

A Linguistic Perspective on Causal Language

Michael Mohler, Mary Brunson, Sean Monahan, Marc Tomlinson

causality@languagecomputer.com

Language Computer Corporation

March 2018, rev. January 2023

Introduction

The nature of causality has been widely speculated about in the fields of philosophy and logic since at least the time of Aristotle and, since that time, there have been various attempts to categorize instances of real or hypothetical causality in order to better understand and model the complex causal mechanisms that underpin physical reality. Aristotle has organized causality based upon its material, its form, its changeableness, and its goal (*Phys.* II 3; Falcon, 2006); David Hume and John Stuart Mill proposed that a cause and its effect could be identified by priority, temporal contiguity (or a chain of succession), and regularity across multiple instances (Hume, 1739; Mill, 1843). Hans Reichenbach introduced a probabilistic model of causality (Reichenbach, 1956) which has since been expanded upon substantially in the Bayesian models of Judea Pearl and in his “Ladder of Causation” which distinguishes causation at the levels of seeing, doing, and imagining (Pearl, 2009; Bareinboim et al., 2022). Other recent work in telecoupling seeks to categorize causality amongst complex systems to identify proximate (immediate) and underlying (latent) causes (Carlson et al., 2018; Busck-Lumholt et al., 2022). John Mackie has leveraged the notion of sufficient and necessary conditions from the field of logic to define sufficient causes and necessary causes but argues instead that most of what we think of as causation is really “*an insufficient but non-redundant part of an unnecessary but sufficient condition*” (INUS causes; Mackie, 1965).

Our goal in this work is not to break new ground in the philosophical or logical field of causality but rather to investigate the usage of causal language in every day speech and writing with an eye to understanding how much (or how little) the speaker intends to convey regarding the relationship between an **antecedent** (an event or state which is logically or temporally prior) and a **subsequent** (an event or state which is logically or temporally after). To this end, we here define 23 types of causality (or semi-causality) organized hierarchically that represent distinct ways of discussing or reasoning over the (real or hypothetical) impact of an antecedent upon a subsequent. We further define a framework for conceptualizing causal transitivity among distinct chains of simpler causal relationships which we refer to as the “Transitive Causal Calculus” (TCC).

Types of Causality and Semi-Causality

For the purpose of this discussion, we will define **causality** broadly as including all cases where an asserted (real, hypothetical, or generic) antecedent has (or is believed to have) an effect on the likelihood, magnitude, or latency¹ of the subsequent. **Semi-causality**, by contrast, does not have an asserted antecedent with some effect on a subsequent, but does nonetheless frequently employ causal language to express the relationship between the two things. In Figure 1, we show a hierarchical representation of 23 types of causal and semi-causal relationships, arranged into four tiers with each tier representing finer-grained modeling of the causal relationship.

¹ That is, the time separating the antecedent from the subsequent.

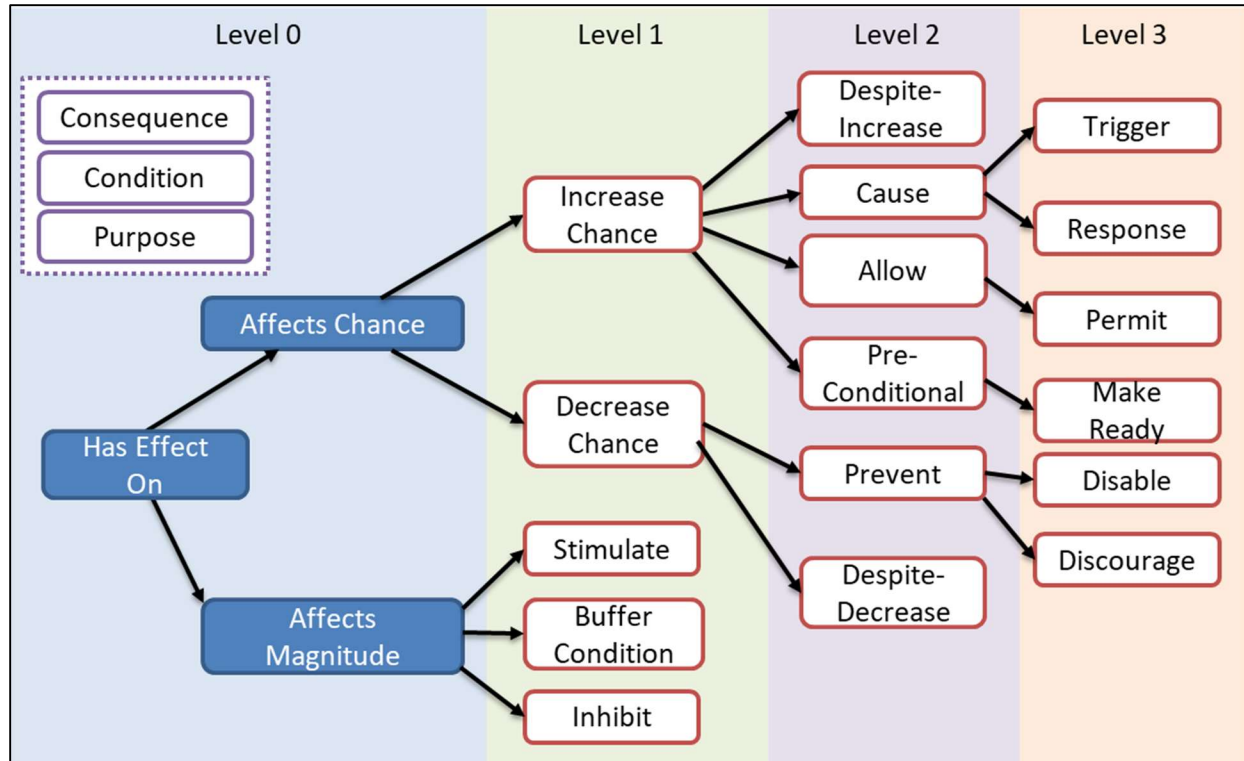


Figure 1: Hierarchical representation of the causal and semi-causal relations uncovered in this investigation. Note that semi-causal relations are shown in the dotted box in the top-left.

Our broad definition of causality above corresponds to the “HasEffectOn” relation, shown in Figure 1, which is subcategorized as either “AffectsChance” or “AffectsMagnitude”. The “AffectsChance” relation is more traditionally causal, insofar as the likelihood of the subsequent occurring (if an event) or holding (if a state) is altered based on the occurrence or holding of the antecedent. On the other hand, for “AffectsMagnitude” relations, the occurrence or holding of the subsequent is not dependent on the antecedent at all, but its magnitude or latency is significantly affected.

In the four tables below, we seek to define each of the remaining types of causality shown in Figure 1 above. These tables are organized to distinguish “Increase” relations (where the likelihood of the subsequent is increased), “Decrease” relations (where the likelihood of the subsequent is decreased), “Magnitude” relations (where the likelihood of the subsequent is fixed, but its magnitude/latency is affected), and “Semi-Causal” relations (where the antecedent does not truly affect the subsequent in a causal way).

In Table 1, we show nine types of causality which are alike insofar as they represent (real or believed) increases in the likelihood of a subsequent occurring or holding. They are distinguished based on the actuality² of the subsequent, the magnitude of the change in likelihood, the necessity and sufficiency of the antecedent, and the mediating presence of agency³ separating the antecedent from the subsequent.

Table 1. Examples of Increase Relations.

Relation	Definition	Example
Increase Chance	The likelihood or expectation of the subsequent occurring/holding is higher with the antecedent occurring/holding.	The stress of losing his job gave John an ulcer.
Despite Increase	The likelihood or expectation of the subsequent occurring/holding increased, but it did not actually occur/hold.	Even though I ran my hardest, I didn't win the race.
Pre-conditional	The antecedent provides a condition which is necessary for the subsequent to occur/hold.	Steve bought a plane ticket. Now , he can come visit us in Paris.
Make Ready	The antecedent provides a condition which is necessary for an agentive subsequent to occur and the resulting condition is sufficient for the agent to perform the action.	I put my number in Mary's phone, so she can call me now.
Cause	The increase in likelihood or expectation of the subsequent occurring is "significant" and non-preconditional.	Smoking is a leading cause of lung cancer in men over 30.
Response	The cause is mediated through an agent who decides (consciously or otherwise) to carry out the effect.	Rick punched Bill in the parking lot, and Bill punched him back.
Trigger	The effect occurs immediately upon the occurrence of the cause with no obvious mediating factors. This is similar to the notion of a 'proximate' cause.	The bomb struck a hospital, thereby killing 20 patients inside.
Allow	The antecedent removed an obstacle to the occurrence/ holding of the subsequent or failed to act to prevent the subsequent.	The guards didn't chase the escapee. They let him go.
Permit	The antecedent removed a legal or normative obstacle to the occurrence/holding of the subsequent providing explicit permission.	The king decreed that selling apples on Sunday is permissible .

² That is, whether or not the subsequent actually occurred or held. See Monahan and Brunson (2014).

³ Agency is here defined as the free and independent act of an agentive entity (e.g., human person, animal, deity) which is logically capable of choosing to perform the subsequent act or not.

In Table 2, we show five types of causality which are alike insofar as they represent (real or believed) decreases in the likelihood of a subsequent occurring or holding. They are distinguished based on the actuality of the subsequent, the change in likelihood, the necessity/sufficiency of the antecedent, and any agency separating the antecedent from the subsequent.

Table 2. Examples of Decrease Relations.

Relation	Definition	Example
Decrease Chance	The likelihood or expectation of the subsequent occurring/holding is lower with the antecedent occurring/holding.	There were no bombings in the city this year, due to increased police presence.
Despite Decrease	The likelihood or expectation of the subsequent occurring/holding decreased, but it actually did occur/hold.	Despite our best efforts, the bill was signed into law Tuesday morning.
Prevent	The decrease in likelihood or expectation of the subsequent occurring is “significant”.	Heavy traffic prevented me from making it into the office on time.
Discourage	The antecedent is mediated through an agent who decides (consciously or otherwise) not to carry out the subsequent.	The PETA ad campaign convinced Sarah not to eat meat.
Disable	The antecedent removes one or more necessary conditions to prevent the subsequent from occurring/holding.	My car was out of gas this morning, so it wouldn’t start.

In Table 3, we show three types of causality which are alike insofar as they represent (real or believed) effects on the magnitude or latency of the subsequent (but not its likelihood). They are distinguished largely based on the direction of the change.

Table 3. Examples of Magnitude Relations.

Relation	Definition	Example
Stimulate	The antecedent causes the subsequent to occur sooner or to a greater extent than it would otherwise.	The high cost of oil contributed to inflation this year.
Maintain (Buffer Condition)	The antecedent causes the subsequent to maintain its current status regarding its magnitude and latency.	Strict enforcement of law keeps the murder rate low.
Inhibit	The antecedent causes the subsequent to occur later or to a lesser extent than it would otherwise.	Refrigerating your vegetables keeps them from going bad too soon.

Table 4 shows the final three categories of relations which we have observed employing causal language, but which are not truly causal in the same sense. These Semi-Causal categories are included here due to their commonality and their ability to confound straightforward causal analysis from text.

Table 4. Examples of Semi-Causal Relations.

Relation	Definition	Example
Consequence	The subsequent is a logical consequence of the antecedent OR the relationship is purely definitional.	This figure is an octagon because it has eight sides.
Purpose	The subsequent is a goal, while the antecedent is performed by an agent in order to increase the chance of the goal being realized.	I bought a fancy car to impress my coworkers.
Conditional	The antecedent is expressed as something which has not (necessarily) occurred and the subsequent is left uncertain.	If you invite Alex to the party, he will bring snacks.

Figure 2 below shows a flowchart which has been used at Language Computer to annotate causality in text documents. It assumes the annotator has encountered explicit causal language (or language where causality can be inferred). The flowchart begins by determining whether the relationship meets our criteria for causality or semi-causality before attempting to place it deeper within the hierarchy of Figure 1.

Antecedent: The event/state (positive or negated) which occurs first, either logically or temporally.
Subsequent: The event/state (positive or negated) which occurs second, either logically or temporally.
Latency: The amount of time between the antecedent and the subsequent.

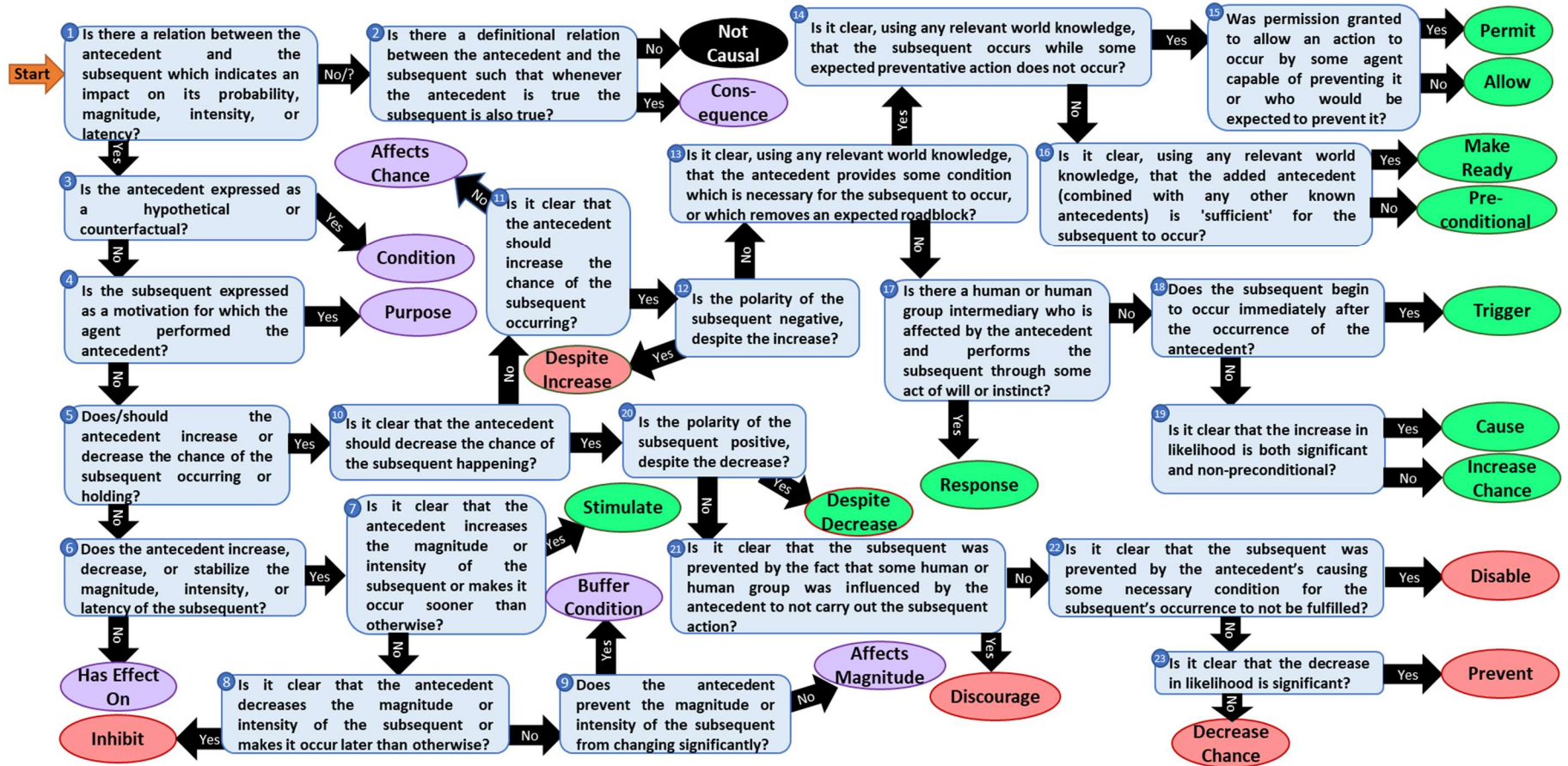


Figure 2: Flowchart for determining the type of Causality (or Semi-Causality) being expressed in a text.

Transitive Causal Calculus

In addition to the hierarchy of causality and semi-causality defined above, we here also define a methodology for reasoning over chains of causal relations in a pragmatic way. The goal is to provide a general sense of how people perceive and describe more remote transitive relationships between an antecedent and a subsequent. An overview of this Transitive Causal Calculus (TCC) is shown in Figure 3 below.

Chains	Closure
Cause(A,B) + Cause(B, C)	Cause(A,C)
Chains	Closure
Cause(A,B) + Prevent(B,C)	Prevent(A,C)
Prevent(A,B) + Cause(B,C)	Prevent(A,C)
Chains	Closure
* Enable(A,B) + Cause(B,C)	Enable(A,C)
Cause(A,B) + Enable(B,C)	Enable(A,C)
Prevent(A,B) + Prevent(B,C)	Enable(A,C)
Chains	Closure
† CEIC(A,B) + Inhibit(B,C)	Inhibit(A,C)
Stimulate(A,B) + Inhibit(B,C)	Inhibit(A,C)
Inhibit(A,B) + Stimulate(B,C)	Inhibit(A,C)
Chains	Closure
CEIC(A,B) + Stimulate(B,C)	Stimulate(A,C)
Stimulate(A,B) + Stimulate(B,C)	Stimulate(A,C)
Inhibit(A,B) + Inhibit(B,C)	Stimulate(A,C)

Chains	Closure
Enable(A,B) + Enable(B,C)	IncreaseChance(A,C)
IncreaseChance(A,B) + CEIC(B,C)	IncreaseChance(A,C)
CEIC(A,B) + IncreaseChance(B,C)	IncreaseChance(A,C)
DecreaseChance(A,B) + Prevent(B,C)	IncreaseChance(A,C)
Prevent(A,B) + DecreaseChance (B,C)	IncreaseChance(A,C)
Inhibit(A,B) + Prevent(B,C)	IncreaseChance(A,C)
Stimulate(A,B) + CEIC(B,C)	IncreaseChance(A,C)
Chains	Closure
Enable(A,B) + Prevent(B,C)	DecreaseChance(A,C)
Prevent(A,B) + Enable(B,C)	DecreaseChance(A,C)
DecreaseChance(A,B) + CEIC(B,C)	DecreaseChance(A,C)
CEIC(A,B) + DecreaseChance (B,C)	DecreaseChance(A,C)
Prevent(A,B) + IncreaseChance (B,C)	DecreaseChance(A,C)
IncreaseChance(A,B) + Prevent(B,C)	DecreaseChance(A,C)
Inhibit(A,B) + CEIC(B,C)	DecreaseChance(A,C)
Stimulate(A,B) + Prevent(B,C)	DecreaseChance(A,C)
Chains	Closure
Prevent(A,B) + Stimulate(B,C)	BufferCondition(A,C)
Prevent(A,B) + Inhibit(B,C)	BufferCondition(A,C)

*Enable = Precondition or Allow

† CEIC = Cause, Enable, or IncreaseChance

Figure 3: The Transitive Causal Calculus

In general, an antecedent (A) causing an intermediary (B) which has a relation (R) with (C) suggests that A has relation R with C as well. When antecedent (A) prevents an intermediary (B) which would have had a relation (R) with (C) the relationship is more complex, requiring the polarity of R to flip and often generalize up the causal hierarchy. For instance, if (A) prevents (B) and (B) would have prevented (C), then we can say that (A) enabled (C) to take place, because B was not available to prevent it.

Conclusion

In this white paper, we have defined a hierarchy of causal and semi-causal relations which are frequently observed in text. While natural language can be used to express a wide variety of causal relationships, speakers typically convey only a general sense of the antecedent's effect on the likelihood, magnitude, or latency of a subsequent event or state. In addition to defining and distinguishing this hierarchy of relations, we have created a diagnostic methodology to enable

their consistent annotation in English text. Furthermore, we have defined a calculus which models the ways in which causal relations can chain to provide wider-ranging causal effects.

References

- Bareinboim, E., Correa, J. D., Ibeling, D., & Icard, T. (2022). On Pearl's hierarchy and the foundations of causal inference. In *Probabilistic and Causal Inference: The Works of Judea Pearl* (pp. 507-556)
- Busck-Lumholt, L. M., Coenen, J., Persson, J., Frohn Pedersen, A., Mertz, O., & Corbera, E. (2022). Telecoupling as a framework to support a more nuanced understanding of causality in land system science. *Journal of Land Use Science*, 17(1), 386-406.
- Carlson, A. K., Zaehring, J. G., Garrett, R. D., Felipe Bicudo Silva, R., Furumo, P. R., Raya Rey, A. N., ... & Liu, J. (2018). Toward rigorous telecoupling causal attribution: A systematic review and typology. *Sustainability*, 10(12), 4426.
- Falcon, A. (2019). Aristotle on Causality. *Stanford Encyclopedia of Philosophy*, 1-26.
- Hume, David. 1739/1978. *A Treatise of Human Nature*. Oxford: Clarendon Press. Edited by L. A. Selby-Bigge & P. H. Nidditch.
- Mackie, J. L. (1965). Causes and conditions. *American philosophical quarterly*, 2(4), 245-264.
- Mill, John S. 1843/1911. *A System of Logic: Ratiocinative and Inductive*. London: Longmans, Green & Co.
- Monahan, S., & Brunson, M. (2014). Qualities of eventiveness. In *Proceedings of the Second Workshop on EVENTS: Definition, Detection, Coreference, and Representation* (pp. 59-67).
- Pearl, J. (2009). *Causality*. Cambridge University Press.
- Reichenbach, Hans. 1956/1971. *The Direction of Time*. Berkeley: University of California Press.