. As such, participants would be even more likely to choose the *there*-construction or the pseudocleft, respectively, in the 5-exemplar condition than the 1-exemplar condition, while the 3-exemplar condition would be in the middle. If multiple exposures work this way, we might also expect to find less variability across participants in the 5-exemplar condition than the 1-exemplar condition. Another possibility is that participants commit to a categorization of the verb upon the first exposure, and subsequent exposures are distracting, or contribute to the participants’ working memory load, or otherwise make categorization more difficult. If this is true, we might expect to find fewer choices of the *there*-construction overall in the 5-exemplar condition compared to 1-exemplar, on the assumption that participants would then be more likely to assume the verb takes an agentive subject. We, also, might expect to find *more* variability in responses in the 5-exemplar vs. 1-exemplar conditions in this case.

We did not have a strong prediction about the effect of the story vs. list context factor. It was included because we intend to adapt this experiment for use with children (see section 6), and we were curious to see whether embedding the novel verbs inside a story context made the learning process more natural and easier (i.e., as indicated by more ‘as-expected’<sup>10</sup> responses in the story condition than the list condition).

---

<sup>8</sup> As noted by Gleitman et al. (2005) although there are a great many individual verb meanings, (most) verbs fit into one of a relatively small number of sentence frames.

<sup>9</sup> Please see Becker and Schaeffer (under review) for an extension of this argument to unaccusatives.

<sup>10</sup> See Section 4.

To summarize, we make the following predictions:

1. Frame should be a strong cue: in particular, given crosslinguistic preferences for subjects to be animate, an inanimate subject should lead participants to categorize a novel verb as *raising*.
2. Definition should be facilitative, and definition plus frame together should have an additive effect (leading to more ‘as-expected’ categorizations than either frame or definition alone).
3. Frequency: 2 possibilities
  - (a) More exemplars should also be facilitative (more ‘as-expected’ categorizations with more exemplars)
  - (b) More exemplars place an additional burden and distract from initial categorization (fewer ‘as-expected’ categorizations with more exemplars)
4. Story vs. list: We conjectured that the story context would facilitate encoding of the meanings and properties of the novel verbs; additionally, this is a more ecologically valid environment to encounter new words, at least for children; however, it could also be the case that it would actually distract participants from focusing on the verbs’ properties.

## 4. Results

Participants’ responses were labelled *correct* (or ‘as-expected’) if they chose the *there*-construction for a novel raising verb, and if they chose the pseudocleft for a novel control verb; *incorrect* otherwise. We ran two analyses of the data. First, we performed an analysis on the correct/as-expected responses. Second, we performed an analysis on the *raising* responses (i.e., choice of the *there*-construction). Let us consider the analysis of as-expected responses first.

### 4.1. As-Expected Responses

We asked which cues are more effective in yielding correct guesses. A mixed effect logistic regression model with interactions between verb type (raising vs. control) and all other factors, and a model without the interaction terms, were fitted. Participants are a random effect; all other effects are fixed. A likelihood-ratio test showed the interactions to be significant,  $\chi^2(8, 13) = 26.87$ ,  $p < .0001$ . The model predicted percentages are given in Table 3. The estimates and significance tests of fixed effects for the model, allowing verb type to interact with the other factors, are shown in Table 4.

Table 3  
*Predicted Probability of Correct Responses*

		1 exemplar			3 exemplars			5 exemplars		
		Def	Frame	Both	Def	Frame	Both	Def	Frame	Both
List	R	66.7	86.2	82.7	59.7	82.3	78.0	56.5	80.2	75.6
	C	<b>96.9</b>	89.9	95.8	97.9	92.9	97.1	99.5	98.2	99.3
Story	R	70.0	87.9	84.8	63.3	84.4	80.5	60.2	82.5	78.3
	C	99.0	96.6	98.7	99.3	97.7	99.1	99.8	99.4	99.8

Table 4  
*Mixed effect logistic regression estimates and significance tests*

	Parameter	Standard Error	z-value	p
Intercept	3.44	.68	5.08	<.0001
3 exemplars	.38	.59	.64	.52
5 exemplars	1.82	.92	1.99	.047
Story	1.17	.60	1.95	.052
Frame cue	-1.25	.72	-1.74	.082
Both cues	-.32	.81	-.39	.695
<i>Raising verb</i>	-2.74	.69	-4.01	<.0001
3 ex - <i>Raising</i>	-.68	-.61	-1.12	.264
5 ex - <i>Raising</i>	-2.25	.93	-2.43	.015
Story - <i>Raising</i>	-1.02	.61	-1.66	.097
Frame - <i>Raising</i>	2.39	.73	3.27	.001
Both - <i>Raising</i>	1.19	.82	1.45	.147

The intercept (in boldface in Table 3) corresponds to the case with 1 exemplar of a control verb, presented with a definition and no story context (list of isolated sentences). For this case, the probability of a correct guess is estimated at 96.9%. Using the “divide by 4” rule of thumb (Gelman & Hill, 2007, p. 82), we roughly estimate that seeing 5 exemplars of a control verb makes correct guesses 46% more likely ( $1.84/4 = .46$ ). The corresponding estimated probability is 99.5%.

Similarly, seeing the novel control verb in a story makes correct guesses 29% more likely ( $1.17/4 = .29$ ), with an associated probability of 99% (note this coefficient approaches significance at the 95% level). Conversely, not seeing a definition in the control case makes correct guesses 31% less likely (probability 90%). From Table 3, we can see that if the novel verb is compatible with a control categorization, the probability of a control guess is usually 90% or higher.

With 1 exemplar of a raising verb, presented with a definition and no story context (list), the probability of a correct guess is estimated at 66.7%. With 5 exemplars (definition provided and list context, but no frame cue), correct raising guesses are made in 56.5% of cases. This suggests that seeing more exemplars where the subjects are animate drives participants to the default control assignment. However, if the raising verb appears with an inanimate subject (1 exemplar, frame-only), the probability of a correct guess is now 86.2% for a list presentation and 89.8% for presentation in a story.

To summarize, it appears that adults perform extremely well (i.e., ‘as expected’) in all cases where the novel verb is compatible with a control categorization. Control seems to be a *default* assignment for verbs encountered in these ambiguous constructions. If the novel verb is compatible with a raising categorization, their performance is closest to optimal if at least one subject is inanimate, but they make incorrect (control) guesses the more animate subjects they see accompanying the novel raising verb.

#### 4.2. Raising Responses

Since our previous analysis suggested control is a default assumption, we ask here which cues make raising categorizations more likely, thus overriding a control assignment. The response here is *control* or *raising* categorization (i.e., choice of the pseudocleft vs. *there*-construction, respectively), rather than *correct* vs. *incorrect* as above. We fitted a mixed effect logistic regression model with interactions between verb type and all other factors and the same model without the interaction terms. A likelihood-ratio test showed the interactions not to be significant,  $\chi^2(8, 13) = 8.40$ ,  $p = .14$ . This indicates the probability that adults will assign a new verb to the raising or control category, which, in turn, is independent of the compatibility of the novel verb (i.e., how we categorized our novel verbs). The predicted probabilities are shown in Table 5, and the estimates and significance tests for the model without interactions are shown in Table 6.

Table 5  
*Predicted Probabilities of Raising Categorizations*

		1 exemplar			3 exemplars			5 exemplars		
		Def	Frame	Both	Def	Frame	Both	Def	Frame	Both
List	R	70.9	88.7	84.1	63.5	84.8	79.0	56.4	80.6	73.7
	C	<b>1.6</b>	4.8	3.3	1.1	3.5	2.4	0.8	2.6	1.8
Story	R	69.6	88.0	83.2	62.0	84.0	78.0	54.9	79.6	72.5
	C	1.5	4.6	3.1	1.1	3.3	2.3	0.8	2.5	1.7

Table 6  
*Mixed effect logistic regression estimates and significance tests*

	Parameter	Standard Error	z-value	p
Intercept	-4.14	.39	-10.55	<.0001
3 exemplars	-.33	.31	-1.08	.28
5 exemplars	-.63	.92	1.99	.047
Story	-.06	.25	-.24	.81
Frame cue	1.16	.31	3.81	.0001
Both cues	.77	.31	2.52	.012
<i>Raising verb</i>	5.03	.30	17.03	<.0001

The intercept, again (boldface), corresponds to the case with 1 exemplar of a control verb, presented with a definition and no story context (list). For this case, the probability of a raising categorization is estimated at 1.6% (hence, control assignments are expected 98.4% of the time). The remainder of Table 5 can be read in an analogous way.

We interpret the model estimates and significance tests as follows: exposure to 5 exemplars instead of 1 exemplar results in 16% fewer raising guesses overall ( $z = -2.00$ ,  $p = .046$ ). Having frame cues with a definition ( $z = 2.52$ ,  $p = .012$ ) or without ( $z = 3.81$ ,  $p = .0001$ ) results in 20% to 30% more raising guesses than in the condition with only a definition (and no frame cues). Providing a story is not significantly different from providing a list of sentences ( $z = -0.24$ ,  $p = .81$ ). Finally, as expected, if the novel verb is raising-consistent (either by virtue of having a raising-compatible definition or by virtue of occurring with an inanimate subject) the likelihood of a raising guess increases ( $z = 17.03$ ,  $p < .0001$ ).

To summarize, there is a strong default categorization to control, which is resistant to *raising* definitions and sensitive to the frequency of exposure. Specifically, the more instances adults see of novel verbs with animate subjects, the more likely they are to give a control interpretation (as determined by their choice of the pseudocleft), even if the definition is compatible with raising-like meanings. The only cue that can override this pattern is exposure to a new verb with inanimate subjects.

## 5. Discussion

Overall, the adult participants in our novel verb learning study showed a strong bias to choose the pseudocleft response (i.e., categorizing novel verbs as control verbs, as opposed to raising verbs). Note that this is, in some sense, expected since control verbs are typical, in that their syntactic subjects are their semantic subjects. This bias was overridden to a significant extent *only* when participants encountered the novel verb with an inanimate subject. Experiencing a novel verb with at least one inanimate subject led participants to guess the novel verb was a raising verb, and to choose the *there*-construction as sounding better. This result is consistent with our first prediction.

One of our predictions was not borne out. We had predicted that providing a lexical definition would help participants correctly categorize verbs. Instead, we found that when participants were provided with a definition and no other cues, they had the *least* success correctly categorizing raising verbs (compared to the other two cue types). When both the definition cue was provided and the

sentence frame was varied (the verb appeared with at least one inanimate subject), the probability of a correct categorization of a raising verb was slightly lower than when the frame cue was given alone. Therefore, the definition appeared to be unhelpful by itself and did not facilitate responses when paired with the frame cue.

Let's consider why this might be. We can think of two reasons the definitions were detrimental to performance on the task. One possibility is that keeping the lexical definition in mind while reading the sentences/stories and evaluating the test sentences placed a higher working memory load on participants, which in turn led to more *default* (i.e., control) interpretations of the novel verb. In the absence of a formal definition, participants were free to conjure their own likely meaning of the verb based on its use in sentence contexts, and this would not require keeping additional information in memory. Another possibility is that receiving a definition is simply not the way anyone learns the argument structure of predicates. Children could not possibly learn word meanings as rapidly as they do if they relied on hearing a definition of each one (or even a large fraction of them), and it is well established in the word-learning literature that children have other means of acquiring word meanings, such as fast-mapping (Carey, 1978), constraints such as whole object and mutual exclusivity (Markman, 1994), as well as information from sentence frames (Gleitman, 1990). In fact, it is precisely the sentence frame type of information that is known to be most informative for learning abstract verbs (Gleitman et al., 2005). Thus, speakers might learn not to rely on explicit definitions (rarely given) to learn argument structure. This result has quite strong implications for the syntactic bootstrapping argument. This learning procedure is based on the notion that children rarely have sufficient experience from the observation of events (or other non-linguistic means), to derive the meanings of verbs, and verb meanings, therefore, must be learned via their syntactic privileges (Gleitman, 1990). Our manipulation shows that, even if learners are provided with a definition, bypassing the need for extracting a meaning from the observation of an event, this still does not facilitate the learning of certain verbs. Properties of the arguments seem to trump lexical meaning information in the learning of abstract verbs.

We had postulated two different effects frequency might have on participants' guesses about the novel verbs. One idea was that more exemplars would facilitate correct categorization by reinforcing the correct analysis of the verb (this was prediction 3a). It turned out, however, that prediction 3b was correct: rather than facilitating correct raising responses, participants gave *fewer* raising responses (and fewer correct responses for the raising verbs) the more exemplars they saw. For example, given the *definition-only* and list condition, participants were correct for raising verbs 66.7% for 1 exemplar, 59.7% for 3 exemplars, and 57.2% for 5 exemplars. For the *frame-only* and story condition, participants were correct for raising verbs 87.8% for 1 exemplar, 84.1% for 3 exemplars, and 82.7% for 5 exemplars. Note that the same decline is evidenced in measuring the probability of raising responses (see Table 5). Thus, it appears that multiple exemplars have a detrimental effect on correct categorization of raising verbs. Why might this be?

Let us consider more closely what the set of multiple exemplars included. If a participant was in the 1-exemplar condition, they received a single sentence containing either a *raising* or a *control* verb. In the *definition-only* version this sentence contained an animate subject and an infinitive predicate, regardless of the verb type. For example:

(8) *raising*: The old man joops to be very tired.

(9) *control*: My sister zids to be at the beach.

In the *frame-only* and *frame+definition* conditions, a raising verb occurred with an inanimate subject:

(10) *raising*: The book joops to be very long.

The control verbs in this condition had the same form as in the *definition-only* condition; that is, they had an animate subject and an infinitive predicate as in (9).<sup>11</sup> In the 3-exemplar condition (frame-only or frame+definition), a raising verb occurred with one animate subject, one referential inanimate subject, and one expletive. A control verb occurred with two control frames and one transitive frame.

- (11) *raising*
- a. The old man joops to be very tired.
  - b. The book joops to be very long.
  - c. It joops to be sunny outside.
- (12) *control*
- a. My sister zids to be at the beach.
  - b. Cats zid the sunshine.
  - c. I zid to be in my warm bed on a cold day.

In the 5-exemplar condition (frame-only or frame+definition), a raising verb occurred with one animate subject, two referential inanimate subjects and two expletive subjects as in (6). Control verbs occurred 3 times with a control structure and twice in a transitive frame. These distributions were the same across the list and story conditions.

The result of this distribution of frames is that there is no multiple-exemplar condition in which *all* sentences for a raising verb contained inanimate subjects. In both the 3- and 5-exemplar conditions, raising verbs occurred with one animate subject. Only in the 1-exemplar condition was occurrence with an inanimate subject 100%, given that the single exemplar had an inanimate subject. We suspect, therefore, that encountering a verb even once with an animate subject is sufficient to weaken the raising categorization of a novel verb. This finding underscores both the strength of the predisposition to analyze animate subjects as agents, and verbs with animate subjects as agentive verbs, and the power of inanimate subjects to reverse this bias and support an alternative analysis of the verb's thematic properties, and by extension, the structure of the sentence.

As a final note about our predictions, the context of presentation (story vs. list) did not significantly affect responses. Anecdotally, some participants remarked that they found the stories distracting. We suspect that this interference negated any potential facilitative effect of the *naturalness* of learning a verb within a richer context.

Before wrapping up, let us return to the possibility that participants in this study were simply inserting a known English vocabulary item in place of the novel verb and responding by making a judgment about that known verb. While we cannot rule out this possibility with certainty, we attempted to forestall this strategy in a couple of ways. One was by constructing definitions that did not map directly onto real lexical items of the same category. For example, 'to probably be a certain way' is similar in meaning to *be likely*, but *be likely* is a raising adjective rather than a raising verb, so you cannot say *The book likelies to be on the desk*. Similarly, the verb *enjoy* used in the definition of *zid* ('to really enjoy being someplace') does not take an infinitive complement (*\*I enjoy to go to the park*). Also, 'to really enjoy/dislike being someplace' is narrower in meaning than *enjoy/dislike*, since *enjoy/dislike* are not limited to locations. Finally, not all English control verbs actually occur in a transitive frame, and although *like* and *dislike* allow this frame (*I like/dislike pizza*), the verb *try*, which is the closest English analog of *sart* ('to make a big effort to be some way'), does not. However, *sart* was used in transitive frames in the *frame* and *frame+definition* conditions (see example (4) above). If

---

<sup>11</sup> Maximum consistency would have been achieved by putting this verb in a transitive frame. However, since not all control verbs easily occur in a transitive frame, we felt it more important to preserve the *naturalness* of the verb in its sentence context and therefore chose to keep these verbs in a control-type of sentence frame. The results show such even responses across the *control* verb items, we do not feel that this choice complicates our findings.

a subject were substituting *try* for *sart*, these items should have been confusing. However, participants responded the same way on these items as on other control items.

## 6. Further Directions

We began with a question about how language learners derive the structure of sentences whose syntax is opaque on the surface. We approached this question by looking at a sentence string that is compatible with (at least) two underlying structures, where the structures are disambiguated by the choice of the main verb. If the main verb is a raising verb, the subject and main verb are not thematically related. If the main verb is a control verb, the subject and main verb *are* thematically related. Our specific question, then, became which cues in the primary linguistic data could drive learners to categorize an unknown verb in the ambiguous context as raising or control. We assume that knowledge of the structure follows from this categorization.

Our method involved a novel verb learning task with adults. Ultimately, of course, we want to answer our question by investigating learning in children. In this section we address two things: (a) we justify our choice of adult participants in this study, and (b) outline plans to study this type of verb categorization in children.

### 6.1. Why Adults?

Researchers have long used adult participants to investigate questions about language learning. In the verb learning literature this is known as the “Human Simulation Paradigm” (Gillette et al., 1999). In brief, adults are placed in a *simulated learning* environment by removing some information from the experiment stimuli (e.g., using novel words, removing all language and only showing videotaped scenes, etc.). The subject is then asked to make a judgment about the linguistic properties of a new word, or to guess which word was uttered in a conversation. The use of adult subjects instead of children has both a practical and a theoretical justification. In practical terms, adult participants are easier to recruit, they have longer attention spans, are less vulnerable to working memory or other cognitive load effects than children, and we do not need to worry about whether a particular construction or lexical item is part of their mental grammar.

The theoretical justification is twofold. One part is that adults’ intuition about verb categories and sentence structures represents the target state that children are moving towards. Although we researchers can reflect on our own internal grammars and make claims such as “raising verbs are compatible with inanimate subjects but control verbs aren’t,” it is important to verify the psycholinguistic correctness of these claims. Therefore, the fact that adults’ responses in our study confirmed our main prediction (i.e., inanimate subjects strongly favoured a *raising* categorization of the novel verb) is significant. The other part is that, as argued by Gleitman et al. (2005), adults can be presumed to have a fully intact conceptual structure. Therefore, if children and adults perform similarly in these learning tasks, explanations for children’s learning patterns cannot be based on their insufficient conceptual apparatus. In our case the idea would be that if both adults and children assume a novel verb with infinitive complement is a control verb unless it occurs with an inanimate subject, then the reason is not likely to be that raising verbs are simply hard, conceptually, to represent. Rather, the reason should be linked to a property of language, or language learning, which is independent of conceptual development.

### 6.2. How to Test With Children?

Armed with the above results from adults, we plan to investigate which cues children can use to discriminate raising from control verbs, and therefore to disambiguate opaque structures such as *John gorps to be clever*.

One consideration is how the stimuli should be presented to children. Should novel verbs be embedded in a story or simply presented in a list of unrelated sentences? It seems to us that presenting a series of unrelated sentences to children might be awkward and unnatural, especially since the sentences would be presented orally, not in written form. But if the stories were *distracting* to adults

and not facilitative, as we had expected, they seem like a poor choice as well. Therefore, we plan to construct dyadic conversations, to be videotaped, each containing a novel verb. Pilot data suggest that this format works well with children ages 5 and up. We believe it can also be made to work with 4-year-olds.

A second consideration concerns whether children will be told what the verbs mean. Because we expected definitions to be facilitative for adults, we had initially anticipated providing them to children. But given the incontrovertible hindrance of definitions, we would not use them with children. Children will be told that the experimenter does not know the meaning of these new words (i.e., ‘special puppet words’) either.

A third consideration is how to evaluate how children have categorized novel verbs. In the adult experiment, we had adults perform a forced-choice sentence grammaticality judgment task, choosing either a pseudocleft or a *there*-construction. Our pilot study with children ages 5 to 9 indicated that children could not reliably make a judgment about these constructions until at least age 6. We believe the problem for 5-year-olds (and, presumably, younger children) lies in the constructions themselves, given that expletive constructions are comparatively rare in speech to children (Hirsch & Wexler, 2007). We do not have data on the prevalence of pseudoclefts in child-directed speech, but we suspect they are likewise rare.

We plan to remedy this situation by measuring children’s grammaticality judgment, using an on-line measure (Reaction Time) in a non-metalinguistic task (answering Yes/No questions). Previous work has shown that children take longer to respond to ungrammatical prompts than grammatical ones (Naigles et al., 1995). We plan to present children with a conversation containing a novel verb (either occurring with animate subjects or inanimate subjects only), and then have children answer yes/no questions involving the novel verb in a *there*-construction. For example, *Did there gorp to be a book on the shelf?* If children heard the verb *gorp* used only with inanimate subjects, during the course of the conversation, they should be faster in responding to the question than they would be, if they had heard *gorp* used only with animate subjects. If children show no difference in their reaction time this would suggest that there is no difference in grammaticality for the children. An advantage of this approach is that it allows us to tap into children’s grammatical representation without asking them to provide a metalinguistic judgment.

A further potential complication with extending this study to children is that there are very few reliable sentential contexts that admit all and only control verbs. Even the pseudocleft admits various types of verbs (e.g., *What John ate was a sandwich; What she believed was that her friend had left*, etc.). Moreover, pseudoclefts cannot straightforwardly be transformed into yes/no questions (e.g., *\*Did what John wanted was a car?*). Transitives are a poor choice since several frequent control verbs cannot occur in that frame (e.g., *\*John hoped the answer; \*John tried the winner*). If pseudoclefts are not comprehended by young children, it will be difficult to present them with control verb test items that might yield clear results. This leaves us with only a concrete measure of whether a verb is classified as a raising verb or not. However, we believe this is not entirely problematic. Given that our adult study showed a strong bias to interpret novel verbs in the ambiguous context as control verbs, the question is not really how to evaluate children’s control categorization of novel verbs (assuming that children have the same bias as adults), but rather how to evaluate children’s raising analysis. Thus, a measure of a raising analysis is all that is needed.

A final question is the age that should be studied. Raising and control verbs are fairly abstract and are probably acquired relatively late (e.g., Becker, 2006, found reliable discrimination between raising and control verbs around age 4). If we want to know how children admit new verbs to these categories, ideally we should study children just at the beginning of this phase, perhaps right at age 4. Testing 4-year-olds would make it all the more imperative to come up with a format of presentation and test constructions that are easily comprehensible.

## References

- Becker, Misha. (2005). Learning verbs without arguments: The problem of raising verbs. *Journal of Psycholinguistic Research* 34(2), 165–191.
- Becker, Misha. (2006). There began to be a learnability puzzle. *Linguistic Inquiry* 37(3), 441–456.

- Becker, Misha. (2009). The role of NP animacy and expletives in verb learning. *Language Acquisition* 16(4), 283-296.
- Becker, Misha, & Schaeffer, Jeannette. (under review). Animacy, argument structure and unaccusatives. Unpublished manuscript.
- Bunger, Ann, & Lidz, Jeffrey. (2004). Syntactic bootstrapping and the internal structure of causative events. In Alejna Brugos, Linnea Micciulla, & Christine E. Smith (Eds.), *Proceedings of the 28th Boston University Conference on Language Development* (pp. 74-85). Somerville, MA: Cascadilla Press.
- Burzio, Luigi. (1986). *Italian Syntax: A Government-Binding Approach*. Reidel, Dordrecht.
- Carey, Susan. (1978). The child as word learner. In Morris Halle, Joan Bresnan & George Miller (Eds.), *Linguistic Theory and Psychological Reality*. Cambridge, MA: MIT Press.
- Clark, Herbert H., & Begun, Jeffrey S. (1971). The semantics of sentence subjects. *Language and Speech* 14(1), 34-46.
- Gelman, Andrew, & Hill, Jennifer. (2007). *Data analysis using regression and multilevel/hierarchical models*. Analytical methods for social research. Cambridge: Cambridge University Press. From <http://books.google.com/books?id=c9xLKzZWoz4C>
- Gentner, Dedre, & France, Ilene M. (1988). The verb mutability effect: Studies of the combinatorial semantics of nouns and verbs. In Steven L. Small, Garrison W. Cottrell, & Michael K. Tanenhaus (Eds.), *Lexical Ambiguity Resolution: Perspectives from Psycholinguistics, Neuropsychology and Artificial Intelligence* (pp. 343-382). San Mateo, CA: Morgan Kaufmann Publishers Inc..
- Gillette, Jane, Gleitman, Henry, Gleitman, Lila, & Lederer, Anne. (1999). Human simulations of vocabulary learning. *Cognition* 73, 135-176.
- Gleitman, Lila. (1990). The structural sources of verb meanings. *Language Acquisition* 1(1), 3-55.
- Gleitman, Lila, Cassidy, Kimberly, Nappa, Rebecca, Papafragou, Anna, & Trueswell, John C. (2005). Hard words. *Language Learning and Development* 1, 23-64.
- Hirsch, Christopher, & Wexler, Kenneth. (2007). The late development of raising: What children seem to think about seem. In William D. Davies, & Stanley Dubinsky (Eds.), *New Horizons in the Analysis of Control and Raising* (pp. 35-70). Dordrecht: Springer.
- Jackendoff, Ray. (1972). *Semantic Interpretation in Generative Grammar*. MIT Press, Cambridge, MA.
- Keenan, Edward L. (1976). Toward a universal definition of subjects. In Charles Li (Ed.), *Subject and Topic*. New York: Academic Press.
- Lidz, Jeffrey, Gleitman, Henry, & Gleitman, Lila. (2004). Kidz in the 'hood: Syntactic bootstrapping and the mental lexicon. In Geoffrey D. Hall, & Sandra Waxman (Eds.), *Weaving a Lexicon* (pp. 603-636). Cambridge, MA: MIT Press.
- Markman, Ellen. (1994). Constraints on word meaning in early language acquisition. In Lila Gleitman, & Barbara Landau (Eds.), *The Acquisition of the Lexicon* (pp. 199-228). Cambridge, MA: MIT Press.
- Naigles, Letitia R. (1996). The use of multiple frames in verb learning via syntactic bootstrapping. *Cognition* 58, 221-251.
- Naigles, Letitia, Gleitman, Henry, & Gleitman, Lila. (1989). Syntactic bootstrapping in verb acquisition: Evidence from comprehension. In Esther Dromi (Ed.), *Language and Cognition: A Developmental Perspective*. Norwood, NJ: Ablex.
- Naigles, Letitia, Fowler, Anne, & Helm, Atessa. (1995). Syntactic bootstrapping from start to finish with special reference to down syndrome. In Michael Tomasello, & William E. Merriman (Eds.), *Beyond Names for Things: Young Children's Acquisition of Verbs* (pp. 299-330). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Perlmutter, David. (1970). The two verbs begin. In Roderick A. Jacobs, & Peter S. Rosenbaum (Eds.), *Readings in English Transformational Grammar* (pp. 107-119). Waltham, MA: Ginn.
- Perlmutter, David. (1978). Impersonal passives and the unaccusative hypothesis. *Proceedings of the Berkeley Linguistic Society* 4 (pp. 157-189). Berkeley.
- Pinker, Steven. (1984). *Language Learnability and Language Development*. Cambridge, MA: Harvard University Press.
- Wagner, Richard K., Torgesen, Joseph K. & Rashotte, Carol A. (1999). *Comprehensive test of phonological processing*. Austin, TX: PRO-ED.