Causation in Semantics and Grammatical Structure
Week 9: Event structure and causal relationships

Prerna Nadathur

December 5, 2019

1 Background: event semantics

Davidson's idea:
Events are basic objects in our semantic system, just like individuals.
- this idea originates in an important 1967 paper, **"The Logical Form of Action Sentences"**

Davidson's original argument has to do with the way we deal with adverbial modifiers (adjuncts):

(1) Jones buttered the toast with a knife in the bathroom at midnight.

- this sentence has five participants:
  1. **a butterer:** Jones (AGENT)
  2. **a thing that gets buttered:** the toast (PATIENT)
  3. **a tool used to do the buttering:** a knife (INSTRUMENT)
  4. **a location for the buttering:** the bathroom (LOCATION)
  5. **a time for the buttering:** midnight (possibly: $\text{LOCATION}_{\text{time}}$)
- so to capture the meaning of (1), we might write something like:

  (2) BUTTER(J, the toast, knife, bathroom, midnight)

  – where there are certain specifications on the features of the relevant arguments – e.g., that the first argument is an individual, the fourth argument is a location, the fifth a temporal location, and so on.
- so: *butter*, based on (1), looks like a 5-place predicate

This starts off looking okay, but we quickly run into some problems:

- the sentences in (3) are also all grammatical uses of *butter*

(3) a. Jones buttered the toast with a knife in the bathroom.
b. Jones buttered the toast with a knife.
c. Jones buttered the toast.

- if *butter* is a 5-place predicate, then we shouldn't be able to decide if the sentences in (3) are true or false, since they don't have enough arguments
- given how we analyzed (1), we'll have to treat each of (3)a-(3)b as using a different verb
- so, there's a 4-place *butter'* that does not require a time argument, a 3-place *butter"* that doesn't require a time or a location, and so on:

(4) a. (3)a $\sim$ BUTTER$'(J$, the toast, a knife, the bathroom)
b. (3)b $\sim$ BUTTER$''(J$, the toast, a knife)
c. (3)c $\sim$ BUTTER$'''(J$, the toast)

- we need an individual lexical entry for each of these *butter*s, because the number of arguments a predicate take (and what roles these arguments play) is part of the lexical specification of a verb

- this is obviously silly: there's a tight connection between all of the uses of *butter*:
  - all of the lexical entries have agent and patient arguments, which stand in the same relationship to each other (one butters, the other gets buttered)
  - if (1) is true, then each of (3)a-(3)c must also be true

(5) Jones buttered the toast with a knife in the bathroom at midnight.
a. **entails:** Jones buttered the toast with a knife in the bathroom
b. **entails:** Jones buttered the toast with a knife.
c. **entails:** Jones buttered the toast.

- and, if we do things this way, we'll need even more lexical entries for *butter*, because the sentences in (6) are also fine!

(6) a. Jones buttered the toast with a knife at midnight.
b. Jones buttered the toast in the bathroom.
c. Jones buttered the toast at midnight.
d. Jones buttered the toast with a knife in the bathroom.

Question: we already have 4-place *butter'*, 3-place *butter"*, and so on! Why don't these entries work for the sentences in (6)?

Answer: the problem is that the lexical entry for a verb specifies the relationship of an argument to the main action

- a 3-place *butter"* that captures (3)b has three participants:
  1. agent: *Jones*
  2. patient: *the toast*
  3. instrument: *a knife*
- so, BUTTER$''(J$, the toast, midnight) would mean that Jones buttered the toast using midnight as a tool, which doesn't make any sense, and definitely is not what (6)c means

- there are systematic entailments between all of these uses of *butter*, but if each of the sentences in (1), (3) and (6) involves an independent lexical entry, there's no natural way to capture the relationships!
  - at best, we could say that the lexicon contains a set of *meaning postulates* (statements expressing logical relationships between lexical items), but we wouldn't be predicting the entailments
  - we'd have to state similar relationships for other verbs, like *spread*

    (7) Jones spread jam on the toast with a knife in the bathroom at midnight.
  - if we do this, we're saying that, in order to learn English properly, you also have to learn/memorize a HUGE set of statements like:

    (8) a. $\forall x, y, z, w, v[\text{BUTTER}(x, y, z, w, v) \rightarrow \text{BUTTER}'(x, y, z, w)]$
    b. $\forall x, y, z, w, v[\text{BUTTER}(x, y, z, w, v) \rightarrow \text{BUTTER}''(x, y, z)]$

    (and one more for each pair of *butter* sentences that has an entailment relationship)
  - but the entailments are very systematic: a *butter* sentence entails another *butter* sentence if the set of participants in the second sentence is a subset of the set of participants in the first
  - we want to capture this with a rule, rather than having to write (and, as language learners, somehow commit to memory) a long list of statements like (8)a,b.
  - the same rule should also tell us how to connect (7) to *Jones spread jam on the toast in the bathroom*, and so on

Logic to the rescue:

- Davidson points out that there's a basic rule of logic that does something similar:

  (9) $\forall p, q$ where p, q are propositions, $p \ \& \ q \vdash p$

- specifically, if we have a set of complete propositions strung together with ANDs, it will entail any expression which is same but missing one of the propositions

  (10) Jones danced and Rachel sang a song and Amy took a photo.
  - a. $\vdash$ Rachel sang a song and Amy took a photo.
  - b. $\vdash$ Jones danced and Rachel sang a song.
  - c. $\vdash$ Jones danced and Amy took a photo.
  - d. $\vdash$ Jones danced.
  - e. ...

- this looks a LOT like the kinds of relationships we're expressing above – in each case, we just take out on or more of the conjoined propositions

- so, if we can find a way to represent (1) so that each of the additional NPs shows up as a conjunction, we'll be able to predict the patterns

- to get this to work right, we need a basic set of facts about *butter*
  - there are two NPs in (1) that can't be dropped: the butterer (*Jones*) and the thing that gets buttered (*the toast*)
  - these are the **arguments** of *butter*: in order to return a complete sentence, *butter* needs to be combined with a butterer and a thing that gets buttered
  - all of the other NPs in (1) are optional (that's why all of the sentences in 3 and 6 are fine): PPs like *in the bathroom, at midnight* are **adjuncts**, not **arguments** of *butter*

We want something like this:

(1) Jones buttered the toast with a knife in the bathroom at midnight.
BUTTER(Jones, the toast) & It was with a knife & It was in the bathroom & It was at midnight

- what does the *it* in *it was with a knife* and *it was in the bathroom* refer to?

- not Jones or the toast, but in fact the whole situation involving Jones and the toast, and what Jones does to the toast – *it* refers to **an event!**

Davidson is proposing that **verbs describe events** – the participants in a buttering situation can then be connected to the event in different ways:

- let e stand for an event:

  (11) Jones buttered the toast with a knife in the bathroom
  $\exists e[\text{BUTTER}(J, \text{the toast}, e)$ & WITH(a knife, e) & IN(the bathroom, $e)]$

- we can connect the adjuncts to the main event using semantic roles:

(12) Jones buttered the toast with a knife in the bathroom
$\exists e$[BUTTER(J, the toast, e) & INSTRUMENT(a knife, e)
& LOCATION(the bathroom, e)]

- with events, we don't need more than one lexical entry for the verb *butter*:

(13) BUTTER(x, y, e) $\sim$ *There is an event e which is a buttering of y by x*

- and we can just use the logical rule in (8) to derive all of the entailments between sentences in (1) and (3)-(6)

- from the perspective of the language learner, this is a much more reasonable idea: you don't have to memorize a set of relationships between lexical entries – you just use a logical rule that you need for lots of other things in any case!

After Davidson (1967, 1969), a number of semantic researchers have made good arguments for treating the subject and object NPs in (1) as independent conjuncts as well:

- a **Neodavidsonian** representation:

(14) Jones buttered the toast with a knife in the bathroom.
$\exists e$[BUTTER(e) & AGENT(J, e) & PATIENT(the toast, e)
& INSTRUMENT(a knife,e) & LOCATION(the bathroom,e)]

(15) BUTTER(x, y, e) $\sim$ *There is an event e which is a buttering, and x is the agent of e and y is the patient of e*

- this is the common approach these days (we don't need to worry about the difference between (11) and (13) too much, for now)

This is a long lead-in to the connections between events and causation, but it's important to spend some time on because it's now standard to use events in semantic representations, and it's important to understand how and why we do this (especially if you are going to go on in semantics)!

2 Individuating events

Okay, we have events: now what?

- events are basic entities in the semantic computational system

- we can describe events in various ways, but we can't break down the notion of an event into concepts or objects that we already have in our semantics

- as you've probably seen in previous classes, we have a set of **basic types** that we work with in formal semantic representations

  – basic types are things like individuals and truth values

- other types are made up of functions from one type to another: a function from individuals to truth values is what we (so far) been treating as an intransitive verb
- now we have an additional basic type, the type of events
- given the suggested way of introducing events, intransitive verbs are now two-place predicates: they take one individual argument and one event argument, and return a truth value

- events are quite similar to individuals, but they're always going to be abstract objects

- even if we can't break an event down any farther, we'd like to get a better understanding of what we're doing by introducing events!
  - intuitively, we can recognize something as an event – Jones buttering the toast, for instance
  - and we can recognize that there are other ways of describing potentially the same event – e.g., Jones buttering the toast at midnight
  - "no entity without identity" – if events are real objects in our interpretation system, we need a way of determining when one event is the same as or different from another

Davidson's (1969) central question: what are we actually *doing* when we try to decide whether the events e and e'' in (16) are the same event or not?

(16) a. Jones buttered the toast.
$\sim$ There is an event e which is a buttering event and Jones is the agent of e and the toast is the patient of e

b. Jones buttered the toast at midnight.
$\sim$ There is an event e' which is a buttering event and Jones is the agent of e and the toast is the patient of e and e took place at midnight

- we know what it means for two sets A and B to be identical: we check if all of the things in A are also in B, and if all of the things in B are also in A

- how do we do this for events?

- looking at (16), one of the things we need to know is WHEN e is located
  - **proposed** criterion for event equivalence: two events e_1 and e_2 are equivalent only if they take place at the same time
  - as Davidson points out, this isn't enough to ensure two events are the same: if Jones butters the toast at midnight and the clock strikes at midnight, these two events take place at the same time, but they obviously aren't the same event!
  - **revised** criterion: two events e_1 and e_2 are equivalent only if they take place at the same time and the same place

- this helps a bit, but it STILL isn't good enough: suppose Jones butters the toast in the bathroom at midnight, and a clock which is also in the bathroom strikes at midnight – the buttering and the clock-striking are still not the same event!

- another thing we seem to need: events can only be the same if they have the same participants

- but again: suppose Jones butters the toast over a candle in the bathroom at midnight, so now he's heating the toast up as well

  - *Question:* Are the buttering events and heating events the same one? They take place at the same place and time, and they both involve Jones and the toast , but there's something funny about saying they are the same event!
  - More intuitions: it seems clear that there is an event that can be described by a sentence like (17)

    (17) Jones heated the toast while he buttered it in the bathroom at midnight.

  - the heating and the buttering are both PARTS of the event in (17), but they also lead independent lives!
  - after all, we could modify the buttering event by saying that Jones used a knife, and now the buttering has a participant which the heating does not have
  - but (17) is still a valid description of the whole scene!

Davidson's proposal: Two events are identical if they have the same causes and effects

(18) $e_1 = e_2$ if and only if $\forall z.(z \text{ caused } e_1 \leftrightarrow (z \text{ caused } e_2)$ AND $\forall z.(e_1 \text{ caused } z \leftrightarrow e_2 \text{ caused } z)$

Why causes and effects?

- cause and effect is a very natural way of relating one event to another – we talk about one event causing another all of the time!

  (19) Dr. Sheppard killed Roger Ackroyd.

  - part of our knowledge of the meaning of *kill* is that if x *kills* y, then x does something that results in y's death
  - in other words, x *kills* y if x participates in an event which causes an event in which y dies
  - suppose (20) is true:

    (20) Sheppard poisoned Roger Ackroyd.

- if the poisoning led to an event in which Roger Ackroyd died, then (19) will also be true
- but, crucially, the poisoning and the killing are NOT the same event – the poisoning causes Ackroyd's death, but the death is PART of the killing event

- the cause and effect condition actually captures the time, space, and participant conditions we suggested above
- it also captures something more – the idea that causation, like time and space, is a way that we experience and perceive the world
- so, implicit in Davidson's proposal for individuating events and deciding when they are or are not the same is the idea that causation is a basic relation in the world

3 Causal structure in and around events

Like DeLancey (1984), Croft (1991) is interested in a **theory of semantic roles,** which should do three things:

(i) provide a list of semantic roles that are not further analyzable
- meaning that, while we can describe the characteristics of a role like AGENT, we can't write a logical expression for AGENT (or PATIENT, or LOCATION, etc) in terms of other semantic objects

(ii) provide definitions for the semantic roles that are independent of the verbs that require them
- the definition of a particular role should not need to refer to a verb like *break* or *kill*
- for instance, we can't define an agent as a *butterer*, because this wouldn't capture the notion of agent across ALL verbs

(iii) provide a complete list of semantic roles, and one that is not too long
- if semantic roles are unanalyzable, then they are part of what we learn when we learn a(ny) language – so the list should be relatively small to be realistic (as well as theoretically useful)

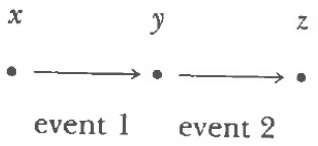
Croft's proposal: define semantic roles with respect to event structure

- the idea here is that, since semantic roles should not be defined with respect to specific verbs, they are selected for by certain parts of a verb's meaning
- these are the parts that occur across a range of different verbs
- if verbs describe **events**, then semantic roles are participants in events of different types

- following Davidson, the structure of events and the relationships between them are defined in terms of **causation**
- so, the consequence of Croft's idea is that the parts of a verb's meaning that define certain kinds of semantic roles have to do with causal structure

Croft develops a model or representation of causal structure that is inspired by **force dynamics** approaches (Talmy 1972, 1976)

- recall: Wolff (2003) also used a version of the force dynamics, in which a causal link represents some force or impulse transmitted from one object to another
- Croft does something extremely similar, except:
- these causal links are between objects (arguments), not events themselves – so he's representing the lexical semantics of verbs in terms of **internal** causal structure
- in other words, the causal links link participants in a verbal event



- in these diagrams, each arrow represents an event, bounded by two participants (represented by dots).
- a verb might describe a single arrow (one segment), or a chain of two or more arrows

Upshot: an event is a transmission of force from one individual to another

- for instance, *event 1* above represents a transmission of force from participant x to participant y, and *event 2* a transmission from y to z
- we could also think of an *event 3*, which contains both events 1 and 2, and so represents a transmission of force from x to z (passing through y)

(21) John hit the boulder with a hammer.
John hand hammer boulder
○ ⟶ ○ ⟶ ○ ⟶ ○

  - (21) describes a transmission of force from John to the boulder, by means of his hand, which wields a hammer
  - the diagram also 'orders' the participants in a causal chain, in terms of the process by which force is transmitted
  - John transmits a (mental) impulse to his hand, which then transmits force to the hammer (by swinging it), which transmits force to the boulder on impact

- thus, the overall event involves a transmission of John's mental impulse to the boulder by a direct causal chain
- ultimately, Croft wants to define and explain semantic roles in terms of this kind of ordering within a causal chain

- in general, it is possible to break events (or links in a causal chain) down into further subevents, depending on the context in which we are describing an event structure:

(22) a. John was sick.
b. The virus attacked John's throat, which became inflamed, resulting in laryngitis, until the immune system succeeded in destroying the infection.

- (22)a and (22)b describe the same event, but at different levels of granularity, which are appropriate for different contexts
- (22)a might be a suitable description for a friend
- (22)b might be a suitable description for a scientific or medical audience: in this case, the description matches a chain in which the virus, John's throat, and his immune system are represented as independent participants, and the 'atomic' events that are relevant involve the transmission of force from one of these participants to another
- if we are less interested in the technical details, it does not make sense to represent John's throat and/or John's immune system as participants in a causal chain independent of John
- the level of detail, and thus the 'size' of the subevents, is determined pragmatically

Not just anything can be an event, however: there are criteria for the 'causal chain' representations:

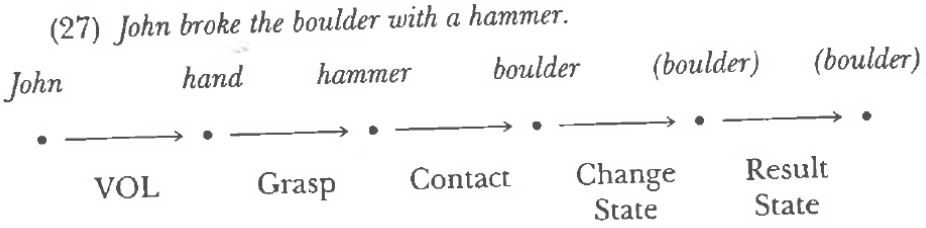
1. **Criterion 1.** An atomic event must be of only one causation type.
   - a causation *type* is determined by the *type* of force that gets transmitted – e.g., whether it is a physical force, or a mental impulse that gets turned into a physical force, and so on.
2. **Criterion 2.** An atomic event must be of a single inherent aspectual type, specifically a state or process.
   - recall Vendler's (1957) **aspectual classes** from Week 2 handout
   - different events have different sort of features – do they take time, do they look the same throughout the duration of the event, and so on

- the distinction Croft wants to make is between 'static' events (states, like "be red", which look the same at any point in time of a being-red event) and 'dynamic' events (processes, like *run*, activities, like *win*, and accomplishments, like *build a house*), which involve motion and/or change (something looks different at the beginning and end)

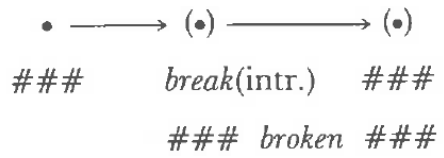
3. **Criterion 3.** An atomic event containing two participants must have those participants aligned in the direction of 'transmission of force'

   - if an event involves a transmission of force from, e.g., John to a boulder, then it can only be represented by a causal chain which puts John before the boulder, not after

4. **Criterion 4.** An atomic event must be a single qualitative unit.



- the first link in the chain for (27) is a **volitional** one, involving a transmission of mental force from John to his hand
  - other causation types: physical causation, affective causation, inducive causation
  - *physical causation:* one physical object acts on another – both *Grasp* and *Contact* links, above
  - *affective causation:* a physical object 'acts on' an object with mental states (for instance, if a tree *scares* John, thus producing a mental state in him)
  - *inducive causation:* a volitional entity acting on an object with mental states – events of convincing, persuading, maybe even forcing (without using physical force)
  - Croft argues that volitional causation is the 'least marked' (which aligns with DeLancey's ideas that the prototypical causal event involves a volitional agent acting on a non-volitional patient to change it in some way)'
  - affective causation is the most unusual
- aspectual class:
  - the grasp and contact events are processes: something changes in them over time. The result state event is different (and must thus be represented with its own link) – it involves the boulder being in a state of being broken, which does not change over time



  - this is consistent with the idea that we want to think of a *breaking* event as having two distinct parts (in the lexical decomposition) – a process, in which some physical force begins and ends (is transmitted), and then a result state, in which the patient object can be described as *broken*
  - the result state necessarily follows the causing process: the patient is not *broken* until the physical force has been transmitted

- Based on criterion 3:
  - we can define a PATIENT argument, then, as the 'right end' of a causal chain – it is the endpoint or final recipient of force transmitted through a chain
  - AGENTS are the left endpoints, where force originates

- Criterion 4 allows us to make distinctions between events that might be entangled in a particular scenario, but different in other:'
  - for instance, an event of a vase falling might also be part of an event in which a vase breaks
  - but since a falling event is not NECESSARILY connected to a breaking event (i.e. doesn't necessarily result in something being broken), we know that falling and breaking should be different segments in a causal chain
  - the same kind of distinction can be made between John's starting to sweat and John's running in an event where John's running causes him to sweat

What does this do for us?

- Croft's framework extends the idea that events are distinguished and determined in terms of causation

- now, we use the internal causal structure to define semantic roles – semantic roles (thematic roles, theta roles) are defined with respect to their position in a causal segment – this is the **causal order hypothesis**
  - as above, AGENTs are where forces originate (prototypically volitionally, but there are exceptions)
  - PATIENTs are at the receiving end
  - various other kinds of roles either fall in between agents and patients – for instance, instruments, which are used to transmit an agent's force to a patient
  - ...or subsequent to a patient – for instance, benefactives, which receive some benefit from the transmission of force to the patient

- Croft's definitions (pp.176–179):
  1. *Agent:* the initiator of an act of volitional causation (a VOL arc)
  2. *Patient:* the endpoint of an act of physical causation (a PHYS arc); acts of volitional causation must be mediated by a physical entity that physically acts on the patient (either some physical extension of the agent, or an instrument)
  3. *Experiencer:* the endpoint of an act of affective causation (an AFF arc)
  4. *Stimulus:* the initiator of an act of affective causation (an AFF arc)
  5. *Comitative:* an entity that participates in a causal chain at the same point and in the same role as the subject of the main verb (potentially also the initiator of a VOL arc)
  6. *Instrument:* an entity that is intermediate in a causal chain between the initiator and the affected entity (possibly restricted so that it is NOT the initiator of a VOL arc)
  7. *Benefactive/malefactive:* the endpoint of an action that causally follows the verbal causal segment. Usually an endpoint of an AFF arc
- Using the segment structure, we can also say more clearly what some of the basic causal entities are:
  1. *Cause:* an event (action or state) that causally immediately precedes the event sequence denoted by the main verb
     - so, if we also adopt Neeleman & van de Koot's (2012) ideas about a CCF – the CCF should be a key participant (probably an initiator) of a *cause* segment
  2. *Result:* an event (action or state) that causally immediately follows the event sequence denoted by the main verb
  3. *Purpose:* an event that is intended by an agentive initiator of the main verb causal segment to follow causally from the event denoted by the main verb causal segment

4 Additional comments

> "Events are identical if and only if they have exactly the same causes and effects. Events have a unique position in the framework of causal relations between events in somewhat the same way objects have a unique position in the spatial framework of objects." Davidson (1969, p.306)

- up to now, we've been trying to define causation (or make sense of its parameters) in terms of verbs and features of their thematic roles
- by moving to an event semantics picture, we've turned this around a bit: now we are defining events in terms of causation

- that is, we're now taking causation to be a basic way in which we experience the world, and we're arguing that the way we describe events (using verbs) is influenced by causal perception or causal cognition
- what we need now is a way of modeling HOW events can be connected
  - we saw the beginnings of this last week, with DeLancey's suggestion that some events are defined in terms of *sufficient* causes, and others by other kinds of causes
  - *next question:* what are the kinds of causes?
  - does everything fall into a causal chain as in Croft's examples?
  - or, does it make more sense to think of events as causally connected to one another in a network – then, different kinds of causal relationships reflect different configurations in a network
- this is the kind of framework we'll pursue in the second half of the course
  - we'll first review some different approaches and the ideas behind them
  - and then see how a particular (network) model of causation – or something like Wolff's (2003) force dynamics – can help us to analyze different kinds of causal language (and explain the differences between different kinds of causal words, and the reasons why we choose to use one causal word instead of another)
- **Important takeaways from this week:**
  - the ideas behind an event semantics framework (what motivates such a framework)
  - how event semantics and semantic roles are connected
  - the idea (an active area of research in semantics!) that events are defined and structured in terms of causation

More broadly:

- philosophical work on causation spends a lot of time trying to work out what causation actually *is*: what does it mean to be a cause, how we might define this in terms of logical relationships
- we're taking a different view, cognitively speaking (somewhat Kant-ian in approach):
  - we don't actually know what the real physical world is like, since we only have access to the way we perceive it and categorize it
  - causation is, basically, one of these ways of experiencing the world (like time and space)
  - so, as linguists, it's more reasonable to ask what causation *does* with respect to how we use language (instead of what it IS in reality)

- in other words, we don't need to find a way of defining causation if we can build a model that does certain things ...
- ... for instance, if the kinds of relationships that exist in our model contrast in ways that correspond to contrasts in language
  * since agentive and non-agentive causation are often marked differently, we'd like a model that reflects this as a structural difference
  * for Croft, this is the difference between a mental and a physical arc

5 References

1. Croft, W. 1991. *Syntactic Categories and Grammatical Relations.* Chicago: University of Chicago Press.
2. Davidson, D. 1967. The logical form of action sentences. In N. Rescher, ed., *The Logic of Decision and Action.* Pittsburg: University of Pittsburgh Press.
3. Davidson, D. 1969. The individuation of events. In N. Rescher, ed., *Essays in Honor of Carl G. Hempel.* Dordrecht: Springer.
4. Kearns, K. 2011. *Semantics.* London: Pallgrave Macmillan.