Talking about Causing Events

Alexis Wellwood, Chris Vogel, Rachel Dudley, J. Brendan Ritchie

March 1, 2013

Abstract

Questions about the nature of the relationship between language and extralinguistic cognition are old, but only recently has a new view emerged that allows for the systematic investigation of claims about linguistic structure, based on how it is understood or utilized outside of the language system. Our paper represents a case study for this interaction in the domain of event semantics. We adopt a transparency thesis about the relationship between linguistic structure and extralinguistic cognition, investigating whether different lexico-syntactic structures can differentially recruit the visual causal percept. A prominent analysis of causative verbs like move suggests reference to two distinct events and a causal relationship between them, whereas non-causative verbs like push do not so refer. In our study, we present English speakers with simple scenes that either do or do not support the perception of a causal link, and vary (between subjects) a one-sentence instruction for the evaluation of the scene. Preliminary results suggest that causative constructions in the linguist’s sense are judged true by competent speakers of English significantly more when cause-based judgments of the scene obtain than do non-causative constructions. Implications for the “new view” and future directions are discussed.

1. Introduction

Since Davidson’s (1967) classic treatment of action sentences as involving reference to the ontological category event, the “event analysis” has only increased in popularity; and, in the face of new linguistic generalizations, it has also increased in complexity. It is now fairly standard linguistic practice to analyze sentences containing a causative verb like move in (1) as involving reference to not one, but two events: the causing event (e.g., Al’s action), and the caused event (the table’s movement).

(1) Al moved the table.
Correspondingly, since at least Michotte (1946) there has been a great deal of research investigating the psychological mechanisms employed in event perception. The emerging view is that there are event percepts that exist independently of how we think about “things that happen”, and which function to discretize the continuous flow of experience into useful units that feed other processes (Liverence & Scholl 2012). The perception of causality between two events has, in particular, been shown to be fast, automatic, irresistible, stimulus-driven, and independent of the animacy of the causer (Scholl & Tremoulet 2000).

Here, we offer a demonstration of how these literatures might be brought together, by asking how event structure as encoded in the logical forms of sentences might interact with extralinguistic perception and cognition. Our major hypothesis is that the event variables appealed to in the logical analysis of sentences like (1) range over representations of events in the psychologist’s sense, and that event percepts, if structured appropriately, can represent the satisfiers of linguistically structured predicates and propositions. In our test case, we show how appeal to a two-place causal concept in the semantics of causative sentences can predict the recruitment of simple, visually-constructed event representations.

This paper thus represents a small contribution to a new and growing literature that uses well-understood properties of visual processing to attempt reverse inferences into the nature of natural logical form itself. Supposing that the interface between language and extralinguistic cognition is natural (Baker 1997) or transparent (Goldman 2007), examining the types of cognitive processes engaged when a sentence is evaluated for truth or falsity can lend insight into the format of the meaning encoded by the sentence. This is dubbed the Interface Transparency Thesis, or ITT.

\[(2) \textbf{Interface Transparency Thesis:} \text{ The verification procedures employed in understanding a declarative sentence are biased towards algorithms that directly compute the relations and operations expressed by the semantic representation of that sentence (Lidz, Hunter, Pietroski, & Halberda 2011, p233).}\]

It has been shown, for example, that if the compositional semantics of a sentence specifies the computation of a greater-than relation, English speakers will evaluate instances of that sentence using visual set selection/enumeration, and not e.g. a one-to-one matching procedure—even when the visual scene biases towards the latter strategy (\textit{ibid.}).

On our view, the compositional semantics of a sentence is a recursive specification of instructions to build complex concepts out of simpler ones (Pietroski 2010, in prep). “Lexical concepts”, novel to the language faculty, form the basis for this composition. Lexical concepts are introduced during language acquisition as a bridge between (arbitrary) phonological forms and (antecedently available) primitive concepts, which themselves imbue sentences with our sense of their meaning. On our view, to decide whether a given sentence is true, one must be able to find an extralinguistic representation that is structured or structurable in the way demanded by the predicates and relations encoded in the sentence. Put another
way, “meanings” can ground judgments of truth just in case the existential claims they express can be verified by correlative representations in perception or cognition.

Unlike other views on the nature of the relationship between language and extralinguistic perception and cognition, this view suggests an interdisciplinary arena in which philosophers, linguists, and psychologists can work together to gain new insight into the nature of language and the conceptual structure of mind. In our case study, we consider the semantics of causal eventive sentences and the psychology of event perception in §2, generating novel predictions that we test experimentally in §3. The punchline is that slight variations in task instructions result in quite different patterns of behavior, where the differences are predicted by the event analysis assigned to the variant sentences.

2. Event language

Much ink has been spilled over how best to capture the fact that certain verbs, like move, are capable of systematically appearing both in a transitive (3a) and intransitive frame (3b). This pattern, dubbed the causative-inchoative alternation, is robust in English and many other languages. Semantically, whenever a verb so alternates, the transitive form (3a) entails the intransitive form (3b).

(3) a. The red ball moved the blue ball.
b. The blue ball moved.

This pattern is interesting insofar as there are other verbs whose meanings do not appear all that different on the surface, yet which do not show the alternation. The non-causative verb push is rigidly transitive—i.e., (4a) is grammatical, but (4b) is not. (A “*” preceding a sentence indicates ungrammaticality.)

(4) a. The red ball pushed the blue ball.
b. *The blue ball pushed.

Historically, there has been a progression from semantic analyses that treat move and push as essentially interchangeable, to ones explicitly encoding generalizations about alternations into their logical forms. For example, on classical and Davidsonian approaches, nothing formally distinguishes CAUSATIVE (move) from NON-CAUSATIVE (push) verbs: classically, both denote two-place predicates (5a-5b), and on Davidson’s (1967) analysis, the classical predicates are made three-place by the addition of an event argument (6a-6b).

\[1\] For reasons of space, we do not review Davidson’s motivations for introducing the event argument. One linguistically-motivated reason was that only event-based analyses adequately capture the “diamond-shaped” pattern of inferences generable from sentences with multiple adjunct phrases like Brutus stabbed Caesar in the back with a knife.
(5)  **Classical**
   a.  move(Red, Blue)
   b.  push(Red, Blue)

(6)  **Davidsonian**
   a.  $\exists e [move(e, Red, Blue)]$
   b.  $\exists e [push(e, Red, Blue)]$

On such views, the truth or falsity of sentences is determined based on whether arbitrary lexical properties of the verb are satisfied.

Parsons (1990) introduced what has come to be known as the Neodavidsonian analysis, in which all but the event argument has been severed from the verb’s denotation, allowing for the explicit labeling of the roles played by various entities in the event described by a sentence. For causative verbs, Parsons posits a “hidden” predicate cause that relates the inchoative event to the causative event (7a); non-causative verbs have no such hidden structure (7b). The “external argument” to the verb is labeled with the thematic role Agent, and the “internal argument” is labeled Theme. On Parsons analysis of causative sentences, Agent and Patient name roles in distinct events, here $e$ and $e'$ respectively; in non-causative sentences, they are related to a single event $e$.

(7)  **Neodavidsonian**
   a.  $\exists e [\text{Agent}(e, Red) \& \exists e'[\text{move}(e') \& \text{Theme}(e', Blue) \& \text{cause}(e, e')]]$
   b.  $\exists e [\text{Agent}(e, Red) \& \text{push}(e) \& \text{Theme}(e, Blue)]$

Notice that, with this shift, it becomes possible to capture the entailment patterns distinguishing causative from non-causative sentences: in (7a), reference to the inchoative event $e'$ denoted by the intransitive (3b) is contained within the reference to the causative event $e$ denoted by the transitive (3a).

However, Thomson (1987; also Fodor & Lepore 1998, Pietroski 2005) cautioned against appeal to a cause predicate in sentences with causative verbs. Imagine a situation where Al sets the house on fire, and as a consequence some water sitting in a pot on the stove boils. Under these circumstances, speakers will assent to (8b) as an accurate description of what happened, but not to (8a): (8a) seems to imply a more direct relation between Al and the boiling event than is satisfied in such circumstances.

(8)  a.  Al boiled the water.

---

2 Again, we do not review his motivations. But notice, for one thing, that severing the arguments makes the job of accounting for nominalizations easier, e.g. those sentences where events are explicitly, nominally referred to: *Al saw Brutus stab Caesar* $\equiv$ *Al saw the stabbing of Brutus by Caesar.*

3 Dowty 1991 discusses ‘proto-roles’, whereby Agents are understood as whatever has the most Agent-like properties, and so on for other thematic roles. See Fillmore 1970 for early discussion of thematic roles.
b. Al caused the water to boil.

Since, the objection continues, both (8a-8b) will invoke cause, they should be synonymous, contrary to fact: hence, Parsons’ analysis cannot be correct.

We think this rejection of Parsons’ analysis is too quick, as it relies on an presumed identity between an implicit predicate in logical form, and an explicit predicate in English. Yet, these predicates differ at least as much as open and closed class lexical items do: the meanings of open class words vary wildly from person to person even within a given language, while the meanings of function words are uniform, constituting as such core grammatical knowledge. The English word cause is part of the “open-class” vocabulary: if we interpret this as cause (on analogy with dog, or push), we can differentiate the predicate appearing in causative sentences that is part of the closed class vocabulary Cause. There is nothing implausible to the claim that the causal predicate introduced reflexively in the syntax of certain constructions has a different interpretation than what underlies the interpretation of the English word cause. Indeed, as Thomson herself pointed out, the semantic relation between a verb and its external argument (the subject of a transitive clause) seems to be primitive in some sense, and the word cause likely came along somewhat later when social attentions shifted in particular directions.

Following Folli and Ramchand (2005), we assume that causative verbs like move surface in syntactic structures at least as complex as that in (9a), and non-causative verbs like push in those like (9b).

(9)

a. vP  
   Al v  
   Cause VP  
   V Bill

b. VP  
   Al V Bill
   push
   move

The interpretation of Cause is given in (10) (cf. Pylkkänen 2002). Its semantic function is to bind the open event argument introduced by the verb move, and introduce a new argument for the causing event.

4 “Open class” lexical items are, roughly, the “content” words, things like dog, cat, run, red, and happy. This class is not fixed in an individual’s youth, normal speakers update their open class lexicon regularly throughout their life. “Closed class” lexical items correspond to function words like of, the, a, various forms of verbal inflection, and words like which and who etc. This class is called “closed”, as it is rare, if not nonexistent, to add items to this class past the critical period for language acquisition.

5 “Little-v” is a functional verbal category headed by Cause; see ibid., and references therein, for relevant crosslinguistic evidence.
Thus, the interpretation of (9a) is (11a) and (9b) is (11b). On what we will call the “cognitive interpretation” of such forms, the predicate *Cause* is true of pairs of event representations \(<e, e'>\), such that the first bears the (cognitive) relation *Cause* to the second. Here, (11a) will be verified by an event representation with two component events, one in which an Agent acts such that a Theme is *Caused* to move, and (11b) by an event representation with only one component, an event in which an Agent acts on a Theme.

\[
(11) \begin{align*}
a. \ & \exists e' [\text{Agent}(e', \text{Red}) \land \exists e [\text{MOVE}(e) \land \text{Cause}(e', e) \land \text{Theme}(e, \text{Blue})]] \\
b. \ & \exists e [\text{Agent}(e, \text{Red}) \land \text{PUSH}(e) \land \text{Theme}(e, \text{Blue})]
\end{align*}
\]

3. Experiments at the interface

We discussed the analysis of small linguistic structures, some of which refer to multiple events and a causal relation between them, and others referring to only one event. We posited that the satisfiers of logical forms are representations in extralinguistic cognition; can we find evidence for this? It is often true that seemingly high level concepts are influenced by perception, e.g. causality and animacy (Scholl and Tremoulet 2000). So simple cases giving rise to cognitive representations of events may be ones in which causal events are perceived.

In designing our study, we used simple stimuli that have been shown to give rise to causal percepts, i.e. the paradigm developed in Michotte (1946). In his stimuli, a ball of one color moves across the screen to come into contact with a ball of a different color, as shown schematically in Figure 1. Michotte reported that, within a certain temporal window between the moment of contact between two balls, subjects report perceiving that the second movement was caused by the first. In more contemporary experiments (see Scholl & Tremoulet 2000 for review), it has been shown that spatial parameters can be varied and still evoke causal percepts, showing contact between the balls is not necessary.

Drawing on these studies, our experiment was designed as follows. We created five temporal **LAG** levels, ranging from 0ms to 200ms (0ms, 50ms, 100ms, 150ms, 200ms), and five spatial **GAP** levels, ranging from 0cm to 2cm (0cm, .5cm, 1cm, 1.5cm, 2cm), for a total
of 250 trials/subject, or 8 per LAG-GAP combination. The balls moved at a rate of 40 degrees of visual angle a second, a velocity that reliably produces causal percepts (cf. Scholl & Nakayama 2002). While experiments investigating causal perception typically provide extensive instructions (and often training; e.g. Schlottman & Anderson 1993) to ensure subjects’ responses track causal events, our interest was in how subjects would respond given minimal, targeted instructions. We reasoned that the event structure they infer from a target sentence’s meaning would guide their performance in the task.

We ran four conditions, each corresponding to a change in a target sentence of the instructions (12A-D). Subjects were told they would be presented with short animated scenes and asked to evaluate, for each scene, whether the TARGET was true of that scene. In addition to push and move, we used sentences containing the “indirect” causative cause to move and the causative verb launch.6

(12) **Experimental conditions by target sentence and “cause level”**

<table>
<thead>
<tr>
<th>Condition</th>
<th>Target</th>
<th>Cause level</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>The red ball pushed the blue ball.</td>
<td>low</td>
</tr>
<tr>
<td>B</td>
<td>The red ball caused the blue ball to move.</td>
<td>mid-low</td>
</tr>
<tr>
<td>C</td>
<td>The red ball moved the blue ball.</td>
<td>high</td>
</tr>
<tr>
<td>D</td>
<td>The red ball launched the blue ball.</td>
<td>high</td>
</tr>
</tbody>
</table>

Pilot data with subjects evaluating an intransitive sentence containing the (non-causative) verb touch (i.e., the red ball and the blue ball touched) revealed a pattern of responses in which only conditions where contact was made were judged true. Based on these data and the results from the causal perception literature, we postulated two possible interpretations of the targets that would give rise to ‘true’ answers: CONTACT, or CAUSE (cf. Young & Sutherland 2009, p733). The (idealized) patterns of responses if subjects choose one or the other interpretation for our experiment is shown in Figure 2.

It appeared to us, intuitively, that any of the sentences in (12A-D) might require contact, and deliver the graph on the right in Figure 2. However, if any invoked a truly causal interpretation, we would expect the graph on the left. Importantly, the latter result would imply a non-trivial relationship between the sentence and recruitment of the causal percept. In terms of “Cause level” in (12), we reasoned that as push is not a causal verb, it shouldn’t recruit the causal percept. In contrast, (B-D) all feature causal constructions, and thus would be more likely to recruit the causal percept. The cause to X construction is ranked lower on expected causality given that it allows for the “indirect” interpretation discussed in the previous section.

We tested thirty undergraduates who received either course credit or $10 for participating,  

---

6Causative, as determined by participation in the causative-inchoative alternation: witness The red ball launched the blue ball; so, the blue ball launched. We included launch as this verb is often used to describe Michotte-like stimuli, see Scholl & Tremoulet 2000.
Figure 2: **Idealized plots reflecting two interpretations of Michotte-type stimuli**

all native speakers of American English as determined in a pre-test questionnaire (8 in each of (A,B), 7 each of (C,D)).

**Results.** Given the large number of factors in this experiment, the full statistical analysis is presented in the Appendix. In this section, we present an overview of the results in light of the preceding discussion.

While preliminary, the results presented in Figure 3 are supportive of the predicted effects. Overall, causative launch showed the highest correspondence to the expected causal graph, with levels of lag and gap together determining a clear declination in its acceptance as a good descriptor of a given scene. The other conditions were not as clear, though analyses reveal interesting differences. The condition with the non-causative verb push gave rise to the least cause-based responses, whereas constructions with move and cause to move were roughly identical, and gave rise to a somewhat more cause-based pattern of responses.

LAG was a marginal predictor for all but (D) launch. Indeed, when both lag and gap are included in the analysis, almost all the predictive power (i.e., model fit) in (A-C) is done by GAP. To get a better sense of the data, we conducted analyses over just the trials where GAP was set to 0 to see if there were differences between the targets for LAG (as in Michotte’s original experiments). Indeed, LAG is then a significant predictor in some but not all conditions, see Table 1. Only (A), push, fails to show an effect of lag with zero gap.

**Discussion.** Our data raises two immediate questions, if the discussion in the preceding section is on the right track. The first is why we had such lower overall cause-based responses in (at least) condition (C) move. Manually inspecting the results by individual subjects, our
impression is that 2/7 subjects chose CAUSE in (C) whereas 5/7 did in (D) launch.\(^7\)

We think there is an explanation for this, which we are currently testing. In Michotte’s original experiments, a ratio of 3.6:1 for Red:Blue’s velocity was determined ideal for evoking the causal percept. However, in our experiments the velocity ratio was always 1:1. The strong

\(^7\)There were only 1/8 subjects that appeared to have chosen CAUSE with each of (A) push and (B) cause to move. This inspection was of course purely qualitative, however individual subject data really did seem to conform to either a CAUSE or a CONTACT pattern like those in Figure 2.
effect we found for (D) launch could thus be a product of the ratio: in normal situations, if something caused another thing to move by force, it is unlikely that the force of the causer’s movement would equal the force of movement evoked in the causee. In our scenes, there was always equal strength of movement, which is natural for the (presumably lexical) requirements of launch, but unnatural for move.

4. Conclusion

We presented data suggesting that English speakers’ behavior in a simple task is sensitive to subtle semantic features of the sentences used in the instructions. We gave subjects sentences that differed little in their surface meaning (move/push on the one hand, and push/launch on the other), yet which differ in their underlying syntax/logical form. The verbs move and launch participate in the syntactic causative-inchoative alternation, but push does not. Our results so far are novel in that we gave subjects nearly no explicit instruction. Construing logical forms as cognitive objects that are used to guide comprehension, we found preliminary support for the idea that linguistic analyses differentially, and correctly, predict behavior in an essentially non-linguistic task.

These data suggest intriguing possibilities for future work. “Syntactic bootstrapping” approaches posit that speakers (young and old) of a language have knowledge of systematic relationships between syntactic and semantic form, which they can use to restrict the range of possible meanings for a novel word (Gleitman 1990). Past work has shown (e.g. Naigles 1990; Fisher, Hall, Rakowitz, & Gleitman 1994; Yuan & Fisher 2009) that children can infer aspects of an event described by a novel verb based on the number of arguments it appears with, even on first exposure. However, it is not known whether children would make judgments corresponding to adults’ based on simple, Michotte-type stimuli after hearing sentences that do/do not support Cause.

The most pertinent future direction we foresee is in event segmentation. The logical forms discussed in this paper suggest that two event representations contribute to the interpretation of a causative sentence. Zacks, Tversky & Iyer (2001) asked participants to segment a continuous stream of complex action into either fine- or coarse-grained units, with the resultant pattern that fine-grained units nested within coarse-grained units. We hypothesize that, if a continuous stream of Michotte-type stimuli were created, different instructions could give rise to different segmentation patterns as a function of the instructions’ semantics. This segmentation methodology could similarly be extended to children, since we know infants as young as 6 months (Sharon & Wynn 1998) or 10 months (Baldwin, Baird, Saylor, & Clark 2001) are sensitive to the number of iterations in an event or to the boundaries of an event, respectively.

Finally, our results supported the particular interpretation we offered as to the nature of the interaction between language and mind. Core language includes a functional vocabulary
which must be related to content in extralinguistic cognition to be qualitatively comprehended. An expression of the language can only be evaluated for truth in case the predicates and relations comprising it find correlates outside of the language faculty. Some of these satisfiers may be perceptually-based, as in the data we report. The moral of this story, so far, is that the “new view”, this cognitive interpretation of logical form, has great potential for fruitful, future interdisciplinary work.

5. Appendix

Since subjects’ responses were binary, the influence of lag, gap and instruction on performance was analyzed by logistic regression using SPSS (SPSS Inc, Chicago, IL). In addition to lag and gap, instruction condition provided significant model improvement\(^8\) in predicting the likelihood of the data, \(\chi^2(3, N = 7500) = 33.462, p < .001\). Further, interaction terms for lag-by-condition and gap-by-condition were also significant based on model improvement, lag-by-condition: \(\chi^2(1, N = 7500) = 24.809, p < .001\); gap-by-condition: \(\chi^2(1, N = 7500) = 89.151, p < .001\). Together, these results warranted investigating the influence of lag and gap on the individual instruction conditions.

Table 2 summarizes the results of regressing subjects’ judgments on lag and gap for each of the four instruction conditions. In all conditions lag and gap were both significant predictors of response (i.e. both increased the likelihood of the data). However, as can be seen from the tests of model improvement, gap was a far better predictor than lag in conditions (A), (B) and (C), while lag and gap were similar in terms of producing model improvement for condition (D). Across all four instruction conditions, the negative coefficients show that increased gaps and lags made subjects less likely to respond that the instruction sentence was true of the scene–and this was especially true for gap.

Inspecting the distribution of response proportions in Figure 3, conditions (A), (B) and (C) have distributions somewhat similar to the CONTACT interpretation (especially (A)), while in (D) subjects clearly took the CAUSE interpretation. However, no condition was purely CONTACT. For example, in (A) lag was not a significant predictor when there was no gap, \(\chi^2(1, 400) = 3.27, p = .071\), in line with CONTACT, but lag was a significant predictor across all gap distances, \(\chi^2(1, N = 1600) = 62.99, p < .001\).\(^9\) Also, in (B) and (C), the slight decrease in proportion of responses with zero gap was significant, respectively: \(\chi^2(1, N = 400) = 7.98, p < .01\), \(\chi^2(1, N = 350) = 10.12, p < .01\). Further, the steep decrease in (D) under the no gap trials was highly significant \(\chi^2(1, N = 350) = 96.163, p < .0001\).

---

\(^8\)The chi-square model improvement is comparable to the \(F\)-change test in multiple linear regression; Pedhazur 1997.

\(^9\)That lag was a significant predictor does not appear to be a result in the difference in sample size. Restricting ourselves to the first gap of 0.5 cm, lag is still a significant predictor, \(\chi^2(1, N = 400) = 62.628, p < .001\).
Given the similarity of the response proportions for (B) and (C) (they are virtually indistinguishable in Figure 3), we tested whether they were significantly different. As expected,
condition as a predictor did not provide a significant model improvement after lag and gap for these two instruction condition, improvement: \( \chi^2(1, N = 3750) = 2.248, p = .134 \). This suggests (B) and (C) were comparable in terms of the interpretations they led subjects to adopt.