How do we reason about the things that go unsaid? Alternatives in English Causative Constructions

Sohini Kar
Massachusetts Institute of Technology
Cambridge, MA
skar@mit.edu

Jennifer Hu **(TA)**
Massachusetts Institute of Technology
Cambridge, MA
jennhu@mit.edu

Abstract

As speakers face various choices in their chosen utterances representing various events and world states, these alternatives enable listeners to draw pragmatic inferences in language understanding. We found that the proportion of people who inferred directness from a verb with no lexical causative form used in the periphrastic form was greater than the proportion who infer directness from a verb with a lexical causative form. We modeled our results using the Rational Speech Act model.

Keywords: Rational Speech Act model, Bayesian inference, language, causatives, Amazon Mechanical Turk

Introduction

Causatives and Constructions

Events may be described in a variety of linguistic constructions, which each convey different meanings and information. Three specific constructions include the intransitive, the periphrastic causative, and the lexical causative. Taking an example of someone breaking a vase, we can create the above constructions as follows: "the vase broke" (intransitive), "he made the vase break" (periphrastic causative), and "he broke the vase" (lexical causative). The lexical causative (LC) has been observed to be preferred over the periphrastic causative (PC) when attributing causation to the agent, unless there was a deliberate action taken. (Wolff, 2003; Fodor, 1970).

Studying the differences in interpretations based on the chosen language construction allows for a greater understanding of how people reason about everyday events using the language and utterances presented to them. Understanding the distinction and why people may prefer one form over the other allows for further questions on people's priors and assumptions for various world states. However, the problem is not quite easily solved, as not all verbs have an LC construction, thus creating two cases in how a construction may be interpreted. First, a verb may be +LC, in which case if the PC is used the listener may infer a lower level of direct causation. For example, if "he broke the vase" (LC) were used instead of "he made the vase break" (PC) for the +LC verb "to break" the listener may attribute causation to the agent. Alternatively, for an -LC verb, listeners will not draw strong inferences about the directness of causation. An example of this would be the verb "to fall" where the PC construction "he made the vase fall" is possible, but the LC construction "he fell the vase" is ungrammatical. In this case, use of the PC construction would not encourage inferences for causation.

This encourages the following relationship:

$$P(\text{direct} \mid \text{Periphrastic}, +LC) < P(\text{direct} \mid \text{Periphrastic}, -LC)$$

However, this equation does not take into account possible likelihoods *a priori*. Specifically, some events may be more or less likely to occur *a priori.* A vase breaking or falling would be less likely to be a direct action taken by an agent, but rather an accident.

To account for these prior likelihoods and situations, the intransitive construction is considered as a baseline control as it provides no explicit information about causation and the agent. Using this information, we may construct a 2×2 experiment design, including sample sentences.

Construction // Verb Class	+Lexical Causative (+*LC*)	−Lexical Causative (−*LC*)
Intransitive (Baseline)	The vase **broke**	The rabbit **disappeared**
Periphrastic	She made the vase **break**	She made the rabbit **disappear**

Rational Speech Act

The Rational Speech Act (RSA) uses Bayesian models to formalize inferences drawn from meanings in context and causative constructions (Goodman and Frank, 2016). RSA models provide a computational framework to linguistic structure, using the idea that a rational speaker would choose utterances to be helpful to the listener. Using Bayesian inference, a listener may update their beliefs about a speaker's utterance using a set of assumptions, including that the speaker aims to maximize usefulness while minimizing language production.

The equations used in the RSA model uses sets of utterances U and world states W. We also define a literal listener L_0, a pragmatic speaker S_1, and a pragmatic listener L_1. The probability of each choosing or interpreting an utterance from the world state is given by the following equations:

$$P_{L_0}(w \mid u) \propto [\![u]\!](w) \cdot P(w)$$
$$P_{S_1}(u \mid w) \propto \exp(\alpha \cdot U_{S_1}(u \mid w))$$
$$P_{L_1}(w \mid u) \propto P_{S_1}(u \mid w) \cdot P(w)$$

Here, $[\![u]\!](w)$ defines if the utterance u is true (1) or false (0) when applied to that world state w. The utility function $U_{S_1}(u \mid w)$ gives the utterance cost of using the utterance given the world state, and uses costs associated with each utterance. Using Bayesian decision-making, we may find the speaker's rationality and the listener rationality with the RSA equations defined above.

In a basic RSA model, we may calculate how the speaker chooses utterances to provide unsaid information to the listener as well as the listener's rationality in understanding that meaning and inferring a world state. Additional layers of complexity may be provided through variables such as the cost function and the utility function, the alpha proportion, and the basic probability of each world state. These each allow for information to be injected into the model based on English grammar as well as the prior likelihoods discussed previously. For example, the utility function has been tested with more complex variations to model speaker choices, and the world state likelihoods are highly variable as well.

Procedure and Paper Overview

Implement Bayesian Model

We will develop a simple Python-based model of RSA, and collect human interpretations. We will then modify the model in order to better fit the human data.

To develop the model, first we will denote the lexicon as well as the notations for each world state. This sets the original sets of utterances and corresponding world states, and provides us with our main reference for the relationships.

The world states will be given as follows:

$W = \{B_0$: break - no agent, B_i : break - indirect agent, B_d : break - direct agent, V_0 : vanish - no agent, V_i : vanish - indirect agent, V_d : vanish - direct agent$\}$

$\mathcal{L}(m,u)$	$B_\emptyset$	B_i	B_d	$V_\emptyset$	V_i	V_d
Y broke	1	1	1	0	0	0
X made Y break	0	1	1	0	0	0
X broke Y	0	1	1	0	0	0
Y vanished	0	0	0	1	1	1
X made Y vanish	0	0	0	0	1	1

Table 1: Lexicon

We then define the denotation function $[\![u]\!](w)$. This indicates if a world state is true (1) or false (0), and it is primarily used by the literal listener in interpreting the utterance. The corresponding table for $[\![u]\!](w)$ to our example lexicon is provided in Table 2.

$\mathcal{L}(m,u)$	$B_\emptyset$	B_i	B_d	$V_\emptyset$	V_i	V_d
Y broke	1	1	1	0	0	0
X made Y break	0	1	1	0	0	0
X broke Y	0	1	1	0	0	0
Y vanished	0	0	0	1	1	1
X made Y vanish	0	0	0	0	1	1

Table 2: Denotation Function

We will now create the cost for each utterance compared to the world states. This is using in the utility function $U_{S_1}(u \mid w)$ in creating the utterance cost associated with each utterance compared to the world state. For example, the cost of saying "X made Y vanish" when the world state is B_i should be very high, as the two are unrelated. We initialize the model with a basic set of costs, provided in Table 3.

$C(m,u)$	$B_\emptyset$	B_i	B_d	$V_\emptyset$	V_i	V_d
Y broke	0	0.5	0.5	∞	∞	∞
X made Y break	∞	1	1	∞	∞	∞
X broke Y	∞	1	0	∞	∞	∞
Y vanished	∞	∞	∞	0	0.5	0.5
X made Y vanish	∞	∞	∞	∞	1	1

Table 3: Cost Function

Finally, we will initialize a set of priors associated with each world state. To begin, we will allow each world state to have the same probability, but we will fine-tune this according to how likely each world state is. Because we start with 6 world states, we initialize the world states as:

$$P(w) = [\frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}]$$

After implementing this base set of definitions, we can then implement the equations outlined in the Introduction using Python. As we collect data, we will fine-tune parameters such as $P(w)$ and α.

Run Pilot Study

The next step is to create a pilot study for collecting human inferences about the same lexicon as implementing using RSA. We will create a set of stimuli testing listener inferences about a variety of utterances, and ask them to select the corresponding world state.

Figure 1: Example images for “she made the signature disappear” with the images shown in order of Baseline, Direct, and Indirect.

An example stimuli is shown in Figure 1, for the statement “she made the signature disappear.” The subject will be asked to select the corresponding cartoon representing the statement. Each image corresponds to a different world state, and the prompt given will correspond to an utterance. Using these, we may recreate the human version of the interpretations for each utterance-world state pair.

In order to display and test the subjects, we will use Amazon Mechanical Turk, which is a crowdsourcing website for businesses to hire remotely located "crowdworkers" to perform discrete on-demand tasks that computers are currently unable to do.

For Mechanical Turk, we will specifically use the psiTurk framework, which aids in collecting behavioral data. Upon opening the job, subjects will be able to begin viewing the prompts and selecting their cartoon choice.

Once we collect the data, we will create two table displaying the proportions found based on the verb type, representing the the human interpretations and corresponding probabilities that difference constructions are chosen based on the displayed cartoon, or world state. We provide an example in Table 4. Each cell represents the proportion of subjects who chosen a world state given by the cartoon, based on the provided utterance given by the prompt.

$P(w\,|\,u, \text{LC})$	Base	Direct	Indirect
periphrastic			
intransitive			

Table 4: Human data representation

Paper Overview

After collecting the human data and implementing the RSA model in Python, we will then analyze the data collected to observe trends, as well fine-tune the parameters of the RSA model to better fit the human model and test how the model compares to human behavior.

Our main prediction is that the perceived directness in causation when the periphrastic is used will be greater for -LC verbs than +LC verbs. This is illustrated as:

$$P(\text{direct}\,|\,\text{Periphrastic}, +LC) < P(\text{direct}\,|\,\text{Periphrastic}, -LC)$$

In the next sections, we will present the data and our corresponding analysis of trends and how they line up with our predictions. We will then present the model results after fine-tuning the parameters to fit the human data.

Data

The data we collected from a sample of 9 participants is presented below.

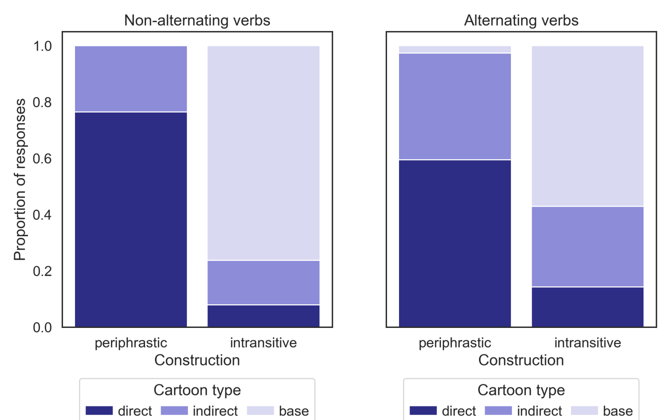

Figure 2: Proportions for n=9 subjects and their choices of cartoon based on presented construction, divided by the verb type

$P(w\,|\,u, \text{-LC})$	Base	Indirect	Direct
intransitive	0.67	0.22	0.11
periphrastic	0	0.22	0.78

Table 5: Approximate proportions of cartoon type chosen based on construction for non-alternating verbs

$P(w\mid u,$+LC)	Base	Indirect	Direct
intransitive	0.5	0.3	0.2
periphrastic	0.05	0.35	0.6

Table 6: Approximate proportions of cartoon type chosen based on construction for alternating verbs

In Table 5, the average proportion of subjects who chose either the base, direct, or indirect cartoon is chosen for the periphrastic and intransitive constructions, using non-alternating -LC verbs. Table 6 displays similar data, but for alternating, +LC verbs. Each cell may be read as $P(w\mid u, \mathrm{LC})$, where the world state (cartoon type chosen) w is the column and the utterance (prompt/construction presented) u is the row.

Analysis

We see that our hypothesis holds true., since $P(\text{direct}\mid\text{Periphrastic}, +LC) = 0.6$ was less than $P(\text{direct}\mid\text{Periphrastic}, -LC) = 0.78$. Our initial hypothesis was:

$$P(\text{direct}\mid\text{Periphrastic}, +LC) < P(\text{direct}\mid\text{Periphrastic}, -LC)$$

This would indicate that people indeed infer additional information from the choice of verb used in presented language constructions. The proportion of people who inferred directness from a -LC verb used in the periphrastic form was greater than the proportion who infer directness from a +LC verb.

This means that more people would interpret "she made the signature disappear" as more direct causation than "the signature disappeared," when compared with "she made the vase break" and "the vase broke," because "to break" is a +LC verb and there is an alternate form in "she broke the vase." Thus, the hypothesis holds that there is additional information inferred from constructions chosen when the speaker is assumed to be rational (Katzir, 2007; Fox and Katzir, 2011).

We may now recreate our results into a format mirroring the previous lexicon from Table 1.

$P(w\mid u,$+LC)	B_0	B_i	B_d	V_0	V_i	V_d
Y broke	0.5	0.3	0.2	-	-	-
X made Y break	0.05	0.35	0.6	-	-	-
X broke Y	-	-	-	-	-	-
Y disappeared	-	-	-	0.67	0.22	0.11
X made Y disappear	-	-	-	0	0.22	0.78

Table 7: Combined data table, mirroring the lexicon's form. "X broke Y" was untested

Next, we will fit our basic RSA model to the identified human data. Assuming each world state priors are equal and $\alpha = 1$, we find the following results:

$P(w\mid u,$+LC)	B_0	B_i	B_d	V_0	V_i	V_d
Y broke	0.7	0.2	0.2	-	-	-
X made Y break	NA	0.19	0.19	-	-	-
X broke Y	-	0.19	0.23	-	-	-
Y disappeared	-	-	-	0.7	0.32	0.32
X made Y disappear	-	-	-	NA	0.29	0.29

Table 8: Model results with $P(w) = [\frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}]$ and $\alpha = 1$

Decreasing α in increments of 0.1, we find that $\alpha = 0.2$ seems to match the data closest to Table 7. This variable is primarily used for adjusting the data, and does not have a representative value.

For the world states, we consider the priors the subjects may be using. Often, people assume that if someone acts upon an object, it is in a causative, direct manner rather than accidental. If it is not causative and direct, usually it is thought of as a base world state rather than indirect. We can model this set of priors with direct > base > indirect. Specifically, we may adjust the model:

$$P(w) = [\frac{2}{12}, \frac{1}{12}, \frac{3}{12}, \frac{2}{12}, \frac{1}{12}, \frac{3}{12}].$$

By adjusting these priors, we are placing assumptions on the listeners' interpretations of possible world states and their likelihoods, as well as how those predictions are used in selecting a cartoon representing those world states.

$P(w\mid u,$+LC)	B_0	B_i	B_d	V_0	V_i	V_d
Y broke	0.56	0.28	0.32	-	-	-
X made Y break	NA	0.3	0.7	-	-	-
X broke Y	-	0.3	0.8	-	-	-
Y disappeared	-	-	-	0.56	0.27	0.34
X made Y disappear	-	-	-	NA	0.25	0.8

Table 9: Model results with $P(w) = [\frac{2}{12}, \frac{1}{12}, \frac{3}{12}, \frac{2}{12}, \frac{1}{12}, \frac{3}{12}]$ and $\alpha = 0.2$

We were able to adjust our RSA model to mirror the human data quite closely, mirroring our hypothesis that $P(\text{direct} \mid \text{Periphrastic}, +LC) = 0.7$ was less than $P(\text{direct} \mid \text{Periphrastic}, -LC) = 0.8$. However, the delta is significantly less than that observed in the original set of data, which is a major area of improvement that may be resolved with further research into the implications and specifics of the cost function.

One trend we were not able to replicate was $B_d < B_i$ for both +LC and -LC verbs in the transitive form. Additionally, there was some proportion of subjects who chose $P(B_0 \mid \text{periphrastic}, \text{+LC})$, but we had set $L(B_0, \text{periphrastic})$ to 0 in the denotation function, meaning it would not be possible. These outliers and their reasoning is another area for potential research.

Conclusion

Our hypothesis held true, that the proportion of people who inferred directness from a -LC verb used in the periphrastic form was greater than the proportion who infer directness from a +LC verb. Additionally, we were able to replicate the RSA model and fit it to the data found, supporting the predicted pattern of behavior.

This would support the theory that a rational listener interprets constructions of the periphrastic causative depending on the existence of alternate, lexical causative forms of the verbs used by a predicted rational speaker.

Further research includes considering the cost function and priors on the world states, and how to adjust those to better fit the model. Additionally, there were some outliers discarded by the denotation function, such as $P(B_0 \mid \text{periphrastic}, \text{+LC})$, so further work may be done in representing those interpretations and proportions.

Acknowledgments

Thank you to Jennifer Hu for the amazing guidance and resources in completing this paper and learning about the intersection of language and Bayesian inference. Additionally, thank you to the 6.804 staff for organizing this class and final project.

References

J. A. Fodor. *Three reasons for not deriving "kill" from "cause to die"*. Linguistic Inquiry, 1:429–438, 1970.

D. Fox and R. Katzir. *On the characterization of alternatives.* Natural Language Semantics, 19(1):87–107, 2011. doi: 10.1007/s11050-010-9065-3. URL https://doi.org/10.1007/ s11050-010-9065-3.

R. Katzir. *Structurally-defined alternatives.* Linguistics and Philosophy, 30(6):669–690, 2007. doi: 10.1007/ s10988-008-9029-y. URL https://doi.org/10.1007/ s10988-008-9029-y.

N. D. Goodman, M. C. Frank. *Pragmatic Language Interpretation as Probabilistic Inference.* Trends in Cognition, 20(11):818-829, 2016. doi: 10.1016/ s0010-0277(03)00004-0.

P. Wolff. *Direct causation in the linguistic coding and individuation of causal events.* Cognition, 88(1):1–48, 2003. doi: 10.1016/s0010-0277(03)00004-0.