



HHS Public Access

Author manuscript

Cognition. Author manuscript; available in PMC 2016 September 01.

Published in final edited form as:

Cognition. 2015 September ; 142: 266–290. doi:10.1016/j.cognition.2015.05.015.

Using Instruments to Understand Argument Structure: Evidence for Gradient Representation

Lilia Rissman^a, Kyle Rawlins^b, and Barbara Landau^b

Lilia Rissman: lrissman@uchicago.edu; Kyle Rawlins: rawlins@jhu.edu; Barbara Landau: landau@jhu.edu

^aDepartment of Psychology, the University of Chicago, 5848 South University Ave., Chicago, IL 60637

^bDepartment of Cognitive Science, Johns Hopkins University, Krieger 237, 3400 N. Charles St. Baltimore, MD 21218

Abstract

The arguments of a verb are commonly assumed to correspond to the event participants specified by the verb. That is, *drink* has two arguments because *drink* specifies two participants: someone who drinks and something that gets drunk. This correspondence does not appear to hold, however, in the case of instrumental participants, e.g. *John drank the soda with a straw*. Verbs such as *slice* and *write* have been argued to specify an instrumental participant, even though instruments do not pattern like arguments given other criteria. In this paper, we investigated how instrumental verbs are represented, testing the hypothesis that verbs such as *slice* encode three participants in the same way that dative verbs such as *lend* encode three participants. In two experiments English-speakers reported their judgments about the number of participants specified by a verb, e.g. that *drink* specifies two participants. These judgments indicate that *slice* does not encode three distinct arguments. Nonetheless, some verbs were systematically more likely to elicit the judgment that the instrument is specified by the verb, a pattern that held across individual subjects. To account for these findings, we propose that instruments are not independent verbal arguments but are represented in a gradient away: an instrument may be a more or less salient part of the force exerted by an agent. These results inform our understanding of the relationship between argument structure and event representation, raising questions concerning the role of arguments in language processing and learning.

Keywords

Argument structure; event representation; verbal semantics; linguistic judgments; instruments

Correspondence concerning this article should be directed to Lilia Rissman, Department of Psychology, the University of Chicago, 5848 South University Ave., Chicago, IL 60637. lrissman@uchicago.edu. 773-206-3954.

¹²In Experiments 1–2, recipients were selected/produced slightly less often than themes. This absolute difference may have arisen due to task-specific effects, however: subjects may have had a mild dispreference for selecting a prepositional phrase (Experiment 1) or producing a third argument (Experiment 2).

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Argument structure is central to most theories of verbal semantics, sentence processing and children's verb learning. Despite the importance of argument structure, however, there is little consensus on how to determine when a phrase is an argument rather than a modifier of a verb (Vater, 1978; Dowty, 1982; Schutze, 1995; Croft, 2001; Dowty, 2003; Koenig, Mauener, & Bienvenue, 2003; Conklin, Koenig, & Mauener, 2004). Perhaps the most fundamental assumption is that arguments correspond to the "event participants" that are specified by the verb. Events of drinking, for example, intuitively involve two participants, the drinker and what is drunk, and *drink* has been analyzed as having two arguments, which surface as syntactically privileged NPs in the sentence. Although such cases are relatively clear, in other cases, a participant may play an essential role in an event without being considered an argument on traditional syntactic and semantic diagnostics. At the extreme, most non-stative verbs specify that an event occurs at a time and in some place, yet temporal and locative phrases are almost universally classified as modifiers (see Bresnan, 1982). A less extreme, but no less problematic case is that of instrument verbs such as *cut* or *slice*. Such verbs seem to specify that an instrument is a key event participant, but it is not at all clear that these instrumental participants should also be considered arguments. Indeed, it is well-known that different argument criteria lead to different conclusions about whether the instrument is an argument or a modifier (see Schutze, 1995; Schutze & Gibson, 1999). In this paper, we consider the general question of how one decides whether an event participant is an argument of a verb, and we do so by focusing on the controversial case of instrumental participants.

In the next sections, we begin by describing in more detail both traditional approaches to diagnosing whether instruments are arguments as well as a less traditional approach introduced by Koenig et al. (2003)/Koenig, Mauener, Bienvenue, & Conklin (2008), who propose that semantic obligatoriness is a more informative diagnostic of argumenthood than syntactic diagnostics. Each approach offers a different conclusion using somewhat different analytic methods. We build on both by proposing a new empirical approach to understanding whether instruments are arguments, drawing on native English speakers' judgments about verbally-specified participants across a wide range of verbs. Our findings suggest a resolution to the question of whether instruments are arguments, and more generally, suggest a broader conceptualization of the interface between event participants and arguments. Specifically, we propose that instruments are not independent arguments of a verb but may be more or less salient within the agentive portion of an event, as determined by the verb.

1.1 Traditional notions of argumenthood

Argument structure plays a crucial role in many theories of language use and learning. In the domain of sentence processing, argument structure has been shown to influence expectations during sentence parsing (Carlson & Tanenhaus, 1989; Clifton, Speer, & Abney, 1991; Liversedge, Pickering, Branigan, & van Gompel, 1998; Schutze & Gibson, 1999; Boland, 2005). For example, Schutze & Gibson found that participants interpreted a *for*-PP more quickly when it was an argument of a noun, as in (1a), than when it was a modifier of a verb, as in (1b):

- (1)
 - a. The company lawyers considered employee demands for a raise.
 - b. The company lawyers considered employee demands for a month.

Arguments also play an important role in theories of acquisition: to learn the meaning of a verb, children are thought to depend on the correspondence between the number of participants entailed by a verb's meaning and the number of NPs that typically appear with the verb (Landau & Gleitman, 1985; Gleitman, 1990; Naigles, 1990; Fisher, 1996; Lidz, Gleitman, & Gleitman, 2003).

Most theories of argument structure maintain a distinction between “root” and “structural” aspects of verbal meaning (Chomsky, 1965; Pinker, 1989; Jackendoff, 1990; Pesetsky, 1995; Goldberg, 1995; Rappaport-Hovav & Levin, 1998; Hale & Keyser, 2002; Borer, 2005; Ramchand, 2008). Root meaning is thought to encode the kind of fine-grained meaning that distinguishes *drink* from *gulp* from *sip*, for example. The structural component, by contrast, which is sometimes considered synonymous with argument structure, contains category, valence, aspect and thematic role information, i.e. exactly which participants are encoded as arguments. Across theories, a verb's arguments have been represented in a variety of ways, including as items in a thematic grid (Grimshaw, 1990; Croft, 1991; Goldberg, 1995; Babby, 2009), a list of θ -roles (Chomsky, 1981; Baker, 1988; Larson, 1988; Reinhart & Siloni, 2005), or as slots in a decompositional event structure (Pinker, 1989; Jackendoff, 1990; Rappaport-Hovav & Levin, 1998). Root and structural meaning are in turn distinguished from “associative” verbal knowledge, e.g. that drinking is associated with thirst, alcoholic beverages, water bottles, etc.

Given the proposed close mapping between a verb's main participants and its syntactic realization, one would expect that syntactic and semantic diagnostics should align and therefore differentiate between participants that are arguments vs. modifiers. In prototypical cases, they do. For example, in the domain of syntax, the verb constrains where an argument may appear in a syntactic structure (Tesnière, 1959; Jackendoff, 1977; Marantz, 1984; Pollard & Sag, 1987; Radford, 1988; Cowper, 1992; Schutze 1995), whereas modifiers may appear more freely and are almost always optional. In the domain of semantics, arguments “complete” the meaning of a verb, whereas modifiers contribute an additional, and relatively invariant, property of an event (Marantz, 1984; Pollard & Sag, 1987; Dowty, 2003). For example, in the sentence *Joan drank lemonade from a glass*, the NPs *Joan* and *lemonade* fill roles already specified by the verb *drink*. By contrast, *from a glass* specifies an additional property of the drinking event. Correspondingly, the syntactic position of *lemonade* is more constrained than that of *from a glass*: you can say *John drank lemonade quickly from a glass* but not **John drank quickly lemonade from a glass*.

1.2 The problematic case of instruments

Despite the alignment between syntactic and semantic diagnostics for these prototypical cases, argument diagnostics function more as guidelines than deterministic rules for identifying a verb's arguments (Vater, 1978; Croft, 2001; Dowty, 2003). The behavior of instruments demonstrates a case where multiple argument diagnostics do not align. For example, instrumental *with*-phrases have traditionally been classified as modifiers based on

their syntactic behavior (Levin & Rappaport, 1988; Sedivy & Spivey-Knowlton, 1994; Goldberg, 2002; Iwata, 2008; see also Jackendoff, 1990). Nonetheless, Schutze (1995) argues that *with*-phrases do depend on the verb for their interpretation. Koenig et al. (2003) argue that for some verbs, the instrumental *with*-phrase completes the meaning of the verb. That is, Koenig et al. propose that verbs such as *slice*, *write* and *dig* semantically “require” an instrument, whereas *eat*, *break* and *open* only “allow” an instrument. Koenig et al. developed this view by drawing on the results of a study in which they asked two informants the following question for 4200 English verbs:¹

- (2) Does one of verb’s senses describe situations in which one participant must, can, or cannot *use* another participant to perform an action?

For verbs including *slice*, *write*, *behead* and *beat*, the informants judged that one participant must use another participant to perform the action. In contrast, for verbs such as *eat*, *open*, *kill* and *break*, the informants judged that an instrument can be used but is not necessary.

This require/allow-instrument distinction was further supported by findings from a separate experiment (Koenig et al., 2003) in which participants read sentences where a *wh*-phrase could be misparsed as an object of the verb rather than an instrument:

- (3) a. Which sword did the rebels *behead* the traitor king with__ during the rebellion?
b. Which sword did the rebels *kill* the traitor king with__ during the rebellion?

Sentences with require-instrument verbs as in (a) were read more quickly than sentences with allow-instrument verbs as in (b), suggesting that require-instrument verbs activate the concept of an instrument. Koenig, Mauner & Bienvenue (2002) also found that when participants completed sentences such as *the farmer split the logs_____* and *Dana ate a bowl of soup_____*, there were more instrumental completions for require- than allow-instrument verbs.

The solution of separating the instrument verbs into two classes provides an explanation for the difficulty of deciding whether instruments are arguments or modifiers. Nonetheless, multiple puzzles remain concerning the representation of require-instrument verbs, and how require- and allow-instrument verbs differ. First, there is disagreement concerning the intuition that verbs such as *slice* and *beat* semantically require an instrument, as generated by Koenig et al.’s two informants. As described in Carlson and Tanenhaus (1989), “direct intuition is less decisive... can one slice things without an instrument? If one uses one’s hand, is that an instrument?” (10). In other words, does *slice* specify an agent, a patient and an instrument, or just an agent and a patient? Body-part instruments (e.g. slicing with a hand, writing with a finger) are possible for most require-instrument verbs, suggesting that these verbs do not, strictly speaking, semantically require an instrument.²

¹Koenig et al. (2008) report also asking coders the question “does one of the verb’s senses describe situations in which one participant must or can or cannot perform the action described by the verb *with* something?”

²Koenig et al. (2008) argue that in such cases, the body part plays the same role as a non-body part such as a knife. In this way, the body part is an instrument and does not contradict the claim that these verbs semantically require instruments. Nonetheless, it is unclear how to determine whether a body part and a non-body part are playing the same role in an event.

Second, the require/allow-instrument distinction is not reflected in syntactic argument diagnostics. Applying traditional word order and constituency tests, we find no distinction between require- and allow-instrument verbs. For example, pseudo-clefting is possible for modifier but not argument phrases, as shown in (4a–b) (Klima, 1962; Vestergaard, 1977):

- (4)
- a. What John did last night was drink a martini.
 - b. *What John did on the table was put his keys.
 - c. What Sarah did with the stick was **beat** the rugs.
 - d. What Sarah did with the chopsticks was **eat** the sushi.

As shown in (4c–d), clefting of an instrumental *with*-phrase is possible for both require-instrument verbs (e.g. *beat*) and allow-instrument verbs (e.g. *eat*), indicating that they are modifiers. Psycholinguistic studies that have investigated whether instruments show an argument-like processing signature have not contrasted the require- and allow- classes, and are thus difficult to interpret with respect to these issues (Ferretti, McRae, & Hatherell, 2001; Blodgett & Boland, 2004; Boland, 2005).

Finally, the data reported by Koenig et al. (2003) are compatible with multiple ways of representing require-instrument verbs. In their 2003 paper, Koenig et al. propose that these verbs “lexically encode” an instrument, where lexical encoding “reduces to semantic categories accessed/activated upon word recognition” (71). This description is consistent with instruments being encoded as part of root, structural or associative verbal meaning, or through some interaction of these components. This description is also compatible with theories where the argument/modifier distinction is non-categorical (Langacker, 1987; MacDonald, Pearlmutter, & Seidenberg, 1994; Manning, 2003). In later work, Koenig & Davis (2006) and Koenig et al. (2008) suggest that the instrument occupies a discrete slot in the structural component of verbal meaning.

In sum, Koenig’s approach strongly suggests that require- and allow-instrument verbs are representationally distinct. Nonetheless, additional empirical study is needed to address the issues raised above: whether/why the instrument judgments are variable across speakers, why the require/allow-instrument distinction is not reflected syntactically, and how instruments are encoded in the verb’s lexical entry (i.e. as a discrete instrumental slot or through some other mechanism). Our studies build on Koenig’s approach in order to generate a richer set of data to address these questions.

1.3 Current studies

Like Koenig et al., we elicited semantic judgments for instrument verbs. But we focused more directly on judgments concerning which participants are crucially involved in an event. We elicited such judgments by instructing people about the idea of an “argument.” For example, participants were instructed that because wanting involves someone who wants and something that is wanted, *want* therefore has two arguments. Participants then read sentences with a range of different verbs and judged whether certain participants in the sentence qualified as arguments of the verb. Crucially, we asked participants to judge whether the instrument was an argument in sentences with a variety of instrument verbs.

We tested several hypotheses. First, we asked whether the require/allow-instrument distinction uncovered by Koenig et al. holds up in a broader group of participants, with the require-instrument verbs identified by Koenig et al. showing more acceptance of instruments as arguments than the allow-instrument verbs. Second, we asked whether the judgments for require-instrument verbs yield a pattern that is similar to judgments for verbs widely analyzed as encoding three participants. If require-instrument verbs encode a discrete slot for the instrument in their argument structure, they should be represented in an equivalence class with dative verbs such as *give*, *teach* and *lend*, which are commonly analyzed as encoding an agent/source, a theme and a recipient (Dowty, 1978; Larson, 1988; Pinker, 1989). Allow-instrument verbs should not behave in the same way, and in particular, should be differentiated from both require-instrument verbs and dative verbs. Third, we asked whether the judgment data conform to the idea of two distinct and separate classes (require-instrument vs. allow-instrument), or alternatively, appear to fall into a more graded pattern, with instrument verbs having a graded probability of “requiring” an instrument.

In some subfields of language research, using metalinguistic judgments to make inferences about linguistic representation is controversial (Hill, 1961; Edelman & Christiansen, 2003; Ferreira, 2005). In this study, we focused on participant judgments as a source of evidence for two reasons: first, few theorists would deny that verbs specify participants at some level of representation.³ That is, the intuitive distinction between *sneeze*, *drink* and *give* is so robust that any theory of verbal meaning would be incomplete without an explanation for it. Given the centrality of this intuition, we aim to understand what such judgments can reveal about more controversial cases, e.g. the representation of instrumental verbs such as *slice*. The second reason we gathered judgments is that one of our central questions is whether individual require- and allow-instrument verbs cluster into distinct classes. Addressing this question requires many data points for each individual verb. Although Koenig et al. (2003) showed that a require/allow-instrument distinction was reflected in a self-paced reading task, the design constraints of such implicit psycholinguistic measures usually require averaging across individual verbs. In our judgment task, we were able to quickly elicit many judgments, allowing for a comparison between individual verbs.

Despite the seemingly robust contrast between *sneeze*, *drink* and *give*, it remains to be shown that verbal participant judgments can be elicited in an informative way from theoretically naïve participants in the lab. Recent studies suggest that most acceptability judgments previously reported by individual linguists can be replicated experimentally (Sprouse, 2007; Culbertson & Gross, 2009; Sprouse & Almeida, 2012). Nonetheless, few studies have tested the reliability of participant judgments, i.e. that giving “involves” three participants, across multiple individuals (see Wittenberg & Snedeker, 2013 for one example). To establish that the participant judgments of speakers in the lab are comparable to the judgments reported in the literature, we included in our study several types of control trials in addition to the judgments concerning instrument verbs and dative verbs. First, we included prototypical arguments and modifiers, testing how often participants judged these to be arguments. These prototypical cases allow us to evaluate how well our argument

³See Parsons (1990), Borer (2005) for possible exceptions

instruction elicited judgments parallel to what linguists have previously reported. We also included “allow-recipient” verbs: motion verbs that are compatible with but do not require a recipient, e.g. *slide* and *roll*. We compared the recipient judgments for dative verbs against the recipient judgments for these manner of motion verbs. Testing this relatively uncontroversial require/allow-recipient distinction serves as another benchmark for interpreting the judgments for the require- and allow-instrument verbs.

Although we used the term “argument” in our instructions, we do not assume that the judgments we elicited based on this instruction directly reflect the representation of argument structure itself. Other labels, such as “key participants” or “central players” would have been appropriate. For this reason, we use the term “participant judgment” to describe the type of judgment elicited in our studies. This term refers to the basic intuition that drinking, for example, involves two participants, without indicating what representation/s give rise to this judgment.

2 Experiment 1

2.1 Participants

Fifty-two adult native English speakers from the Johns Hopkins community participated (M = 13, F = 39). In a written questionnaire, five of the 52 participants indicated that they were equally confident speaking English as well as some other language. 12 participants reported having taken a college-level linguistics course. Participants received either course credit or \$12 for their participation.

2.2 Design and materials

Participants saw three types of trials: instrument trials, recipient trials and filler trials. The instrument trials included 24 instrument verbs, 12 each of which had been previously designated as either requiring or allowing an instrument by Koenig et al. (2008).⁴ The recipient trials included 16 recipient verbs, 8 of which required a recipient and 8 of which allowed a recipient; these were drawn from Levin’s (1993) classes of Verbs of Change of Possession and Verbs of Sending and Carrying. The filler trials included 20 additional verbs (see Table 1 for instrument and recipient verbs; see Appendix A for filler verbs).

Semantic subclasses figure prominently in the analysis in Koenig et al. (2008); we thus selected for the require-instrument verbs four verbs of contact (*touch*, *hit*, *beat* and *poke*) and four verbs of incision (*cut*, *stab*, *chop* and *slice*). The set of require-recipient verbs were drawn from previous studies of dative argument structure (Bock, 1986; Huttenlocher, Vasilyeva, & Shimpi, 2004; Rappaport-Hovav & Levin, 2008; Thothathiri & Snedecker, 2008). The allow-recipient verbs were manner of motion verbs consistent with a recipient. None of the recipient verbs syntactically requires a recipient *to*-phrase.⁵

⁴Koenig et al. did not explicitly classify *touch*, *lift*, *wash*, *clean* or *grow*. We designated these as either requiring or allowing an instrument based on the diagnostic question in (1). For *wash*, the instrumental role of interest was an object such as a sponge, not water.

⁵The dative verb *give* was not included in the study due to concerns that the *to*-phrase is syntactically obligatory for *give*. That is, we judged sentences such as *the woman gave something on Monday* to be somewhat infelicitous.

Trials were designed to invite participants to decide which of several entities in a sentence qualified as an “argument,” after instruction. Each trial included a single-clause sentence in which the verb appeared in capital letters and two phrases appeared in brackets, as in (5):

- (5) a. Martha **CHOPPED** something [with a bread knife] [in the operating room].
 b. Josh **SERVED** something [to the musicians] [gracefully].

The participants’ task was to choose one of the bracketed phrases as an argument of the verb, or to choose that neither phrase was an argument. In the test sentences, recipients and instruments were structurally parallel: they both appeared in a bracketed, syntactically optional prepositional phrase.

The two bracketed phrases constituted several key contrasts between two possible participant types. These contrasts are shown in Table 2 for the instrument and recipient trials:

The contrasts in the top half of Table 2 pitted instruments and recipients against prototypical modifiers (location, time and manner phrases). If subjects judge instruments and recipients to be essential participants for a verb, these should be chosen significantly more often than the prototypical modifiers. The contrasts in the bottom half of Table 2 pitted prototypical modifiers against each other. These modifier vs. modifier trials included sentences such as *Martha CHOPPED something [violently] [in the forest]*. We predicted that in the modifier vs. modifier trials, participants would judge that neither phrase was an argument of the verb.

The order of the bracketed phrases was counterbalanced such that some participants saw (5a), whereas others saw *[In the operating room] Martha CHOPPED something [with a bread knife]*. Each instrument and recipient verb appeared six times, once for each of the six contrasts in Table 2, for a total of 144 instrument trials and 96 recipient trials.

Instruments appeared in a *with*-PP; recipients appeared in a *to*-PP. Event locations were introduced by *in*. Times were introduced by *at* or *on*, e.g. *at noon*, *on Tuesday*. Manner modifiers were *-ly* adverbs such as *quickly* and *violently*. Direct objects were semantically neutral NPs such as *something* or *someone*.⁶

For each verb, there were multiple tokens of each modifier type. For example, each of the instrument verbs was paired with six unique instrument tokens, e.g. *with a bread knife*, *with a switchblade*, *with a baseball bat*. Each verb was randomly paired with participant tokens generated from the Corpus of Contemporary American English (Davies, 2008) and infelicitous pairs were excluded, e.g. *poke with a computer*; *stab with a clock*. This process yielded a distribution of typical and atypical tokens for each verb.

Experiment 1 also included 120 filler trials. In these trials, themes were pitted against participant locations (e.g. *carry [books] [in a tote bag]*) and beneficiaries (e.g. *read [the message] [for the detective]*), as well as against sources, purpose phrases and

⁶For *dig*, the direct object was always *a hole*, given the relative unnaturalness of sentences such as *Jim dug something with a shovel*.

quantificational adverbs. The function of these trials was to increase the variety in the stimuli and to assess whether participants would choose the theme as an argument. To balance the design between the filler and instrument/recipient trials, each of the filler verbs also appeared in modifier vs. modifier contrasts (e.g. *read something [from a newspaper] [for the detective]*). As with the instrument/recipient trials, each filler verb appeared 6 times. For the filler trials, there was a single pair of tokens for each verb/contrast pairing.

Summing across the instrument, recipient and filler trials, each participant in Experiment 1 viewed a total of 360 trials.

2.3 Procedure

Participants were first instructed about the nature of “arguments.” They were given a brief description to read, stating, for example, that “arguments are entities that are not part of the verb itself but are essential to its meaning.” The description stated that arguments need not appear explicitly in a sentence: “the arguments of a verb are understood to be part of the action, even when we don’t mention them... in *Mary was hugged*, we understand that Mary was hugged BY someone.” Participants were given five positive examples demonstrating that verbs can have different numbers of arguments, e.g. *run* has one, *hug* has two and *put* has three arguments. Arguments were described in terms of verb-specific roles rather than generalized thematic roles, e.g. that wanting involves someone who wants and something that is wanted. The instruction also included two negative examples, e.g. in *John ran until he was sick*, the phrase *until he was sick* is not an argument. This instruction did not include explicit information about the verbs or participant types that participants would be tested on. The experimenter read these instructions aloud while they were visible on a computer screen. See Appendix B for full instructions.

After reading these instructions, participants completed 10 interactive practice trials of three different types. In the first, participants read a sentence, were told how many arguments the verb has and were asked to indicate the arguments. For example, participants read the sentence *Chloe chased a sea turtle*, were told that *chase* has two arguments, and had to type the arguments into a box on the screen. The second type was the same as the first, except that not all arguments were present in the sentence. For example, participants read the sentence *Jim was cooking*, were told that *cook* has two arguments, and had to indicate which argument of *cook* was present and which was absent. In the last, participants were asked to list a verb’s arguments without being told how many arguments the verb has. For example, participants were asked *what is/are the arguments of “smile?”* Feedback was given after each trial, including explanations, e.g. that *steal* has three arguments because you cannot steal something without stealing it from some place/person. Participants received the same explanations regardless of their responses in the practice trials.⁷

Following the practice trials, participants viewed a display such as in Figure 1 for each trial:

⁷For *smile*, some subjects indicated that the person you smile at is an argument. After collecting data from 28 subjects, we added the explanation that this participant is not an argument of *smile* because you can smile when you are alone. The presence or absence of this additional information did not have a significant effect on subjects’ judgments.

Participants pressed the [1], [2] or [3] key on the keyboard to indicate their choice. Participants were provided a summary of the argument instructions on a printed handout. The study was presented in Psyscope.

2.4 Results

Preliminary analyses showed that, for prototypical arguments, participants' judgments were consistent with a conventional understanding of how verbs specify participants. That is, for trials in which a theme was pitted against a modifier (e.g. *Jim CARRIED [the books] [in a tote bag]*), participants chose the theme on 97% of trials. Results were then analyzed by examining the key contrasts of interest, as follows.

2.4.1 Require/Allow distinctions—When two prototypical modifiers were pitted against each other (e.g. *Martha CHOPPED something [violently] [in the forest]*), we predicted that participants would choose “neither.” They did so on 94% of recipient trials and 96% of instrument trials, again indicating that the judgments we observed are consistent with the participant judgments previously reported in the literature. Given this finding, we then proceeded to ask whether instruments (for instrument verbs) and recipients (for recipient verbs) were chosen more frequently than the prototypical modifiers, and whether this varied across the require/allow sets for the two verb classes. Figure 2 shows the mean rates of choosing the recipient or instrument, as opposed to one of the modifiers or the “neither” option, for each of the four instrument and recipient verb classes:

Consistent with our predictions for the modifier vs. modifier trials, participants chose the modifier as an argument on 1% or fewer trials. Recipients and instruments were chosen considerably more often: for the recipient verbs, the recipient was chosen on 87% (SE = .9%) of require-recipient trials and 53% (SE = 1.4%) of allow-recipient trials. For the instrument verbs, the rates of choosing the instrument were 35% (SE = 1.1%) vs. 21% (SE = .9%) for require- and allow-instrument verbs, respectively. These results suggest that the require-allow distinction holds for both classes of verbs, and that there are differences in the degree to which participants view recipient vs. instrument as likely arguments.

To assess whether these require/allow distinctions were statistically significant, we modeled the data using mixed-effects logistic regression. Although participants could choose 1 of 3 possible responses for each trial, they almost never selected one of the modifiers as an argument. Therefore, we collapsed the modifier and “neither” responses and modeled the data as a binary choice. We fit regression models using the lme4 package for R (Bates & Maechler, 2009). Possible random effects were Subject, Recipient/Instrument Token and Modifier Token. The main fixed effect of interest was Verb Type (require vs. allow).

We also factored into the model the typicality of each token with respect to the verb, e.g. how typical a bread knife is for a chopping event. Such associative knowledge is rapidly deployed in sentence processing (Taraban & McClelland, 1988; McRae, Ferretti, & Amyote, 1997; Speer & Clifton, 1998; Snedeker & Trueswell, 2004; Bicknell, Elman, Hare, McRae, & Kutas, 2010; Roland, Yun, Koenig, & Mauener, 2012) and may have influenced the argument judgments. To assess this factor, we conducted a norming study with 24 adult native speakers of English who had not participated in Experiment 1. Participants read

sentences in which the verbs from Table 1 were paired with one of their participant tokens, e.g. *Stephen CHOPPED something [with a bread knife]*. Participants judged on a 7-point scale how common they thought the bracketed phrase was for the event type indicated by the verb. Given these norming data, possible additional factors in the model were Recipient/Instrument Typicality (continuous) and Modifier Typicality (continuous) as fixed effects.

Models were evaluated through nested model comparison. Starting with an intercept-only model, if the addition of a particular factor significantly increases the likelihood of the model according to the log-likelihood ratio test, this indicates a main effect of this factor (Quené & van den Bergh, 2008).⁸

We analyzed the recipient and instrument verb data separately. The best-fitting model of recipient choice included Subject and Recipient Token random effects and Verb Type and Recipient Typicality fixed effects. Participants were more likely to select the recipient as an argument if the verb was a require-recipient verb ($\beta = 2.50$, $SE = .18$, $p < .001$) and if the recipient token was relatively typical for that verb ($\beta = -.24$, $SE = .073$, $p < .01$).⁹ The best-fitting model of instrument choice included Subject and Instrument Token random effects and the Verb Type fixed effect. Participants were more likely to select the instrument as an argument if the verb was a require-instrument verb ($\beta = 1.23$, $SE = .14$, $p < .001$). The typicality of the instrument token did not increase the likelihood of the model.

We then assessed whether these require/allow distinctions were reflected at the level of individual participants. Figure 3A shows how often each individual subject chose the recipient for allow-recipient verbs (x-axis) vs. require-recipient verbs (y-axis). Figure 3B shows the parallel relationship for the instrument trials:

Points above the line indicate participants who were more likely to choose the recipient/instrument for the require- verbs than the allow- verbs. For recipient verbs, 41 participants chose the recipient more often for require- than allow- verbs, more than expected by chance (2-choice exact binomial, $p < .001$). For instrument verbs, 42 participants chose the instrument more often for the require- verbs, more than expected by chance (2-choice exact binomial, $p < .001$). These data show that the require/allow distinction is observed on an individual basis for most participants, for both recipient and instrument verbs.

The 12 participants who had taken a college-level linguistics course did not show any distinct trend in their judgments: descriptively, the values for these participants are scattered evenly throughout the individual subject data points in Figures 3A-B. Adding this binary factor (course vs. no course) did not increase the likelihood of either the recipient or the instrument model, and there were no interactions with Verb Type (require vs. allow).

2.4.2 Individual verb analyses—One of our main questions was whether the division of instrument verbs into two discrete classes -- require and allow -- is the best way of understanding the data across individual verbs. In order to examine this, we analyzed

⁸Given two models that differ by one independent variable, the difference between the log-likelihoods of these models multiplied by -2 is distributed as a chi-squared random variable with $df = 1$.

⁹The coefficient for typicality is negative because on the 7-point scale, 1 = very typical and 7 = not at all typical

individual verb differences in each of the recipient and instrument verb classes. Figure 4 shows the proportion of trials where the recipient was chosen for each verb:

Figure 5 shows the rates of choosing the instrument as an argument for individual instrument verbs:

To assess whether there was significant variability within each verb class, we analyzed the data for each class separately and tested whether the factor Verb (*eat* vs. *drink* vs....) contributed significantly to the model. We found significant variability within all four verb classes (require-recipient: $\chi^2(7) = 19.2, p < .01$; allow-recipient: $\chi^2(7) = 43.5, p < .001$; require-instrument: $\chi^2(11) = 49.8, p < .001$; allow-instrument: $\chi^2(11) = 56.5, p < .001$). The individual verb data also indicate a possible influence of semantic features on the argument judgments. Among the require-instrument verbs, for example, the four incision verbs (*slice*, *cut*, *chop* and *stab*) clustered together and were all significantly higher than the grand mean for the require-instrument verbs (*slice*: $\beta = .83, SE = .21, p < .001$; *cut*: $\beta = .64, SE = .21, p < .01$; *chop*: $\beta = .59, SE = .21, p < .01$; *stab*: $\beta = .45, SE = .21, p < .05$).

2.5 Discussion

In many ways, the results of Experiment 1 confirm a conventional understanding of how verbs specify participants: subjects categorically selected the theme as the argument and categorically avoided choosing manner, time and event location phrases. Participants also chose the recipient and 87% of dative verb trials, compatible with the theory that these verbs encode three participants. Subjects distinguished require-recipient verbs from allow-recipient verbs by a wide margin. Finally, participants were significantly more likely to choose the instrument as the argument for require- than allow-instrument verbs, supporting the evidence reported by Koenig et al. (2002, 2003).

Although subjects' judgments conformed to our predictions in multiple respects, questions remain about whether specific aspects of our procedure and design influenced the judgments. That is, to what extent are the main findings from Experiment 1 robust to changes in our method for eliciting judgments? In Experiment 1, subjects viewed verbs and participants in the context of a sentence, e.g. *Martha CHOPPED something [with a bread knife] [in the operating room]*, and judged which participant was an argument of the verb. This method for gathering judgments may have had several task-specific effects, effects which are not reflective of verbal meaning itself.

First, we found that participants in Experiment 1 almost always chose the theme as the argument for the filler trials, as we predicted. Although we assume that this finding reflects intuitions about the semantic importance of the theme to the verb's meaning, it is possible that participants categorically selected the theme for more surface-level reasons, e.g. because the theme always appeared next to the verb. In the argument instruction, participants were given many examples of arguments in direct object position; indeed, every NP in direct object position was an argument.

Second, subjects in Experiment 1 viewed examples of the participants as part of the stimuli, e.g. *[to the woman]*, *[with a bread knife]*. Seeing examples of recipients and instruments

presumably communicated to subjects that these were potential participants of interest, which may have introduced a bias to choose recipients and instruments as arguments. In other words, subjects may have been more likely to choose a participant as a result of seeing it in a sentence. Such a bias could undermine the important finding that participants were highly likely to choose a recipient for a dative verb.

Third, in Experiment 1 we did not fully control for the typicality of the items. Although we conducted a norming study to gauge how typical each item was for each verb, e.g. how typical it is to chop with a bread knife, we did not estimate the typicality of the two participants in conjunction, e.g. how typical it is to chop with a bread knife in an operating room. It is therefore unknown whether and to what extent the conjoined typicality might have influenced the judgments.

To address these concerns, Experiment 2 used a method for eliciting participant judgments in which the verbs were not viewed in the context of a sentence, with particular NPs and a particular syntax. If the results mirror those in Experiment 1, the patterns in the data are likely not due to specific features of the particular participants/NPs viewed in the test sentences.

3 Experiment 2

We asked subjects to judge the arguments of a verb without viewing the participants of interest in a sentence. That is, subjects were simply asked “what is/are the arguments of ‘chop’?” Because these test sentences do not include particular NPs, this design controls for the three sentence-specific influences described above.

3.1 Participants

Twenty-two adult native speakers of English participated ($M = 10$, $F = 12$). Eleven of these participants were from the Johns Hopkins community, and another 11 were from the University of Chicago community. One participant reported being bilingual. Three reported having taken a college-level linguistics course. None had participated in Experiment 1. Three additional participants were tested but were excluded from analysis: one was excluded due to experimenter error, and another two failed to describe the arguments in terms of verb-specific roles. One participant, for example, just listed wh-words such as *who* and *what*, which are difficult to interpret with respect to their participant type.

3.2 Procedure, Design & Materials

Participants received the same argument instruction as in Experiment 1. For each of the 60 verbs in Experiment 1, participants were then asked *What is/are the arguments of “VERB”?* Participants typed their response into a box on the computer screen, and were given unlimited time. Participants judged each verb once, and viewed the verbs in a random order. Each participant saw a unique random order.

3.3 Coding

We coded subjects’ responses as to whether they included four different types of participants: themes, recipients, participant locations and instruments. Themes included

either an object being acted upon, e.g. *who/what is being killed; the things being blended*, or an object moving or being transferred, e.g. *thing being lent; person/thing rolling*. Subjects marked recipients in multiple ways, e.g. *person told; audience offered to; the person that they are serving that thing to*. A participant needed to be marked as animate to be coded as a recipient, in contrast with participant locations, e.g. *place it was thrown to; the place it is sent*. Instruments were marked with *with* or *use*, e.g. *the thing used to cut; the thing being drawn with*, or were marked as tools, e.g. *the chopping tool*. A second coder coded a random subset of the responses; agreement across coders was 98%.

3.4 Results

Subjects never listed a time or manner as an argument. There was one example of an event location, which specifies where the entire event takes place: one subject listed the arguments of *paint* as “the painter, what is being painted, where the painting job takes place.” These results are consistent with the results from Experiment 1, where subjects almost never chose times, manners or event locations as arguments. Table 3 shows how often subjects in Experiment 2 produced themes, recipients, participant locations and instruments for the relevant verb classes (standard error of the mean is shown in parentheses):

Comparable to the results from Experiment 1, subjects almost always listed a theme as one of the arguments of the verb. We analyzed the recipient verb and instrument verb data separately. In a logistic regression model predicting whether subjects would list a recipient as an argument for a recipient verb, adding the factor Verb Type (require vs. allow) significantly increased the likelihood of the model: $\chi^2(1) = 298.1, p < .001$. In a model predicting whether subjects would list an instrument as an argument for an instrument verb, adding the factor Verb Type (require vs. allow) also significantly increased the likelihood of the model: $\chi^2(1) = 25.1, p < .001$.

Across Experiments 1 and 2, the means for the individual verbs, i.e. the rate of choosing an instrument for *slice* vs. the rate of producing an instrument for *slice*, were highly correlated (recipient verbs: $r(14) = .90, p < .001$; instrument verbs: $r(22) = .80, p < .001$). In Experiment 1, we observed significant variation across the require-recipient verbs. This pattern also emerged in Experiment 2, with rates of producing a recipient ranging from 95% (*teach, tell, offer*) to 77% (*serve*). The means for individual allow-recipient verbs in Experiment 2 were low compared to those of the require-recipient verbs: across the eight allow-recipient verbs, only *toss* and *throw* elicited recipient responses (14% and 5% of trials, respectively). Combining recipients and participant locations, rates of producing these ranged from 5% (*bounce*) to 45% (*slide*).

3.5 Discussion

The results of Experiment 2 parallel those of Experiment 1 in multiple ways: subjects were highly likely to list a recipient as an argument of a dative verb, and were much more likely to list a recipient for a require-recipient verb than for an allow-recipient verb. Subjects were also more likely to list an instrument for a require-instrument verb than for an allow-instrument verb, and there was a substantial difference between the require-recipient verbs and the require-instrument verbs. In addition, the rates of selecting a recipient/instrument as

an argument and the rates of producing a recipient/instrument as an argument were highly correlated across individual verbs. Experiment 2 shows that these results do not depend on subjects viewing the participants of interest in the context of a sentence.

The biggest difference between the two experiments concerns the allow-recipient trials. In Experiment 1, subjects judged an animate recipient to be an argument on 54% of trials. In Experiment 2, by contrast, subjects listed an animate recipient on only 2% of trials. The data on participant locations in Table 3 suggest a partial explanation for this contrast: for the allow-recipient verbs, a participant location was listed relatively frequently, e.g. *the place where something is thrown to*. In Experiment 1, subjects judged sentences such as *Martha THREW something [to the woman] [on Monday]*. While the NP *the woman* fills the role of an animate recipient, it also fills the role of the location where something is thrown. The relative salience of this location role may have led subjects to judge the recipient as an argument of allow-recipient verbs such as *throw*. That is, participants were not instructed whether they were judging the argument status of *the woman* qua animate recipient vs. spatial goal.

Another difference between the two experiments concerns the instrument data: participants in Experiment 2 were less likely overall to list an instrument as an argument than subjects in Experiment 1 were to choose the instrument as an argument. This difference may have been a result of the main methodological difference between the two experiments: if subjects in Experiment 1 were uncertain about the argument status of an instrument, seeing the instrument in a sentence may have encouraged them to choose it as an argument. Crucially, however, there was little difference between the require-recipient data for the two experiments: a recipient was chosen/produced on 87%/85% of dative trials. For participants that are strong examples of arguments, seeing or not seeing a recipient in a sentence does not appear to have an effect. Returning to the motivating questions for these studies, we compared dative verbs with require-instrument verbs to assess whether the latter are represented as specifying three participants. The data from Experiments 1 and 2 strongly suggest that this is not the case. In Experiment 1, participants were much more likely to choose a recipient for a require-recipient verb than an instrument for a require-instrument verb (87% vs. 35%). The gap was even wider in Experiment 2: 85% vs. 13%. The individual subject data from Experiment 1 also indicate a qualitative difference between the two verb classes. In Figure 3A, which shows judgments on recipients, 80% of the points lie above the .8 line on the y-axis. That is, most participants agreed that the recipient is an argument for require-recipient verbs, although judgments varied widely for allow-recipient verbs. By contrast, Figure 3B shows that for the instrument data, few participants judged that the instrument is an argument of a require-instrument verb on 80% or more trials. Thus the relatively low mean of 35% in Experiment 1 does not mask a distribution in which a subset of participants treated dative verbs and require-instrument verbs equivalently. Taken together, these data are inconsistent with the hypothesis that dative verbs and require-instrument verbs are in an equivalence class.

We also sought to determine in these experiments whether instrument verbs fall into two discrete categories, require- and allow- as proposed by Koenig et al. Experiments 1 and 2 both show that instruments were judged to be better arguments for require-instrument verbs

than for allow-instrument verbs, but the data do not indicate a clear boundary between these verb classes. In Experiment 1, there was significant variability within each of the four verb classes. Even in the free-response task in Experiment 2, where instruments were disfavored overall, subjects occasionally listed an instrument for an allow-instrument verb. We also found evidence that the variability within verb classes is partially explained by differences in the root meanings of the verbs. Among the require-instrument verbs, the four verbs encoding incision (*slice*, *chop*, *cut* and *stab*) clustered together as most likely to have the instrument be judged as an argument. Verbal root meaning may also have influenced the judgments for the allow-recipient verbs: *toss*, *take* and *throw* all necessarily involve spatial dislocation of an object, whereas *push*, *bounce* and *kick* do not. We return to the role of verbal root meaning in the general discussion.

We found in Experiments 1–2 that participants were highly likely to choose and produce a recipient for the require-recipient verbs. One unexpected finding, however, was that there was variation across the require-recipient verbs, in both experiments. This variation does not appear to be random: splitting the eight require-recipient verbs into upper and lower halves across the two experiments, we find that *lend*, *tell* and *teach* are in the upper half in both experiments, and *sell*, *pay* and *serve* are in the lower half. This variation calls into question the analysis that the require-recipient verbs all categorically encode three argument slots. We return to this issue in the general discussion. For now, we interpret the require-recipient data as providing a benchmark of how verbs that strongly encode three participants pattern on our task. Experiments 1–2 show that require-instrument verbs do not meet this standard.

Thus the findings so far indicate that require- and allow-instrument verbs are represented differently with respect to instruments, but that this difference cannot be characterized as one in which require-instrument verbs specify three participants and allow-instrument verbs specify two. Instead, we propose that require-instrument verbs specify two participants, the agent who acts, and the patient being acted upon, but that the instrument is represented via a gradient mechanism. Specifically, the root meanings of many of the instrument verbs indicate that the agent acts in a particular way, e.g. for *slice* the agent wields an object with a bladed shape. Given differences in root meaning across the instrument verbs, the instrument is represented as a more or less salient part of the force exerted by the agent, explaining the gradience in judgments across the instrument verbs. We present the details of this proposal in the general discussion.

Before turning to this discussion, however, we present several additional studies that test alternate interpretations of Experiments 1–2, as well as the validity of our participant judgment elicitation method more generally. In Experiment 3, we ask whether the gap between the recipient and instrument judgments in Experiment 1 is driven by a salient difference between the stimuli for the two verb classes, namely animacy. If people are more likely to choose animates as arguments, this in itself could lead to more frequent choices of recipients than instruments. In Experiment 4, we explore the relationship between associative knowledge and the participant judgments: we ask to what extent the variability in the judgments across verbs in Experiments 1–2 was driven by subjects' knowledge of how often instruments and recipients participate in real-world events involving lending, sending, slicing, and the like. In Experiment 5, we ask whether the results observed in Experiment 1

are specific to verbal participant judgments, or whether they could have been generated by other ways of reflecting on the stimuli sentences. Finally, we assessed in a corpus study the relationship between participant judgments and how often instruments and recipients are mentioned in spoken speech.

4 Experiment 3

Overall, participants in Experiment 1 selected recipients more often than instruments as arguments. This finding may reflect not a difference in argument status between require-recipient verbs and require-instrument verbs but rather a sensitivity to animacy: the *to*-recipients in Experiment 1 were always human, whereas the *with*-instruments were always physical objects. To test the influence of animacy, we compared participant judgments for the allow-recipient verbs with animate recipients vs. spatial goals, e.g. *to the parking meter*. For the purpose of this experiment, we refer to spatial goals as “inanimate recipients.” Require-recipient verbs for their part are largely incompatible with inanimate recipients: **John paid \$.50 to the parking meter*.

4.1 Participants

Twenty-four adult native speakers of English from the Johns Hopkins community participated in Experiment 3 (M = 9, F = 15). One of these participants reported being bilingual. One reported having taken a college-level linguistics course. None had participated in the previous experiments.

4.2 Design & Materials

The require-recipient trials were removed from Experiment 1 and animacy was introduced as a between-participants variable for the allow-recipient trials. Half of the participants saw the same allow-recipient trials as in Experiment 1, which had animate recipients. The other half saw allow-recipient verbs with inanimate recipients in (6):

- (6) a. Rachel BOUNCED something [to the end of the street] [loudly].
 b. [In the park] Paco KICKED something [to the fence].

Each allow-recipient verb was paired with 6 unique inanimate recipient tokens which were chosen pseudo-randomly, as described above for Experiment 1. Experiment 3 included the same instrument and filler trials as in Experiment 1. With the require-recipient trials removed, each subject viewed 312 trials.

4.3 Procedure

The procedure was the same as in Experiments 1 and 2.

4.4 Results and discussion

The proportions of trials where participants chose the recipient or instrument as an argument are shown in Table 3, separated by Animacy (animate recipient vs. inanimate recipient):

Table 3

Argument choice means for recipients and instruments in Experiment 3

Verb type: Choice	Recipient type	
	Animate	Inanimate
Allow-recipient verb: recipient	.37 (.03)	.34 (.03)
Require-instrument verb: instrument	.35 (.02)	.35 (.02)
Allow-instrument verb: instrument	.25 (.02)	.21 (.02)

The regression model of these data shows no effect of Animacy: ($\chi^2(1) < .0001$, $p > .5$). Participants were no more likely to judge that the recipient was an argument when it was animate than when it was inanimate. Collapsing across animacy, the argument choice means for individual allow-recipient verbs ranged from 18% (*kick*) to 44% (*toss*) and correlated highly with the allow-recipient means from Experiment 1, Pearson's $r(6) = .81$, $p < .05$.

The data from the instrument trials replicate the findings from Experiments 1 and 2: participants were more likely to select the instrument for require- than allow-instrument verbs ($\chi^2(1) = 38.0$, $p < .001$). There was no effect of recipient animacy on the instrument judgments ($\chi^2(1) = .13$, $p > .1$). The argument choice means for individual instrument verbs were also highly correlated across experiments, Pearson's $r(22) = .87$, $p < .001$.

In sum, animacy did not have an effect on participants' judgments concerning the argument status of the recipient or the instrument, suggesting participants' propensity in Experiments 1–2 to choose/produce the recipient more often than the instrument is not an effect of the recipient being animate.¹⁰

5 Experiment 4

In our discussion of the results of Experiments 1–2 in Section 3.5, we suggested that the variation across the instrument verbs is due to differences in the root meanings of these verbs, specifically how these verbs constrain the action performed by the agent. In Experiment 4, we test an alternate explanation: that cross-verb variability reflects differences in how often instruments and recipients occur as part of events in the real world. Events of slicing, for example, more often involve an instrument than events of breaking; it may be this difference, rather than differences in the verbal root, which drives the variation across verbs. We assume that this type of associative knowledge is distinct from knowledge about verbally-specified participants. In Experiments 1 and 2, for example, we instructed subjects that although we often smile AT someone, this participant is not an argument of *smile* because you can sit alone in an empty room and smile. Nonetheless, these two types of information are correlated, and may have influenced the participant judgments. In estimates of real-world participant frequency, we expect to find require/allow distinctions: by

¹⁰In Experiment 1, the recipient was selected on 54% of allow-recipient trials. By contrast, the recipient was selected on only 37% of the animate allow-recipient trials in Experiment 3. This difference raises the possibility that the presence of the require-recipient verbs in Experiment 1 caused subjects to be more likely to judge the recipient as an argument for allow-recipient trials.

hypothesis, what it means for a verb to require a participant is that those events always involve that participant.

5.1 Participants

Twenty participants on Amazon Mechanical Turk participated in Experiment 4 ($M = 8$, $F = 12$). All participants were living in the United States and reported being native speakers of English. Participants received \$.50 for their participation.

5.2 Design & Materials

Participants were asked to provide estimates about the real-world occurrence of instruments and recipients for the verbs in Experiments 1–2. The estimate question for instrument verbs is shown in (7a), illustrated with *chop*. The question for recipient verbs is shown in (7b), illustrated with *bounce*:

- (7) a. When people perform actions of CHOPPING, what percentage of the time do people use an object or tool that is separate from their body?
- b. When people perform actions of BOUNCING, what percentage of the time do people intend to bounce something to someone?

Participants answered 24 instrument questions and 16 recipient questions. These questions were intermixed and randomized. As discussed in Section 1.2, one of the difficulties in determining whether verbs require instruments is the issue of whether body parts count as instruments. To avoid this difficulty, participants in Experiment 4 were instructed to estimate how often events involve tools outside the body.

5.3 Procedure

Participants were instructed that they would answer questions about people's physical actions in the world. They were provided an example question with the verb *smile*: "when someone smiles, how often does someone see them smile?" They were instructed that smiling people are often observed, but smiling people are also frequently alone, and they should assign a percentage value to how often they think smiling people are observed. They were then told that they would make estimates about how often events in the world involve tools and recipients. For each trial, participants typed a number between 0 and 100 in a box on the computer screen.

5.4 Results and discussion

As expected, participants judged that require-recipient events are more likely to involve a recipient in the real world than allow-recipient events: 87 ($SD = 19$) vs. 50 ($SD = 25$); Mann-Whitney $U = 22552$, $p < .001$, two-tailed. This require/allow pattern also emerged for the instrument verbs: 76 ($SD = 26$) vs. 64 ($SD = 29$); Mann-Whitney $U = 13393$, $p < .001$, two-tailed. Require-recipient verbs had higher estimates than require-instrument verbs (Mann-Whitney $U = 36116$, $p < .001$, two-tailed).

The critical question is whether real-world estimations explain the cross-verb variability in argument judgments that we found in Experiment 1. In Figure 6, we plotted the values from

Experiments 1 and 4 for each verb and verb class. Real-world estimates are plotted on the x-axis and range from 0–100; argument choice means are plotted on the y-axis and range from 0–1:

For the require-recipient verbs, there is no correlation between real world estimation and argument judgment: $r(6) = .04, p > .5$. For the allow-recipient verbs, there is a strong correlation between the two measures: $r(6) = .79, p < .05$. For the require-instrument verbs, there is a moderate correlation between real-world estimation and argument judgment: $r(10) = .62, p < .05$, but this correlation may be driven by the verb *touch*, which has low ratings on both measures. When *touch* is removed, the correlation is relatively weak: $r(9) = .42, p > .1$. For the allow-instrument verbs, there is a weak and non-significant correlation between real world estimation and argument judgment: $r(10) = .22, p > .1$.

These findings suggest that the instrument judgments in Experiment 1 are not readily explained through associative knowledge about how often tools are used when these events are performed. First, although participants in Experiment 4 judged that require-recipient verbs involve a recipient more often than require-instrument verbs involve a tool outside the body (87 vs. 76), we consider it unlikely that this difference could account for the wider difference in the argument means for the two verb classes (.87 vs. .35 in Experiment 1, .85 vs. .13 in Experiment 2). For many of the require-instrument verbs, the estimation rates were as high or higher than those of the require-recipient verbs, a pattern that diverges sharply from the participant judgments.

Second, people's real-world estimates do not account substantially for the variability within the instrument verb classes. Although the incision verbs *slice*, *cut*, *stab* and *chop* were distinguished from other require-instrument verbs with respect to their argument status, this clustering was not reflected in the real-world estimations: *stir*, *draw* and *write* were judged equally likely to involve an external tool as the incision verbs. We also did not find a significant correlation among the allow-instrument verbs. For example, participants estimated that events of eating and drinking are relatively likely to involve a tool, yet these verbs' argument judgment means were the lowest among all instrument verbs. Additionally, subjects judged that events of cleaning and painting are as likely to involve a tool as events of cutting and stabbing.

6 Experiment 5

Experiment 4 addressed whether associative knowledge about how often participants are involved in events in the real world influenced the judgments in Experiment 1. Experiment 5 addresses a converse question: whether participants in Experiment 1 needed to draw on their knowledge of how verbs specify participants to generate the contrasts we observed. That is, are the properties of the stimuli such that other types of judgment tasks would yield the same results? We evaluated in Experiment 5 whether judgments of “importance” would yield the same results as judgments about “arguments.” Rather than give participants the argument instructions, we asked them to consider the meaning of each of the bracketed phrases in Experiment 1 in a sentence and determine which of those phrases was “most important” in the sentence.

6.1 Participants

Twelve adult native English speakers of English from the Johns Hopkins community participated (M = 6, F = 6). One of these participants reported being bilingual. One reported having taken a college-level linguistics course. None had participated in the previous experiments.

6.2 Materials

Participants in Experiment 5 viewed the same sentences as in Experiment 1.

6.3 Procedure

Participants were told that they would read a variety of sentences on a computer screen and would have to choose which bracketed phrase was more important. Participants were not directed towards a particular interpretation of “important.” As in Experiment 1, they could choose either the first or second bracketed phrase as more important. The third, alternate choice was to choose that the bracketed phrases were “equally important.” Given that “important” is a scalar term and “argument” is a non-scalar term, we could not provide an exact parallel to the “neither” option in Experiment 1. Nonetheless, the “equally important” and “neither” choices both give participants the option of not distinguishing between the bracketed phrases.

6.4 Results and discussion

Figure 7 shows the mean rates at which participants chose the theme, instrument or recipient vs. one of the modifiers or the “equally important” option in each of the verb classes:

The pattern of judgments in this experiment was quite different from the one in Experiment 1. Whereas participants in Experiment 1 judged that the theme was an argument on 97% of trials, the theme was selected as more important on only 42% of trials. By contrast, modifiers were more often judged as important than as arguments: 15% of trials in Experiment 5 vs. 1% of trials in Experiment 1. Experiment 5 did not replicate the require/allow distinctions from Experiment 1: there was no effect of Verb Type (require vs. allow) for either recipient trials ($\chi^2(1) = .31, p > .1$) or instrument trials ($\chi^2(1) = 1.2, p > .1$). There was also no effect of Verb for either group of verbs (recipient verbs: $\chi^2(15) = 15.3, p > .1$; instrument verbs: $\chi^2(23) = 15.4, p > .1$).

The results suggest that our findings in Experiment 1 would not have arisen from just any way of reflecting on the sentences. First, the importance judgments did not yield a categorical distinction between prototypical arguments and modifiers. Second, importance judgments failed to reflect a require/allow distinction, for either recipient or instrument verbs. The results from Experiment 5 are also informative in that instrumental phrases were selected as more important more often than they were selected as arguments (51% vs. 28%). This finding suggests that the relatively low rate of choosing the instrument as an argument was not due to a general dispreference for selecting the *with*-phrases. Experiment 5 does not prove that no other way of judging the experimental sentences could lead to the results from Experiment 1. It does, however, indicate that the main effects in Experiment 1 do not come for free given general properties of the stimuli sentences.

7 Corpus study (Experiment 6)

In Experiments 3–5, we showed that the difference between the recipient and instrument judgments is likely not an effect of animacy, that knowledge of how often tools are involved in the real world is a weak predictor of variability across individual verb judgments, and that evaluating the sentences in terms of the importance of the bracketed phrases does not generate the patterns observed in Experiment 1. In a final study, we ask about the relationship between the judgments from Experiment 1 and the frequency with which instruments and recipients co-occur with verbs in spoken language. Participant judgments and corpus statistics both serve as windows into verbal representation; nonetheless, it is unclear from previous research whether we should predict a positive correlation between the likelihood that a participant is judged to be an argument and how often this participant is mentioned in spoken speech. We addressed this issue in Experiment 6.

Previous studies have observed relationships between verbal meaning and how often arguments are mentioned in linguistic corpora (Resnik, 1996; Hare, McRae & Elman, 2003). For example, Resnik (1996) found that for verbs that place more specific selectional restrictions on their direct objects, these objects are omitted more often in language use. In the present study, it may be that verbs that strongly encode a participant, e.g. *teach* → recipient, are also more likely to elicit that participant in spoken speech. On the other hand, the opposite relationship may hold: if verbs such as *slice* and *chop* strongly encode an instrument (relative to other instrumental verbs), then instruments may actually be produced less often in communication, because they are implicit in the meaning of the verb.

We may also fail to find a correlation between participant judgments and corpus frequencies. Koenig et al. (2003), for example, conducted a study in which subjects completed sentence fragments containing require- and allow-participant location verbs (e.g. *the collectible doll was advertised* vs. *the collectible doll was sold*). Koenig et al. did not find a correlation between rates of producing a participant location for each verb in their task and frequencies of participant locations in the Brown corpus. This result suggests that verbal meaning may influence language behavior in different ways in different contexts, and that not all of these behaviors will be correlated with each other. To test these alternatives, we coded a sample of spoken verbal utterances from the Corpus of Contemporary American English (COCA) (Davies, 2008), measuring how often instruments and recipients are produced for each of the verbs in Experiment 1.

7.1 Utterance coding

The spoken language subcorpus in COCA contains unscripted conversation from TV and radio programs from 1990-present. To avoid formal and register-specific language, we analyzed only spoken language utterances. Each word in COCA is tagged for its part of speech. For each of the instrument and recipient verbs in Experiment 1, we selected utterances containing this verb at random, tagged with a verbal POS tag. These utterances were then filtered to exclude abstract, metaphorical, idiomatic and unrelated senses of the verbs, as in (8):

- (8) a. The company announced plans to close plants and **cut** 25,000 jobs.

- b. The mandatory evacuations **kicked** in at 6:00 a.m. this morning.

We excluded metaphorical uses to better match the sentences in Experiment 1, which all denoted physical actions involving concrete objects and people. In many cases, the abstract senses of a verb are not semantically compatible with an instrument or recipient, as in (8a–b). Guidelines for distinguishing concrete from abstract senses were determined on a verb-by-verb basis. For example, uses of *move* needed to convey a trajectory of a physical object through space, excluding *I was moved to vote yes on the ballot measure*. At least 100 concrete uses were obtained for each verb.

7.2 Participant coding

The concrete utterances were then coded as to whether they included an instrument, recipient, event location, temporal modifier or manner adverb. Multiple syntactic/morphological variants for each thematic role type were included. For example, the underlined phrases in (9) and (10) were coded as instruments and recipients, respectively:

- (9) a. I hit the baseball with a baseball bat.
 b. He used the wipes to clean his cell and wash off fresh fruit.
 c. We just got shovels and dug out three areas.
- (10) a. They're going to sell oil cheap to Germany.
 b. You're not going to serve her because of what she's wearing?

Including multiple surface realizations of the same role allows for a more accurate estimate of how often this role is expressed for each verb. Only constituents in the same sentence as a verb were coded as filling a particular thematic role.

7.3 Training and reliability

Three coders and the first author determined the verb-specific criteria for distinguishing abstract from concrete uses. During training, the first author and the three coders established a 90% reliability rate with each other. Subsequently, each verb was coded by a single coder. The first author coded a random sample of the utterances: there was 93% agreement for the recipient verb utterances and 94% agreement for the instrument verb utterances. The coders discussed disagreements with the first author, and the first author made a final decision.

7.4 Results and discussion

Spoken utterances with a require-recipient verb were more likely to contain a recipient than utterances with an allow-recipient verb ($\beta = .36$, $SE = .13$, $p < .05$). This require/allow contrast also emerged for instrument verbs ($\beta = .047$, $SE = .016$, $p < .01$). The relationship between the argument judgments and the corpus data is illustrated in Figures 8A–B. These figures show the rates of producing instruments/recipients, event locations, times and manner phrases in the corpus (x-axis) vs. the rates of judging these participant types as an argument in Experiment 1 (y-axis), for each of the instrument and recipient verbs. The judgment means for the event locations, times and manners are drawn from the modifier vs. modifier trials in Experiment 1.

The rate of choosing the recipient as an argument in Experiment 1 correlated with the frequency of recipient/verb co-occurrences in the corpus: $r(14) = .72, p < .01$. We also found a significant correlation between instrument judgments and instrument/verb frequency: $r(22) = .60, p < .01$. For event locations appearing with instrument verbs, argument judgments and corpus frequencies were correlated: $r(22) = .53, p < .01$. This result appears to be driven by the values for *dig*, however: when *dig* was removed, the correlation for event locations dropped to $.19, p > .1$. None of the other participant types yielded significant correlations (instrument verbs: times: $r(22) = .06, p > .1$; manners: $r(22) = .37, p = .08$; recipient verbs: event locations: $r(14) = .06, p > .1$; times: $r(14) = .37, p > .1$; manners: $r(14) = .07, p > .1$).

The significant correlations we observed between participant judgments and corpus frequencies for recipients and instruments suggest that the more salient a participant is for a verb, the greater likelihood that this participant will be mentioned in conjunction with that verb in spoken language. Recall that Koenig et al. (2003) did not find a correlation between corpus frequencies and rates of sentence completion. A possible explanation for this difference is that our corpus study only included concrete verbal uses. Non-concrete uses (e.g. cutting taxes, breaking vows) are not generally compatible with an instrument. Crucially, verbs differed widely as to how often they conveyed a concrete meaning: most uses of *stab*, for example, were concrete, but less than 10% of *cut* uses were concrete. As concrete utterances are more likely to contain instruments in general, analyzing all the verbal uses in a corpus could obscure a correlation.

The observed correlations between participant judgments and corpus frequencies raise the possibility that the participant judgments were generated directly from the corpus, rather than both being modulated by an underlying verbal representation. Frequency statistics have been found to play a major explanatory role in language processing (see Jurafsky, 2003; Diessel, 2007 for review). Given the strength of these effects, McDonald, Pearlmutter & Seidenberg (1994) argue that the argument/adjunct distinction reflects how frequently a participant follows a lexical head: “an argument is a PP that is strongly (frequently) linked to a word... and an adjunct is one that is weakly (infrequently) linked” (694). In other words, subjects performing the judgment task may be probability matching the frequency of verb-participant cooccurrences in their linguistic environment (see Anderson, 2000).

The data from the corpus study are not consistent with a simple version of this matching strategy: across the instrument verbs, the speakers in the corpus were as likely to produce an event location, time or manner phrase as to produce an instrument phrase. Nonetheless, participants in Experiment 1 almost never judged that event locations, times and manners were arguments. If the judgments in Experiment 1 were simply matched against spoken language probabilities, the judgments for instruments and modifiers should have been more parallel. Subjects' responses in the judgment tasks may have been driven by statistics in a more complex way, however. As described in Koenig et al. (2003), event locations, times and manners are poor candidates for being arguments because they are not lexically specific: almost all non-stative verbs are compatible with an event location, time or manner modifier. Subjects may have ruled out non-specific participants as possible arguments, but then have based their judgments concerning recipients and instruments, which are lexically specific, on statistics regarding how often these participants are mentioned in spoken language.

We favor the explanation that subjects' knowledge of how verbs specify participants influences both participant judgments and corpus frequencies, yielding the observed correlations. This interpretation offers an explanation of why semantically similar verbs cluster together, e.g. why *slice* and *stab* were relatively likely to have the instrument be judged as an argument, and relatively likely to be accompanied by an instrument in spoken speech. As we describe in the next section, the similarity between the root meaning of these two verbs leads them to specify event participants in similar ways. Under the interpretation that the participant judgments were generated from the corpus, this semantic commonality is unexplained.

8 General discussion

We tested three questions in this paper: whether the require/allow-instrument distinction proposed by Koenig et al. would be reflected in the judgments of participants in the lab, whether require-instrument verbs are in an equivalence class with dative verbs, and whether require- and allow-instrument verbs are categorically distinguished from each other. We found that subjects' participant judgments did reflect a distinction between require- and allow-instrument verbs, as predicted by Koenig et al. (2003). The judgments did not, however, indicate an equivalence between require-instrument verbs and dative verbs: in Experiments 1 and 2, the recipient was chosen/produced on 87%/85% of require-recipient trials, whereas the instrument was chosen/produced on only 35%/13% of require-instrument trials. This asymmetry suggests that require-instrument verbs do not encode an instrumental "slot" in their argument structure. Finally, there was variation across individual verbs in each of the four verb classes, with no clear cut-off between require- and allow-instrument verbs. This variation appeared to be partially explained by differences in verbal root meaning: the four verbs whose root encodes incision, i.e. *slice*, *cut*, *chop* and *stab*, were most likely to be judged as having an instrument argument.

In Experiments 3–5 and a corpus study (Experiment 6), we investigated how multiple sources of information about verbs and properties of the stimuli may have influenced the participant judgments. These additional studies do not substantially alter our interpretation of Experiments 1 and 2. Experiment 3 indicates that the gap between the rates of choosing the recipient vs. the instrument was likely not an effect of animacy. Experiment 4 indicates that the frequency of instruments occurring in real-world events is a weak predictor of the variation across individual instrument verb judgments. Experiment 5 suggests that participants needed to be reflecting on verbal meaning per se to generate the main findings from Experiment 1. The corpus study revealed a correlation between how often instruments and recipients were judged to be arguments and how often they are produced in spoken speech. We favor an interpretation of these data in which verbal representation influences both participant judgments and corpus frequencies, yielding the observed correlations.

The divergence between the judgments for require-instrument verbs and dative verbs suggests that the require-instrument verbs do not encode an independent slot for the instrument in their argument structure. For this reason, the distinction between require- and allow-instrument verbs, as well as the gradient result across verbs within these categories, is difficult to explain if we assume that the participant judgments reflect only a discrete list of

slots. In the section that follows, we propose an account of these seemingly divergent findings. Crucially, we propose two different mechanisms whereby verbs specify the number of participants in an event: a discrete mechanism, which we will call “argument structure,” and a gradient mechanism, driven by the interaction between a verb’s argument structure and its root meaning. Given this gradient mechanism, instruments may be judged to be crucial event participants without their being independent arguments.

8.1 Proposal

Prior to presenting our proposal, we describe some basic assumptions about how events are structured in terms of participants, how a verb’s argument structure relates to this participant structure, and how verbs lexicalize particular aspects of events through their root component. We then argue that the verbal root can influence the participant structure of an event, and that certain intransitive activity verbs demonstrate this influence. Turning to the instrumental verbs studied in this paper, we argue that require-instrument verbs lexicalize agent-oriented aspects of events. This finding motivates our main claim, that the require-instrument verbs highlight some aspect of the force exerted by the agent in the event, to a greater or lesser degree. In this way, require-instrument verbs specify an instrumental participant without the instrument being an independent argument. In addition, require-instrument verbs are distinguished from allow-instrument verbs, but this distinction is not a categorical one. After presenting this proposal, we sketch an account of how the argument judgments in our task were generated and conclude by discussing broader implications of this proposal for theories of argument structure.

8.1.1 Event representation preliminaries—Sentences with non-stative verbs indicate the existence of an event (Davidson, 1967; Parsons, 1990; Higginbotham, 2000; Pietroski, 2005). Events are structured in terms of how they unfold over time (Moens & Steedman, 1988; Krifka, 1989; Piñon, 2008) but are also structured in terms of their participants. For example, the sentence *John drank lemonade in the garden* indicates a drinking event *e*, which occupies a certain amount of event-space by virtue of its participants. The event-space includes distinct parts corresponding to the participants, which we will call “subregions” of the event. In this example, there are subregions corresponding to John, the lemonade, which are all contained in the subregion corresponding to the garden.

If no NPs were present in the sentence, the verb itself would still specify the existence of participant subregions in the event-space. For example, the sentence *drinking happened* indicates an event that is divided into agent and patient subregions. Verbs encode a discrete means of subdividing an event into participants: an event of breaking absolutely requires a participant that gets broken. An event of drinking absolutely requires someone who drinks and something that gets drunk, and these participants have equal, independent status. In other words, *drink* encodes a two-place argument structure which imposes a two-way division on a drinking event-space. For the purpose of this paper, we refer to this discrete mechanism as “argument structure,” although we make no assumptions about how this level of representation influences syntactic argument realization, if at all. In a separate mechanism, verbs can specify that these participants have particular thematic role properties, e.g. that an agent is causative and intentional (see Dowty, 1991).

In Experiments 1 and 2, we found that dative recipients were judged to be much better examples of arguments than instruments for require-instrument verbs. Given this asymmetry, we propose that both require- and allow-instrument verbs encode a two-place argument structure. This argument structure divides the event-space for the instrumental verbs into an agent subregion and a patient subregion. The agent subregion contains the event-space of the agent as well as the agent's actions. For example, the agent subregion for *drink* includes the body of the drinker as it moves through space and exerts force on the substance being drunk, which is contained within a distinct patient subregion.

Properties of verbal root meaning also play a key role in our proposal. Root meaning lexicalizes particular aspects of an event (Pinker, 1989; Levin & Rappaport-Hovav, 2010; Koenig et al., 2008). For example, consider an event of a baseball player hitting a home run. The verb *swing* lexicalizes the part of this event where an object moves in a pendular manner, whereas the verb *fly* lexicalizes the motion of an object through the air. Some verbs lexicalize agent-oriented properties of events, such as *gaze*, *memorize* and *love*, whereas other verbs lexicalize result-oriented properties, such as *shatter*, *empty* and *die*.¹¹

A verb's argument structure is primarily responsible for dividing an event-space into participant subregions. Nonetheless, we propose that verbal root meaning can also influence the partitions in event-space. Consider the verb *rollerskate*. *Rollerskate* is a one-argument activity verb, parallel to *run*, *skip* or *walk*. The sentence *rollerskating happened* indicates an event with an agentive participant region, contributed by the single agent argument of *rollerskate*. At the same time, the root meaning of *rollerskate* specifies the presence of rollerskates. In so doing, it introduces an additional subdivision of the event space: the agentive subregion must itself contain a subregion that has the characteristics of rollerskates. In this way, both argument structure and root meaning constrain how events are divided into participants. In this example, one participant is contained within the event-space subregion of another participant.

8.1.2 How verbs specify instrumental participants—As described above, we propose that the argument structure of the instrument verbs divides the event-space for these verbs into an agent subregion and a patient subregion. We propose, however, that for instrumental verbs, root meaning can introduce an additional subdivision of the agentive subregion, just as in the case of *rollerskate*. In each of the instrumental sentences in Experiment 1, the instrument occupies a particular position in the causal chain between the agent and the patient. The crucial factor for the instrument judgments is whether the position of the instrument in the causal chain corresponds to the subregion of the event lexicalized by the root meaning of the verb. For example, the require-instrument verbs *slice*, *cut*, *chop* and *stab* lexicalize an incision in the patient created via contact with a (relatively) sharp object. In *John SLICED something [with a knife]*, the knife corresponds to the incision-making object. We propose that if the instrument in a sentence corresponds to the lexicalized portion of the event, the verb will be more likely to be judged as having an instrument argument.

¹¹In some theories, verbs are limited as to the complexity of the information they can lexicalize: Rappaport-Hovav & Levin (2010) propose that verbs can lexicalize manner or result components of an event, but not both (see also Beavers & Koontz-Garboden, 2012; Husband, 2011).

Among the other require-instrument verbs, *hit* and *beat* lexicalize forceful contact between the agent and the patient. The instrument, e.g. *John hit something [with a bat]*, is the object that comes into contact with the patient. *Write*, *draw* and *stir* lexicalize an agent-driven process yielding a particular result, such as creation of orthographic symbols. For these verbs, the instrument plays an essential role in the process, i.e. it is the means by which the symbols are created. Whether the instrument highlights a key component of the agent's action is a matter of degree, giving rise to the gradient results across verbs.

For many of the allow-instrument verbs, by contrast, the instrument does not correspond to the lexicalized subregion of the event. *Break*, *open*, *move* and *grow* are strongly result-oriented: they specify a particular change undergone by the patient and may appear in the inchoative frame, e.g. *the door opened*. For these verbs, the subregion of the event that includes the instrument is prior to the patient's change of state. *Eat* and *drink* are not purely result verbs, but the instrument is external to what these verbs lexicalize, i.e. an animate agent ingesting a solid or liquid substance. Because the process of ingesting occurs inside the body of the agent, an instrument such as *[with a fork]* is external and temporally prior to this process.

The generalization that emerges from this analysis is that instruments correspond to agent-oriented lexicalizations. For example, the incision verbs (e.g. *slice*, *cut*) specify that something sharp, manipulated by an agent/causal force, comes in contact with the patient. The forceful contact verbs (e.g. *hit*, *beat*) specify that the agent acts with force. Result-oriented verbs (e.g. *open*, *break*), by contrast, were unlikely to be judged as having an instrument argument. Given this generalization, we propose that instruments are represented within the scope of the agent. Just as with *rollerskate*, the root meaning of verbs such as *slice* and *write* imposes an additional subdivision of the agentive subregion. In this way, verbs specify an instrument to a greater or lesser degree without the instrument being an independent argument. In Section 1.2, we discussed the problematic issue of body-part instruments: if I can slice something with my hand, does this mean that *slice* is not a require-instrument verb? Under our proposal, this question is a non-issue. If *slice* simply requires that there be a sharp object somewhere within the agent's scope, it is irrelevant whether the sharp object is part of an actor's body, as in *John sliced the playdoh with his hand*, or is separate, as in *John sliced the bread with a knife*. In either case, John and the hand/knife together form a complex agentive participant that fills the single agent argument role of *slice*.

A possible alternative to this proposal is one where the presence of an instrument is part of the "associative" knowledge linked to the require-instrument verbs – that is, associated in the same way that *drink* is semantically associated with alcohol and water bottles. Several pieces of evidence dispute this alternative. As shown in Experiment 4, estimates of how often instruments are involved in the real world were a weak predictor of the participant judgments for individual instrument verbs. The connection that we propose between agents and instruments also accounts for a variety of grammatical behavior. For example, in contexts where the agent is not particularly salient, an instrument may appear in subject position, e.g. *this bug spray kills cockroaches quite effectively* (see Fillmore, 1968; Marantz, 1984; Schlesinger, 1989; DeLancey, 1991). If instruments are represented within the scope

of the agent, the linking between the instrumental role and grammatical subject does not need to be stipulated. Additionally, the instrumental markers *with* and *use* are themselves sensitive to properties of the agent: *with* requires a causative agent, ruling out *John knew the answer with a calculator*, whereas *use* requires a causative and intentional agent, ruling out *Nancy tripped and accidentally used the scissors to cut her dress* (see Lakoff, 1968; Schutze, 1995; Rissman, 2011). Finally, our proposal establishes a semantic link between instrumental *with* and comitative *with*, as in *John walked to the store with Mary*. In this sentence, John and Mary are both agents of the walking event, forming a complex agent similar to the one we propose for actors and instruments.

8.1.3 How participant judgments are generated—With this framework in mind, we suggest here an account of how the participant judgments in Experiment 1 were generated. For each sentence in the experiment, e.g. *John SLICED something [with a knife] [on Tuesday]*, participants constructed an event representation with the property indicated by the verb (i.e. slicing). They then interpreted the role of each of the NPs in the sentence, i.e. that the knife was something John used to slice, not that John sliced an object that had the attribute of having a knife. Given the meaning of the verb, they accessed an intuition about how this verb divides an event-space into participant subregions. In the case of slicing, this event-space includes an agent and a patient, with the additional specification that somewhere in the agentive subregion, there is an object with a bladed shape. Participants then assessed whether the roles of the bracketed phrases matched any of the subregions specified by the verb. For this example, there would be a match between the phrase *[with a knife]* and the bladed-object subregion of the slicing event-space. In Experiment 2, subjects simply accessed the intuition about how a verb divides an event space into participants.

This matching process favored a match between a bracketed phrase and a primary participant, explaining why subjects were so likely to choose the themes and the dative recipients as arguments. This favoring resulted from the nature of our argument instructions, in which we only provided examples of matches between NPs and primary event participants (e.g. in *Jim was cooking*, Jim corresponds to the agentive participant of a cooking event). Although our description of an “argument” only demonstrated a match with a primary participant, some subjects generalized this instruction to include matches between a bracketed phrase and an event subregion specified by the verb’s root meaning. This type of generalization accounts for our finding that subjects sometimes judged the instrument to be an argument.

We also suggest that spatial goals, as in *Jim THREW something [to the outfield]*, are another type of probabilistically represented event participant, which may be more or less salient in an event depending on the verbal root. *Throw*, for example, encodes two primary participants: a person who throws and a thing that is thrown. The root of *throw* encodes movement through space, making the spatial goal, i.e. the place where the thing that gets thrown ends up, relatively salient. This degree of salience explains why subjects in Experiments 1–3 were somewhat likely to choose/list a recipient or location as one of the arguments of the allow-recipient verbs.

In summary, we propose that the participant judgments reported in this paper most closely reflect event representation, particularly how an event is divided into participant subregions. A verb's discrete argument structure plays a primary role in dividing event-space, but verbal root meaning may impose a more fine-grained division of the space. This analysis of the judgments for the require-instrument verbs explains why these verbs differed from both allow-instrument verbs and dative verbs.

8.1.4 Argument structure of require-recipient verbs—In this paper, we reasoned that if require-instrument verbs encode three argument slots, then participant judgments for these verbs should be similar to judgments for dative verbs, which have been analyzed as encoding three arguments. Although we found in Experiments 1–2 that require-recipient verbs strongly encode an instrument, the results of these experiments potentially cast doubt on the analysis that each of the require-recipient verbs discretely encodes three arguments. For *serve*, for example, rates of choosing/producing the recipient were lowest across the eight require-recipient verbs in both experiments: 82%/77%, respectively. While these values are high compared to the instrument judgments, they are not predicted under the analysis that *serve* absolutely encodes three distinct arguments.

One possibility is that only some of the verbs in our sample of require-recipient verbs actually encode a three-place argument structure. These verbs could include *lend*, *tell* and *teach*, which each had means of 88% percent or higher on both Experiments 1–2.¹² For *serve*, by contrast, representation of the recipient may be more diffuse or underspecified. For example, a waiter at a cocktail party can put a tray of food down on a buffet table and say *I served the bruschetta*, even if no one actually takes any bruschetta for another ten minutes. For events of teaching and lending, the relationship between giver and receiver may need to be more immediate and direct. In other words, a recipient is highly salient for each of the require-recipient verbs, but encoding of the recipient may still be subject to gradient effects.

A more extreme possibility is that only the verb *give* encodes three arguments in a discrete, absolute way, and the require-recipient verbs tested here are all subject to the same gradient encoding of recipients. *Give* is often taken to be a prototypical trivalent verb (see Goldberg, Casenhiser & Sethuraman, 2004) but was not tested in this paper for methodological reasons (see fn. 5). Further research is needed to understand the variation across the require-recipient verbs, and test whether *give* patterns similarly to the verbs studied here. This debate does not, however, detract from our central claim with respect to instruments: if *serve* does not necessarily encode a third argument, then *slice* and *beat* certainly do not.

8.2 Broader implications

This proposal raises multiple questions about the role of arguments in sentence processing and verb learning. As described in Section 1, Koenig et al. (2003) found that in a self-paced reading study, participants who encountered a require-instrument verb more rapidly interpreted the sentence as involving an instrument. It is unknown, however, whether instruments are subject to the same processing advantages as more prototypical arguments such as dative recipients. That is, the bias to expect a recipient upon hearing *give* may be stronger than the bias to expect an instrument upon hearing *slice*. Such a finding would

indicate that the parser is sensitive to the grammatical distinction between arguments and participants contributed via the verbal root. Alternatively, the instrument bias may be as strong as the recipient bias, indicating that the parser treats these two types of participant information as equivalent when interpreting sentences.

Our proposal also raises questions about how children learn verbal meanings. Many proponents of a syntactic bootstrapping model argue that children are biased to map the number of noun phrases in a sentence to positions in a semantic argument structure, crucially in a one-to-one fashion (Fisher, 1996; Lidz, Gleitman, & Gleitman, 2003; Fisher, Gertner, Scott, & Yuan, 2010). For example, children who hear a novel verb in a transitive context (e.g. *Cookie Monster is blicking Big Bird*) are more likely to interpret *blicking* as referring to a two-person causative action rather than a single-person activity (Naigles, 1990; Fisher, 1996; Naigles, 1996; Lidz, Gleitman & Gleitman, 2003). An important feature of such models is that children ignore modifier phrases such as *in the garden* and *on Tuesday* when determining how many arguments a verb has. Under our analysis, children would not correctly learn the meanings of require-instrument verbs such as *slice* if they ignored instrumental *with*-phrases. In *John sliced the bread with a knife*, for example, both John and the knife map to a single argument slot for *slice*. To learn this mapping, children would need to draw on both the instrumental phrase and their developing knowledge of the root meaning of *slice*. How these separate pieces interact with each other over development, ultimately yielding the adult representation of *slice*, is not well understood.

As discussed in Section 1.2, one of the difficulties in determining the argument status of instruments has been that *with*-phrases pattern as syntactic adjuncts, with no distinction between require- and allow-instrument verbs. Our proposal suggests the hypothesis that only the verb's discrete argument structure is reflected in the mapping to the verb's syntactic complements. Under this hypothesis, root-contributed participants such as instruments would only surface as syntactic adjuncts. The behavior of other participant types discussed by Koenig et al. (2003) supports this explanation. For example, Koenig et al. argue that verbs such as *boil*, *advertise* and *hide* semantically require a participant location, as in *Martha boiled the soup in a pot*. Participant locations do not appear to pattern as syntactic complements, however: *Martha boiled the soup quickly in a pot*, for example, is felicitous. A broader investigation of root-contributed participants and oblique arguments is needed to fully test this hypothesis. In *Martha baked her mother a cake*, for example, the opposite mapping holds: the benefactor *her mother* is in a syntactically privileged position but is not semantically required by the verb *bake*.

9 Conclusion

This paper offers a new perspective on the question of whether instruments are arguments: why instruments are intuitively an important part of events such as slicing and writing, but yet show a variety of non-argument like behavior. This apparent conflict emerges, we argue, because event participant representations are determined by multiple factors: the verb's discrete argument structure imposes a coarse partition on event-space, and root meaning further contributes a more fine-grained subdivision of the space. These two ways of dividing participant space do not have equal status with respect to the participant judgments we

studied in this paper, and may not exert the same influence during sentence comprehension and production. Beyond the case of instruments, our approach offers a broader framework for understanding the relationship between a verb's arguments and how verbs specify event participants. Many verbs are in fact problematic with respect to this distinction: is the fax machine a "key participant" in a faxing event? What about the location in a hiding event or the tears in a crying event? Our theory of event-space allows such participants to be among the set of verbally-specified participants, without their being part of argument structure *per se*.

Acknowledgments

This research was supported by NSF IGERT grant #9972807 awarded to the Johns Hopkins University Cognitive Science department. This work benefited greatly from comments and suggestions offered by Paul Smolensky, Akira Omaki, Colin Wilson, Jenny Culbertson, members of the Johns Hopkins Semantics and Language and Cognition labs, and audiences at WCCFL 31 and the University of Maryland PHLING research symposium. Thank you to Danny Salevitz, Yijia Hu, Michelle Chu, Lina Montoya, Allison Bellows and Christine Cheseborough for their hard work with data collection and coding. Additional thanks to all research participants for their time and effort.

References

- Anderson, JR. Learning and memory: An integrated approach. 2. Hoboken, NJ, US: John Wiley & Sons Inc; 2000.
- Babby, LH. The syntax of argument structure. Cambridge, UK; New York: Cambridge University Press; 2009.
- Baker M. Theta Theory and the Syntax of Applicatives in Chichewa. *Natural Language & Linguistic Theory*. 1988; 6(3):353–389.
- Bates, D.; Maechler, M. lme4: linear mixed effects models using Eigen and Eigen4 (Vol. R package version 0.999375-31). 2009.
- Beavers J, Koontz-Garboden A. Manner and Result in the Roots of Verbal Meaning. *Linguistic Inquiry*. 2012; 43(3):331–369.10.1162/LING_a_00093
- Bicknell K, Elman JL, Hare M, McRae K, Kutas M. Effects of event knowledge in processing verbal arguments. *Journal of Memory and Language*. 2010; 63(4):489–505.10.1016/j.jml.2010.08.004 [PubMed: 21076629]
- Blodgett A, Boland JE. Differences in the Timing of Implausibility Detection for Recipient and Instrument Prepositional Phrases. *Journal of psycholinguistic research*. 2004; 33(1):1–24. [PubMed: 15002169]
- Bock JK. Syntactic persistence in language production. *Cognitive Psychology*. 1986 Jul; 18(3):355–387.
- Boland JE. Visual arguments. *Cognition*. 2005; 95(3):237–274.10.1016/j.cognition.2004.01.008 [PubMed: 15788159]
- Borer, H. Structuring sense: Volume 2. The normal course of events. Oxford; New York: Oxford University Press; 2005.
- Bresnan, J., editor. The Mental representation of grammatical relations. Cambridge, Mass: MIT Press; 1982.
- Carlson, GN.; Tanenhaus, MK. Thematic roles and language comprehension. In: Carlson, GN.; Tanenhaus, MK., editors. *Linguistic structure in language processing: Studies in theoretical psycholinguistics*. Dordrecht; Boston: Kluwer Academic Publishers; 1989. p. 413
- Chomsky, N. Aspects of the theory of syntax. Cambridge: M. I. T. Press; 1965.
- Chomsky, N. Lectures on Government and Binding. Foris, Dordrecht: 1981.
- Clifton C Jr, Speer S, Abney SP. Parsing Arguments: Phrase Structure and Argument Structure as Determinants of Initial Parsing Decisions. *Journal of Memory and Language*. 1991; 30(2):251–271.

- Conklin K, Koenig JP, Mauner G. The role of specificity in the lexical encoding of participants. *Brain and Language*. 2004; 90(1–3):221–230. [http://dx.doi.org/10.1016/S0093-934X\(03\)00435-8](http://dx.doi.org/10.1016/S0093-934X(03)00435-8). [PubMed: 15172540]
- Cowper, E. A concise introduction to syntactic theory: the government-binding approach. Chicago, IL: University of Chicago Press; 1992.
- Croft, W. Syntactic categories and grammatical relations: the cognitive organization of information. Chicago: University of Chicago Press; 1991.
- Croft, W. Radical construction grammar: syntactic theory in typological perspective. Oxford; New York: Oxford University Press; 2001.
- Culbertson J, Gross S. Are linguists better subjects? *The British journal for the philosophy of science*. 2009; 60(4):721–736.
- Davidson D. The logical form of action sentences. *Essays on actions and events*. 1967; 5:105–148.
- Davies, M. The Corpus of Contemporary American English: 450 million words, 1990-present. 2008. Available online at <http://corpus.byu.edu/coca/>
- DeLancey, S. Event Construal and Case Role Assignment. *Proceedings of the 17th Annual Meeting of the Berkeley Linguistics Society*; 1991. p. 338-353.
- Diessel H. Frequency effects in language acquisition, language use, and diachronic change. *New Ideas in Psychology*. 2007; 25(2):108–127.
- Dowty D. Governed Transformations as Lexical Rules in a Montague Grammar. *Linguistic Inquiry*. 1978; 9(3):393–426.10.2307/4178070
- Dowty, D. Grammatical relations and Montague grammar. In: Jacobson, P.; Pullum, GK., editors. *The nature of syntactic representation*. Dordrecht: D. Reidel Publishing Company; 1982. p. 79-130.
- Dowty D. Thematic proto-roles and argument selection. *Language*. 1991; 67(3):547–619.
- Dowty, D. The dual analysis of adjuncts/complements in Categorical Grammar. In: Lang, E.; Maienborn, C.; Fabricius-Hansen, C., editors. *Modifying adjuncts*. Berlin: Mouton de Gruyter; 2003. p. 33-66.
- Edelman S, Christiansen MH. How seriously should we take Minimalist syntax? *Trends in cognitive sciences*. 2003; 7(2):60–61. [PubMed: 12584020]
- Ferreira F. Psycholinguistics, formal grammars, and cognitive science. *The Linguistic Review*. 2005; 22:365.
- Ferretti TR, McRae K, Hatherell A. Integrating Verbs, Situation Schemas, and Thematic Role Concepts. *Journal of Memory and Language*. 2001; 44(4):516–547.10.1006/jmla.2000.2728
- Fillmore, CJ. The case for case. In: Bach, EW.; Harms, RT., editors. *Universals in linguistic theory*. New York: Holt, Rinehart and Winston; 1968. p. 210
- Fisher C. Structural limits on verb mapping: The role of analogy in children’s interpretations of sentences. *Cognitive psychology*. 1996; 31(1):41–81. [PubMed: 8812021]
- Fisher C, Gertner Y, Scott RM, Yuan S. Syntactic bootstrapping. *Wiley Interdisciplinary Reviews: Cognitive Science*. 2010; 1(2):143–149.
- Gleitman L. The Structural Sources of Verb Meanings. *Language Acquisition*. 1990; 1(1):3–55.
- Goldberg, AE. *Constructions: a construction grammar approach to argument structure*. Chicago: University of Chicago Press; 1995.
- Goldberg AE. Surface Generalizations: An Alternative to Alternations. *Cognitive Linguistics*. 2002; 13(4):327–356.
- Goldberg AE, Casenhiser DM, Sethuraman N. Learning argument structure generalizations. *Cognitive Linguistics*. 2004; 15(3):289–316.
- Grimshaw, JB. *Argument structure*. Cambridge, Mass: MIT Press; 1990.
- Hale, K.; Keyser, SJ. *Prolegomenon to a theory of argument structure*. Vol. 39. MIT press; 2002.
- Hare M, McRae K, Elman JL. Sense and structure: Meaning as a determinant of verb subcategorization preferences. *Journal of Memory and Language*. 2003; 48(2):281–303. [http://dx.doi.org/10.1016/S0749-596X\(02\)00516-8](http://dx.doi.org/10.1016/S0749-596X(02)00516-8).
- Higginbotham, J. On events in linguistic semantics. In: Higginbotham, J.; Pianesi, F.; Varzi, AC., editors. *Speaking of events*. Vol. 49. Oxford University Press; 2000. p. 80

- Hill A. Grammaticality. *Word*. 1961; 17:1–10.
- Husband, EM. Rescuing manner/result complementary from certain death. *Proceedings of the 47th annual Chicago Linguistics Society*. CLS; 2011.
- Huttenlocher J, Vasilyeva M, Shimpi P. Syntactic priming in young children. *Journal of Memory and Language*. 2004; 50(2):182–195.
- Iwata, S. *Locative alternation: a lexical-constructional approach*. John Benjamins Publishing Company; 2008.
- Jackendoff, R. *X-bar syntax: a study of phrase structure*. Cambridge, Mass: MIT Press; 1977.
- Jackendoff, R. *Semantic structures*. Cambridge, Mass: MIT Press; 1990.
- Jurafsky, D. Probabilistic modeling in psycholinguistics: Linguistic comprehension and production. In: Bod, R.; Hay, J.; Jannedy, S., editors. *Probabilistic linguistics*. MIT Press; 2003.
- Klima, ES. Structure at the lexical level and its implication for transfer grammar. Paper presented at the 1961 International Conference on Machine Translation of Languages and Applied Language Analysis; London. 1962.
- Koenig JP, Davis AR. The KEY to Lexical Semantic Representations. *Journal of Linguistics*. 2006; 42(1):71–108.
- Koenig JP, Mauner G, Bienvenue B. Class Specificity and the Lexical Encoding of Participant Information. *Brain and Language*. 2002; 81(1–3):224–235. <http://dx.doi.org/10.1006/brln.2001.2519>. [PubMed: 12081394]
- Koenig JP, Mauner G, Bienvenue B. Arguments for Adjuncts. *Cognition*. 2003; 89(2):67–103. [PubMed: 12915295]
- Koenig JP, Mauner G, Bienvenue B, Conklin K. What with? The Anatomy of a (Proto)-Role. *Journal of Semantics*. 2008; 25(2):175–220.
- Krifka M. Nominal reference, temporal constitution and quantification in event semantics. *Semantics and contextual expression*. 1989; 75:115.
- Lakoff G. Instrumental Adverbs and the Concept of Deep Structure. *Foundations of Language*. 1968; 4(1):4–29.
- Landau, B.; Gleitman, LR. *Language and experience: evidence from the blind child*. Cambridge, Mass: Harvard University Press; 1985.
- Langacker, RW. *Foundations of cognitive grammar*. Stanford, Calif: Stanford University Press; 1987.
- Larson RK. On the Double Object Construction. *Linguistic Inquiry*. 1988; 19(3):335–391.10.2307/25164901
- Levin, B. *English verb classes and alternations: a preliminary investigation*. Chicago: University of Chicago Press; 1993.
- Levin B, Rappaport M. Nonevent -er Nominals: a Probe into Argument Structure. *Linguistics*. 1988; 26(6):1067–1083.
- Lidz J, Gleitman H, Gleitman L. Understanding how input matters: verb learning and the footprint of universal grammar. *Cognition*. 2003; 87(3):151–178.10.1016/s0010-0277(02)00230-5 [PubMed: 12684198]
- Liversedge SP, Pickering MJ, Branigan HP, van Gompel RPG. Processing arguments and adjuncts in isolation and context: The case of by-phrase ambiguities in passives. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. 1998; 24(2):461–475.10.1037/0278-7393.24.2.461
- MacDonald MC, Pearlmutter NJ, Seidenberg MS. The lexical nature of syntactic ambiguity resolution. *Psychological review*. 1994; 101(4):676–703.10.1037/0033-295x.101.4.676 [PubMed: 7984711]
- Manning, CD. Probabilistic syntax. In: Bod, R.; Hay, J.; Jannedy, S., editors. *Probabilistic linguistics*. MIT Press; 2003. p. 289–341.
- Marantz, A. *On the nature of grammatical relations*. Cambridge, Mass: MIT Press; 1984.
- McRae K, Ferretti TR, Amyote L. Thematic Roles as Verb-specific Concepts. *Language & Cognitive Processes*. 1997; 12(2):137–176.10.1080/016909697386835
- Moens M, Steedman M. Temporal ontology and temporal reference. *Computational Linguistics*. 1988; 14(2):15–28.

- Naigles L. Children use syntax to learn verb meanings. *Journal of Child language*. 1990; 17(02):357–374. [PubMed: 2380274]
- Naigles LR. The use of multiple frames in verb learning via syntactic bootstrapping. *Cognition*. 1996; 58(2):221–251. [http://dx.doi.org/10.1016/0010-0277\(95\)00681-8](http://dx.doi.org/10.1016/0010-0277(95)00681-8). [PubMed: 8820388]
- Parsons, T. *Events in the semantics of English: a study in subatomic semantics*. Cambridge, MA: MIT Press; 1990.
- Pesetsky, D. *Zero syntax: experiencers and cascades*. Cambridge, MA: MIT Press; 1995.
- Pietroski, P. *Events and semantic interpretation*. Oxford: Oxford University Press; 2005.
- Pinker, S. *Learnability and Cognition: the Acquisition of Argument Structure*. Cambridge, MA: MIT Press; 1989.
- Piñón C. Aspectual composition with degrees. *Adjectives and adverbs: Syntax, semantics, and discourse*. 2008:183–219.
- Pollard, C.; Sag, IA. *An information-based syntax and semantics. Volume 1: Fundamentals*. Leland Stanford Junior University, Center for Study of Language and Information; 1987.
- Quené H, van den Bergh H. Examples of mixed-effects modeling with crossed random effects and with binomial data. *Journal of Memory and Language*. 2008; 59(4):413–425.
- Radford, A. *Transformational grammar: a first course*. Cambridge; New York: Cambridge University Press; 1988.
- Ramchand, GC. *Verb meaning and the lexicon: A first phase syntax*. Vol. 116. Cambridge University Press; 2008.
- Rappaport-Hovav, M.; Levin, B. Building verb meanings. In: Butt, M.; Geuder, W., editors. *The projection of arguments: Lexical and compositional factors*. Stanford, CA: CSLI publications; 1998. p. 97-134.
- Rappaport-Hovav M, Levin B. The English dative alternation: The case for verb sensitivity. *Journal of Linguistics*. 2008; 44(1):129.
- Rappaport-Hovav, M.; Levin, B. Reflections on Manner/Result Complementarity. In: Doron, E.; Sichel, I., editors. *Syntax, Lexical Semantics, and Event Structure*. Oxford, UK: Oxford University Press; 2010. p. 21-38.
- Reinhart T, Siloni T. The Lexicon-Syntax Parameter: Reflexivization and Other Arity Operations. *Linguistic Inquiry*. 2005; 36(3):389–436.
- Resnik P. Selectional constraints: an information-theoretic model and its computational realization. *Cognition*. 1996; 61(1–2):127–159. [http://dx.doi.org/10.1016/S0010-0277\(96\)00722-6](http://dx.doi.org/10.1016/S0010-0277(96)00722-6). [PubMed: 8990970]
- Rissman, L. *Instrumental with and use: modality and implicature*. Paper presented at the Proceedings of SALT 21; 2011.
- Roland D, Yun H, Koenig JP, Mauner G. Semantic similarity, predictability, and models of sentence processing. *Cognition*. 2012; 122(3):267–279. [PubMed: 22197059]
- Schlesinger IM. Instruments as Agents: On the Nature of Semantic Relations. *Journal of Linguistics*. 1989; 25(1):189–210.
- Schutze CT. PP Attachment and Argumenthood. *MIT Working Papers in Linguistics*. 1995; 26(Sept): 95–151.
- Schutze CT, Gibson E. Argumenthood and English Prepositional Phrase Attachment. *Journal of Memory and Language*. 1999; 40(3):409–431.
- Sedivy, JC.; Spivey-Knowlton, M. The use of structural, lexical, and pragmatic information in parsing attachment ambiguities. In: Clifton, C.; Frazier, L.; Rayner, K., editors. *Perspectives on sentence processing*. Hillsdale, NJ: Lawrence Erlbaum; 1994. p. 389-413.
- Snedeker J, Trueswell JC. The developing constraints on parsing decisions: The role of lexical-biases and referential scenes in child and adult sentence processing. *Cognitive psychology*. 2004; 49(3): 238–299. [10.1016/j.cogpsych.2004.03.001](http://dx.doi.org/10.1016/j.cogpsych.2004.03.001) [PubMed: 15342261]
- Speer SR, Clifton C Jr. Plausibility and Argument Structure in Sentence Comprehension. *Memory & cognition*. 1998; 26(5):965–978. [PubMed: 9796230]
- Sprouse, J. *A program for experimental syntax*. College Park, MD: University of Maryland dissertation; 2007.

- Sprouse J, Almeida D. Assessing the reliability of textbook data in syntax: Adger’s Core Syntax. *Journal of Linguistics*. 2012; 48(3):609–652.
- Taraban R, McClelland JL. Constituent Attachment and Thematic Role Assignment in Sentence Processing: Influences of Content-Based Expectations. *Journal of Memory and Language*. 1988; 27(6):597–632.
- Tesnière, L. *Elements De Syntaxe Structurale*. Paris: C. Klincksieck; 1959.
- Thothathiri M, Snedeker J. Syntactic priming during language comprehension in three- and four-year-old children. *Journal of Memory and Language*. 2008; 58(2):188–213.
- Vater, H. *Valence, semantic case, and grammatical relations*. Amsterdam: John Benjamins; 1978. On the possibility of distinguishing between complements and adjuncts.
- Vestergaard, T. *Prepositional phrases and prepositional verbs: a study in grammatical function*. The Hague: Mouton; 1977.
- Wittenberg E, Snedeker J. It takes two to kiss, but does it take three to give a kiss? Categorization based on thematic roles. *Language and Cognitive Processes*. 2013:1–7. ahead-of-print.

Appendix A

Filler verbs

Carry	Read	Like
Bring	Learn	Admire
Pull	Study	Respect
Measure	Memorize	Envy
Describe		Hate
Discuss		Pity
Mix		
Blend		
Combine		
Bake		

Appendix B

Full argument instructions

In this task, you will be asked about the meanings of verbs. People who study language often talk about verbs having “arguments.” Basically, arguments are entities that are not part of the verb itself, but are absolutely necessary for the verb’s meaning. For example, in order to describe a situation as “hugging,” there has to be a person who hugs someone and a person who is hugged. Thus the arguments of “hug” are the person doing the hugging and the person who is hugged. To describe a situation as “wanting,” there must be a person doing the wanting and something that is wanted. Thus the arguments of “want” are the person who wants and the thing that is wanted.

Page 2

The arguments of a verb are often mentioned when we use that verb.

For example, the arguments of “hug” and “want” are shown below:

- (1) The boy wanted a popsicle.
- (2) The girl hugged her pet turtle.

The arguments of a verb don't NEED to be mentioned in a sentence, however.

In (3), for example, only one argument is mentioned, the person or thing being hugged:

- (3) Mary was hugged.

The arguments of a verb are understood to be part of the action, even when we don't mention them. That is, in (3) we understand that Mary was hugged BY someone.

Page 3

Verbs can have different numbers of arguments.

“Hug” and “want” each have two arguments, but some verbs only have one:

- (1) John ran.
- (2) The ice melted.

To describe a situation as “running,” there only needs to be one entity involved, the person doing the running. Similarly for “melt,” there only has to be one entity involved, the thing that is melting.

Some verbs have more than two arguments:

- (3) John put the book on the table.

To describe a situation as “putting,” there has to be a person doing the putting, a thing that gets put, and a place where it gets put.

Page 4

Things that are striking or unusual in a sentence aren't necessarily arguments.

In (4) and (5), John is still the only argument of run:

- (4) John ran until he was sick.
- (5) John ran because a zombie was chasing him.

Remember, arguments are absolutely necessary for the verb's meaning. We can understand that John was running without knowing WHY he was running. So, people think that the phrase “because a zombie was chasing him” is not one of the arguments of “run.”

The ‘reason’ phrase “because a zombie was chasing him” and the ‘duration’ phrase “until he was sick” aren't arguments of “run.” That doesn't necessarily mean, however, that reason and duration phrases will NEVER be an argument of any verb. That is, some types of phrases might be arguments for some verbs but not others. Other types of phrases might always be arguments, or might never be arguments.

Highlights

- We investigate whether instruments are arguments of verbs such as “slice”
- We present a new method for gathering judgments about verbal meaning
- We propose that instruments are not independent arguments, but may be more or less salient parts of the force exerted by the agent
- We propose that event participant representation has both a discrete component and a gradient component

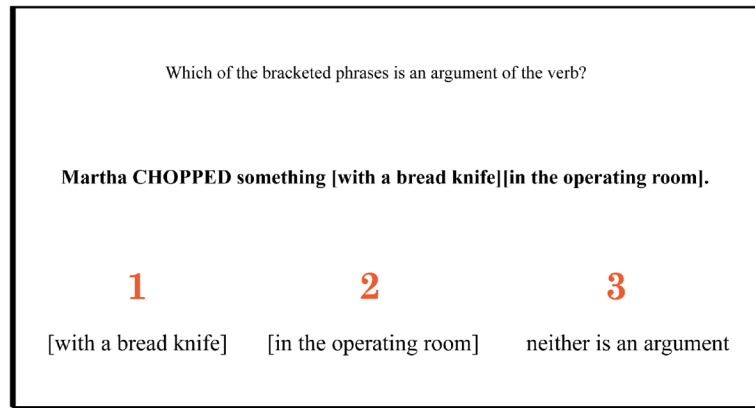


Figure 1.
Sample display for argument judgment task

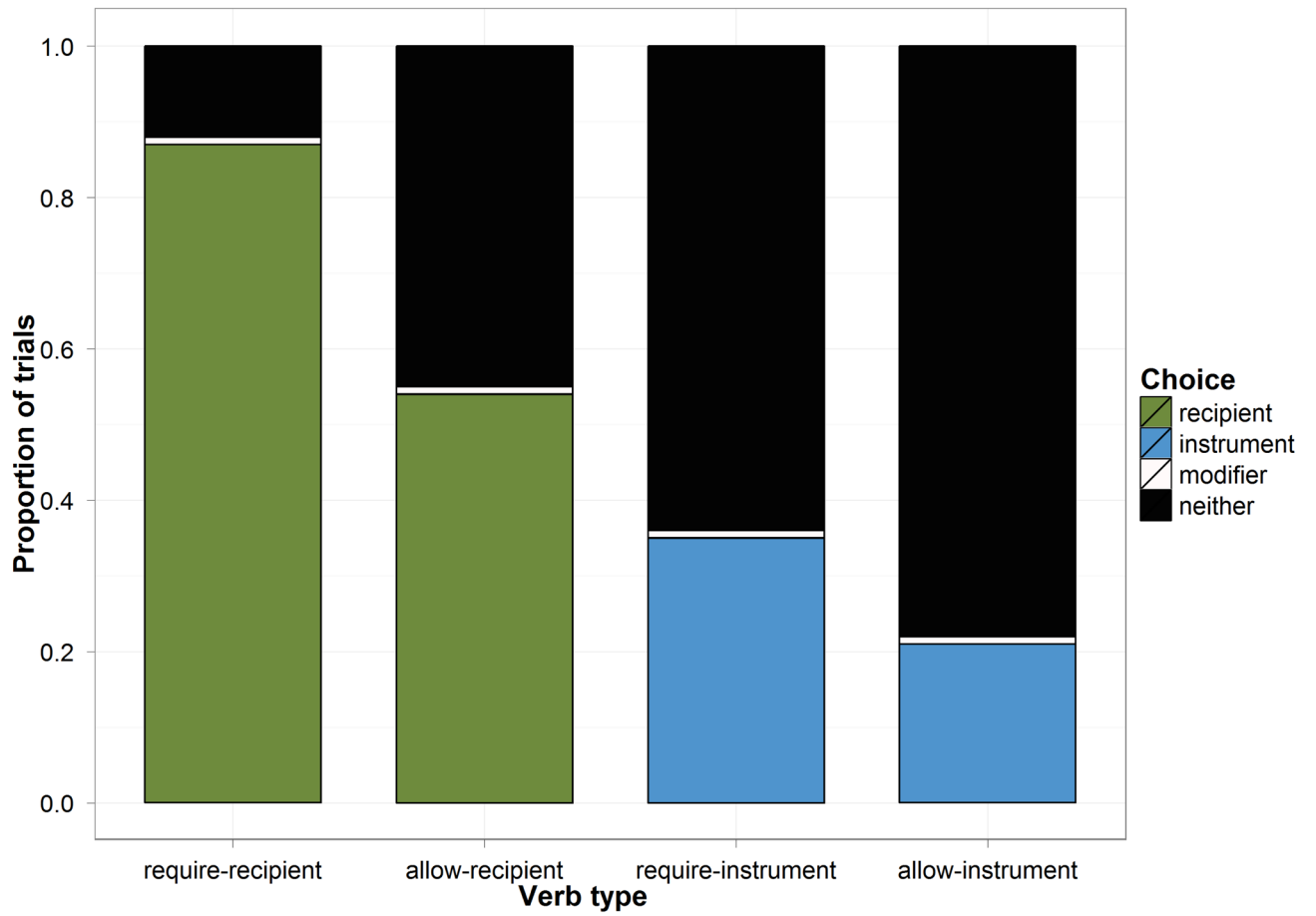


Figure 2.
Mean choice of recipient/instrument vs. modifier vs. neither across verb types

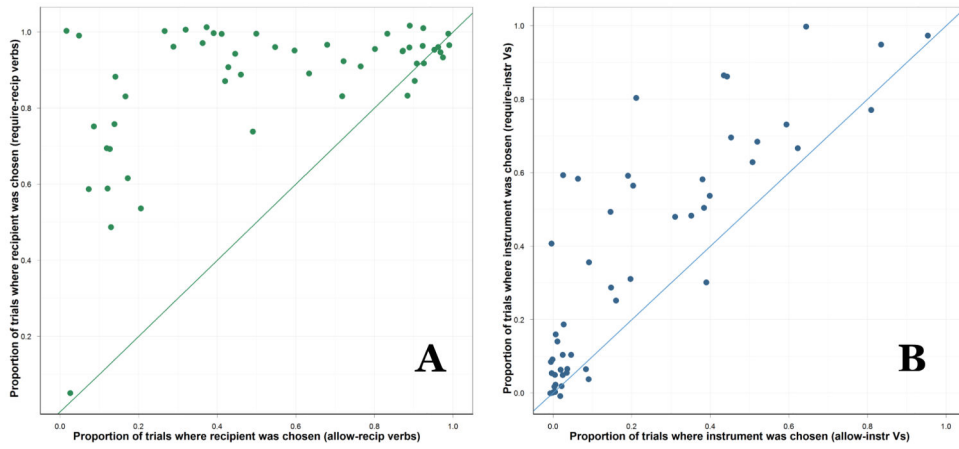


Figure 3. Rates of choosing the recipient (A) or the instrument (B) for allow vs. require verbs. Each dot represents a single subject

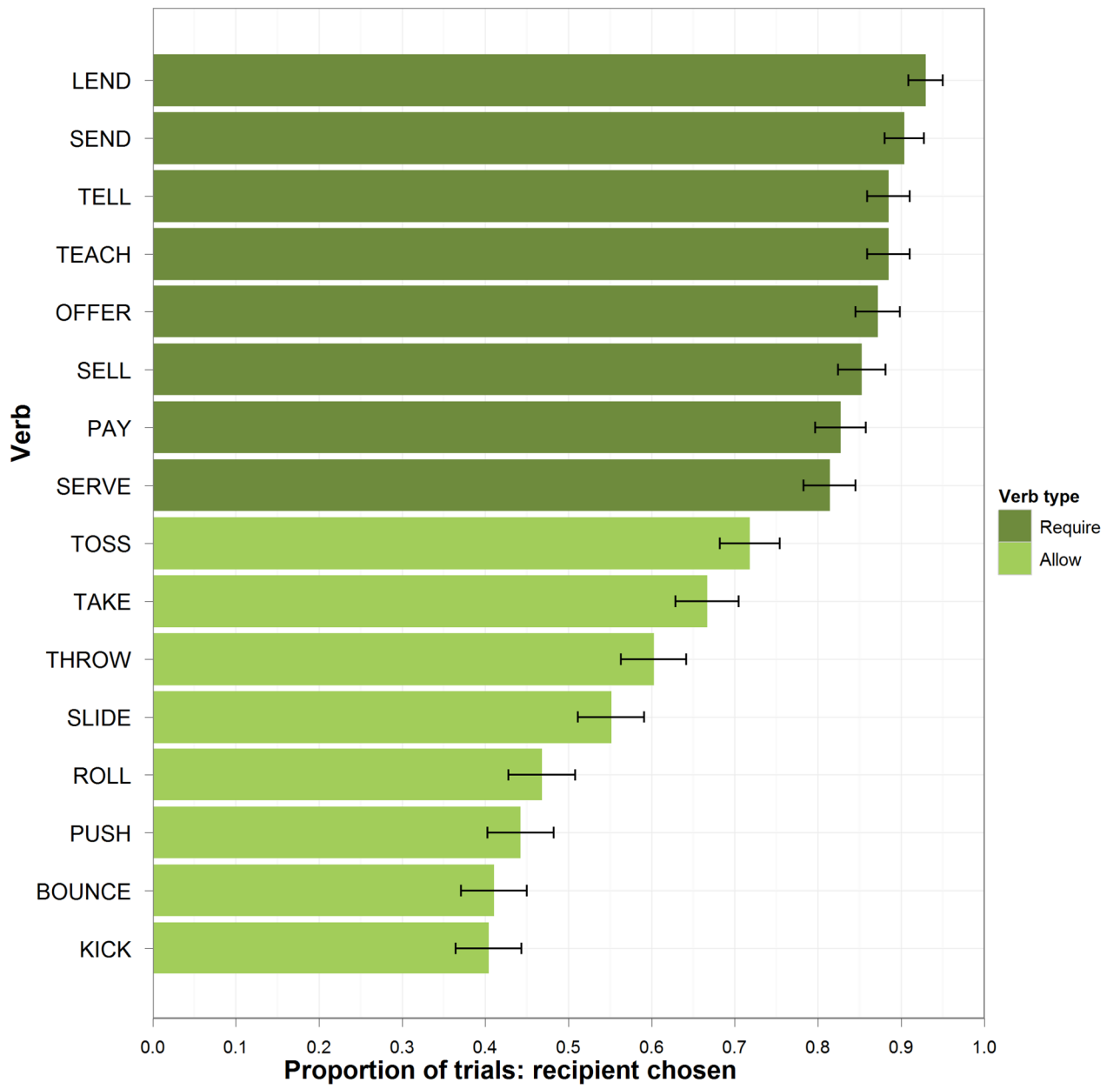


Figure 4.
Argument choice means for individual recipient verbs

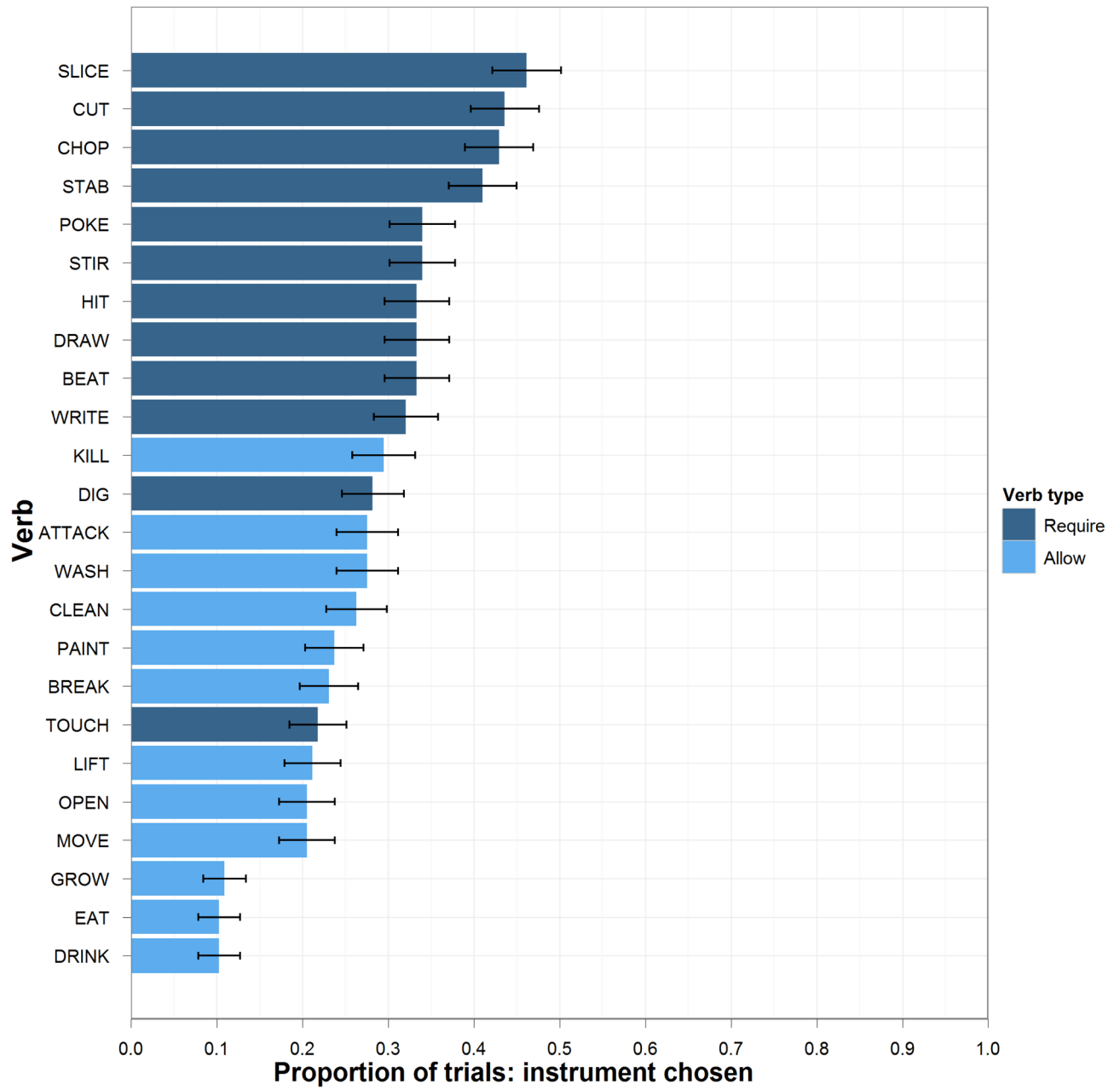


Figure 5.
Argument choice means for individual instrument verbs

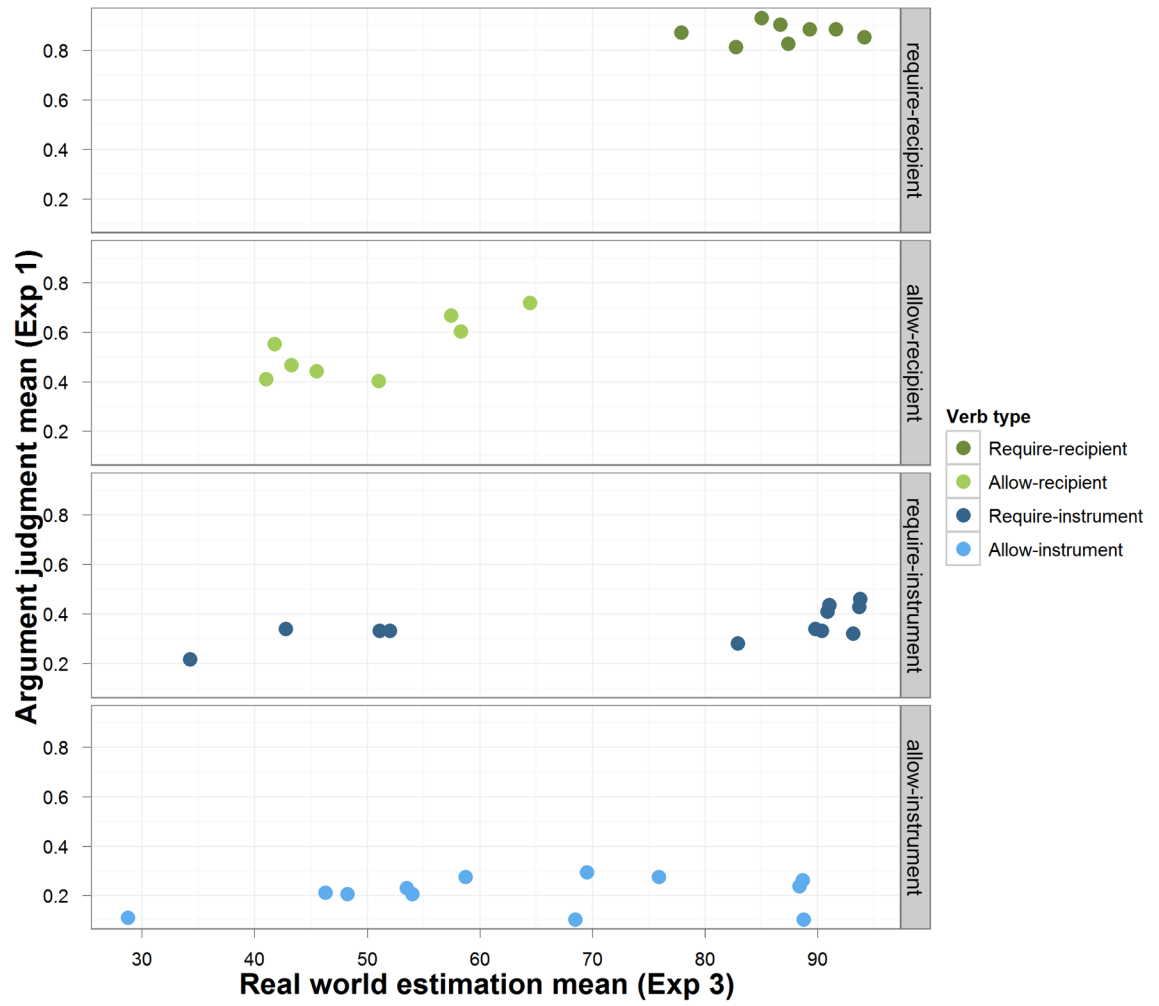


Figure 6. Real-world estimation mean from Experiment 4 vs. argument judgment mean from Experiment 1 for each verb type. Each dot represents an individual verb.

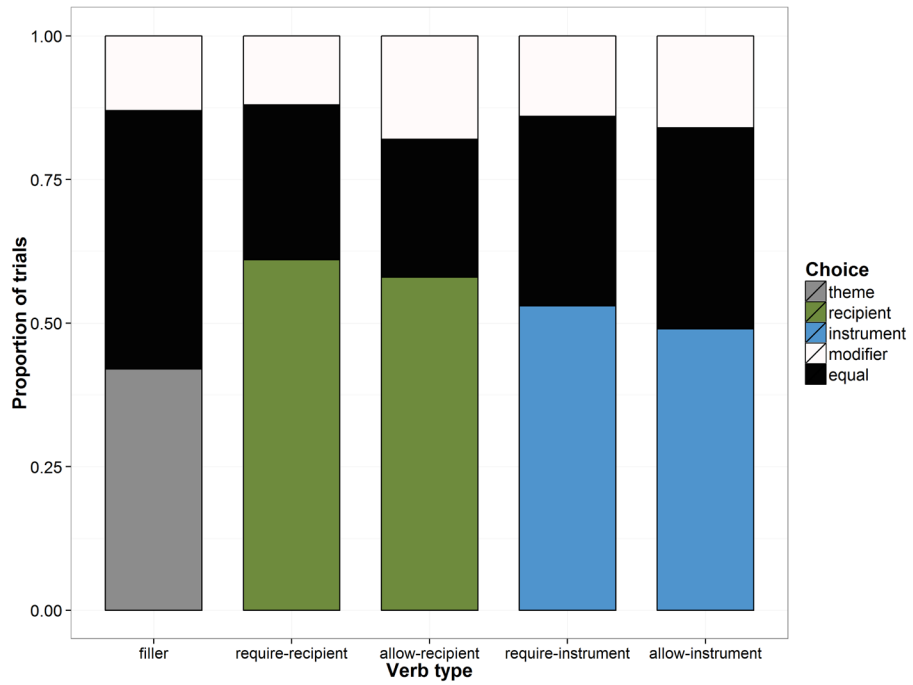


Figure 7.
Importance judgment means for each verb type in Experiment 4

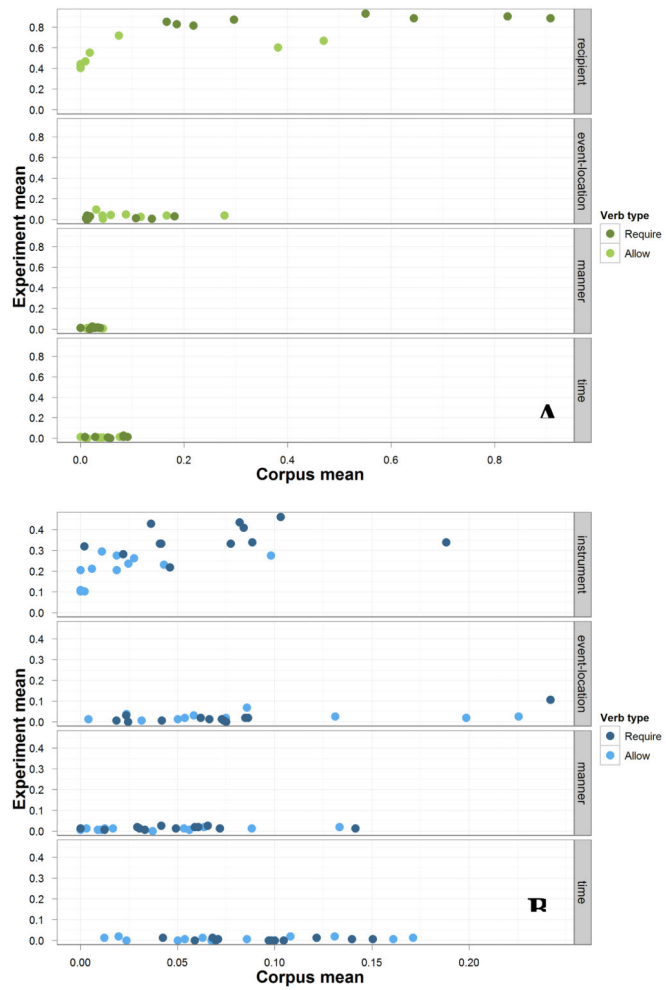


Figure 8. Corpus frequencies of recipients, instruments, locations, times and manners vs. argument choice means from Experiment 1 for recipient verbs (A) and instrument verbs (B). The A and B figures differ in scale on both the x- and y-axes.

Table 1

Instrument and recipient verbs

Require instrument	Allow instrument	Require recipient	Allow recipient
Touch	Eat	Serve	Kick
Hit	Drink	Teach	Throw
Beat	Break	Send	Toss
Poke	Open	Tell	Roll
Stab	Kill	Sell	Push
Cut	Attack	Lend	Slide
Chop	Paint	Pay	Take
Slice	Grow	Offer	Bounce
Write	Move		
Draw	Lift		
Dig	Clean		
Stir	Wash		

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 2

Instrument/recipient verb contrasts

Instrument verbs		Recipient verbs	
Contrast type	Contrast	Contrast type	Contrast
Instrument vs. modifier	Instrument vs. event location	Recipient vs. modifier	Recipient vs. event location
	Instrument vs. manner		Recipient vs. manner
	Instrument vs. time		Recipient vs. time
Modifier vs. modifier	Event location vs. manner	Modifier vs. modifier	Event location vs. manner
	Event location vs. time		Event location vs. time
	Time vs. manner		Time vs. manner

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 3

Results from Experiment 2

	Theme	Recipient	Participant location	Instrument
Filler verb	.98 (.01)			
Require-recipient verb	.94 (.02)	.85 (.03)	.03 (.01)	
Allow-recipient verb	.97 (.01)	.02 (.01)	.22 (.03)	
Require-instrument verb	.94 (.01)			.13 (.02)
Allow-instrument verb	.97 (.01)			.03 (.01)

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript