Learning reference biases from language input

Toth, Abigail Grace; Taatgen, Niels; Hendriks, Petra; van Rij, Jacolien

Published in:
Proceedings of the 19th International Conference on Cognitive Modelling (ICCM 2021)

Learning Reference Biases from Language Input: A Cognitive Modelling Approach

Abigail Toth (a.g.toth@rug.nl)
Department of Artificial Intelligence, University of Groningen
Nijenborgh 9, 9747 AG Groningen, The Netherlands

Niels Taatgen (n.a.taatgen@rug.nl)
Department of Artificial Intelligence, University of Groningen
Nijenborgh 9, 9747 AG Groningen, The Netherlands

Petra Hendriks (p.hendriks@rug.nl)
Department of Linguistics, University of Groningen
Jatstraat 26, 9712 EK Groningen, The Netherlands

Jacolien van Rij (j.c.van.rij@rug.nl)
Department of Artificial Intelligence, University of Groningen
Nijenborgh 9, 9747 AG Groningen, The Netherlands

Abstract

In order to gain insight into how people acquire certain reference biases in language and how those biases influence online language processing, we constructed a cognitive model and presented it with a dataset containing reference asymmetries. Via prediction and reinforcement learning, the model was able to pick up on the asymmetries in the input. The model predictions have implications for various accounts of reference processing and demonstrate that seemingly complex behavior can be explained by simple learning mechanisms.

Keywords: implicit causality; reference; cognitive modelling

Introduction

If you congratulate someone, most often it is because of something they did, whereas if you apologize to someone, most often it is because of something you did. This preference to attribute the cause of an event to a particular entity is known as the *implicit causality* (IC) *bias* and as such, verbs like *apologize* and *congratulate* are known as IC verbs. These verbs can be further separated depending on whether causality is attributed to the grammatical subject or object. For example, when asked, participants consistently attribute the cause in (1a) to Kaitlyn and the cause in (1b) to Marie (e.g., Brown & Fish, 1983; Rudolph & Forsterling, 1997).

(1) a. Kaitlyn angered Marie.
b. Kaitlyn comforted Marie.

Implicit causality has mainly been applied to investigate reference processing, in particular pronoun resolution (e.g., Garnham, Traxler, Oakhill, & Gernsbacher, 1996; Järvikivi, Van Gompel, Hyönä, & Bertram, 2005; Koornneef & Van Berkum, 2006). In a self-paced reading study Koornneef and Van Berkum (2006) had participants read sentences like those in (2) and found significantly slower reading times for sentences like (2b), where the pronoun was inconsistent with the bias set-up by the verb, compared to sentences like (2a), where the pronoun was consistent with the bias set-up by the verb. Furthermore, the effect was significant immediately following the pronoun, suggesting that IC information is used proactively, influencing comprehenders' expectations about subsequent reference.

(2) a. Linda praised David because <u>he</u> had been able to complete the difficult assignment with very little help.
b. David praised Linda because <u>he</u> had been able to complete the difficult assignment with her help only.

Other studies have investigated implicit causality in cases where the pronoun cannot ultimately be disambiguated by gender information. When participants are presented with sentences like those in (3), continuations for (3a) are predominantly about Molly, the subject, whereas continuations for (3b) are more evenly distributed between the subject and object referents (e.g., Kehler, Kertz, Rohde, & Elman, 2008; Stevenson, Crawley, & Kleinman, 1994). This has been taken as evidence that IC can modulate well established structural biases, such as the *first mention* and/or *subject bias*, which reflect the typical pattern of people interpreting ambiguous pronouns as referring back to first-mentioned and/or subject referent, which in English are most often confounded (e.g., Gernsbacher, 1989; Järvikivi et al., 2005). Eye-tracking studies have also shown that implicit causality affects the online processing of pronouns (e.g., Järvikivi, Van Gompel, & Hyönä, 2017).

(3) a. Molly apologized to Sophie. She ______.
b. Molly congratulated Sophie. She ______.

It is not exactly clear how knowledge about certain interpersonal exchanges, like *congratulating* and *apologizing*, ends up influencing language processing. One possibility is that when language users encounter an IC verb (and are not already privy to the causal information), they naturally generate an expectation about which referent will be referred to subsequently. In the case of subject-biased IC verbs, this expectation manifests as a continuation, such that the listener expects to continue to hear about the subject referent. In the case of object-biased IC verbs, the expectation instead manifests as a shift, such that the listener expects attention to shift from the (grammatically prominent) subject to the (less grammatically prominent) object. These expectations, about which referent will be referred to subsequently, can in turn interact with expectations about the form of the reference (i.e., whether a name or a pronoun will be used). For instance, there is evidence that subject pronouns like *she* are more likely to be used to refer back to preceding subjects, whereas names are more likely to be used to refer back to preceding objects (e.g., Arnold, 1998; Kehler et al., 2008; Stevenson et al., 1994). Thus, in the case of subject-biased IC verbs, listeners may not only expect to hear about the subject referent, but also that a pronoun will be used.

Not only is it unclear how knowledge of implicit causality ends up influencing language processing, it is also unclear how such knowledge is acquired. One possibility is that over time language learners pick up on asymmetries present in the input (and then make use of this knowledge when processing language). Unfortunately it is quite difficult to assess the relative frequency of the different reference possibilities in one's actual language input (see Sukthanker, Poria, Cambria, & Thirunavukarasu, 2020). To our knowledge no corpus study has been conducted to investigate the frequency of referring to the subject versus the object following implicit causality verbs. Similarly, it is difficult to assess how often pronouns like *she* actually refer back to the preceding subject. In a corpus of children's books, Arnold (1998) found that third person subject pronouns co-referred with the previous sentence's subject in 64% of cases. However, it is unclear if this would hold in a larger and more diverse corpus that also includes natural spoken language. Most of what we know actually comes from sentence completion studies (e.g., Ferstl, Garnham, & Manouilidou, 2011), which are limited in the sense that participants have to switch from being the comprehender to the producer. This can result in task demands that do not reflect natural language processing.

Present Study

The aim of the present study was to gain insight into how certain reference biases come to exert their influence on language processing. Specifically we explored whether simple learning mechanisms, such as prediction and reinforcement learning, could help explain why people display certain reference biases. We had a naive cognitive model learn reference biases from an input dataset containing reference asymmetries. The model was presented with simple transitive sentences and had to predict the subsequent referent (subject vs. object), as well as the form of reference (name vs. pronoun). When the model predicted correctly, it was issued a reward. We wanted to determine if the model could pick up on the asymmetries present in the input, as well as investigate how predictions changed as the model was presented with more input.

Methods

Input Data

Our input dataset consisted of 1000 unique items. All items consisted of a simple sentence, containing a transitive verb with its subject and object arguments, followed by a critical referring expression. An example of the four possibilities for a single verb can be seen in (4) below. The information in brackets indicates the referent of the referring expression.

(4) a. Kaitlyn angered Marie. Kaitlyn (Kaitlyn)
b. Kaitlyn angered Marie. Marie (Marie)
c. Kaitlyn angered Marie. She (Kaitlyn)
d. Kaitlyn angered Marie. She (Marie)

All items were created by sampling a verb from a list of 10 verbs that differed with respect to their associated implicit causality: 5 subject-biased verbs (*apologized, repulsed, angered, fascinated, disappointed*), 3 object-biased verbs (*congratulated, feared, comforted*), and 2 non-IC verbs (*filmed, interrupted*). We used an unequal number of each verb type because of our critical assumption that real-world asymmetries in the input are what cause people to display biases. The subject and object referents of each item were randomly sampled from a list of 40 unique female names. The second sentence was determined based on two unique probabilities. The first probability determined which *referent* would be referred to subsequently (i.e., subject or object), for which we used probabilities from Ferstl et al. (2011)'s implicit causality sentence completion corpus. For example, in their study, participants' continuations following an *anger* sentence were about the subject 85% of the time. Therefore, for all of our anger items the sampling probability of the referent being the subject versus object was 0.85/0.15. The second probability determined the *form of reference* (i.e., name or pronoun). Across all verb types, we opted for a general pronoun bias when referring to subjects (with a pronoun sampling probability of 0.75) and a general name bias when referring to objects (with a name sampling probability of 0.75). Given the lack of corpus data, these values were inspired by sentence completion literature.

PRIMs Cognitive Model

Our model was implemented using the cognitive architecture PRIMs (*Primitive Information Processing Elements*, Taatgen, 2013, 2014), which evolved from the ACT-R cognitive architecture (Anderson, Bothell, Lebiere, & Matessa, 1998; Anderson, 2007). Like other cognitive architectures, PRIMs serves as a unified theory of cognition, as well as an interface for implementing models. Like ACT- R, PRIMs assumes that

Table 1: Operators responsible for processing input.

Operator	PRIMs	Description
retrieve-V1	V1<>nil	slot 1 of the input buffer (V) is not empty
	RT1=nil	slot 1 of the retrieval buffer (RT) is empty
	WM1=nil	slot 1 of the working memory buffer (WM) is empty
	==>	
	lexical-entry->RT1	retrieve a 'lexical-entry' chunk from the declarative-
	V1->RT2	where slot 2 of the chunk matches the information currently in V1
store-V1	V1=RT2	slot 1 of the input buffer is the same as slot 2 of the retrieval buffer
	WM1=nil	slot 1 of working memory is empty
	==>	
	RT2->WM1	store the information in slot 2 of the retrieval buffer in slot 1 of working memory

the cognitive system is modular and thus has different modules for specific cognitive functions (e.g., vision, motor control, working memory, declarative memory, etc.). The different modules communicate with each other through their respective buffers. Each buffer has a number of slots that can each hold a single piece of information. Together all the buffers comprise the global workspace of the system. Instead of production rules (as in ACT-R), the exchange of information within the workspace is achieved through the use of operators, which reside as chunks in declarative memory. All operators consist of condition 'PRIMs' and action 'PRIMs' and are retrieved on the basis of their activation. If the conditions of the retrieved operator are met, the actions are carried out; if not, the operator with the next highest activation is retrieved. We will begin by describing how a single trial unfolds within the model. For the current model it is useful to distinguish between operators responsible for processing input (i.e., the sentences the model is presented with) and operators responsible for making predictions.

The model is first presented with a complete sentence (e.g., 'Kaitlyn angered Marie') and then processes each word by retrieving an associated lexical chunk from declarative memory and storing the results in working memory. The two operators responsible for processing the first word are presented with descriptive detail in Table 1 above. The ==> arrow separates condition 'PRIMs' from action 'PRIMs'. The processing of the entire first sentence ultimately results in a completed event representation being held in working memory. It should be noted that because the buffer slots in PRIMs do not have names, the order matters. For example, in our model WM1 is always used to store information about the subject and WM3 is always used to store information about the object.

Next the model predicts the subsequent referent (subject vs. object). The two operators responsible for this are presented in Table 2 below. These operators have the exact same conditional PRIMs and thus in a completely naive model have an equal chance of firing. The crucial difference is which information gets copied into WM5 (the slot reserved for holding information about the subsequent referent). Next the model makes a prediction about the form of reference (name vs. pronoun). The operators responsible for this are presented in Table 3 on the next page. The first three operators again have equal conditional PRIMs and thus an equal chance of firing. The final operator in Table 3 (retrieve-PRO) fires in cases where the model predicts a pronoun in order to account for the fact that different pronouns would be needed to refer to referents depending on their gender and number. However, in the current model the correct pronoun is always *she*. After the model predicts the form of reference, a 'read-next' action fires and the model is presented with the critical referring expression and information about the referent. In cases where

Table 2: Operators responsible for predicting referent.

Operator	PRIMs	Description
predict-subj	WM3<>nil	slot 3 of working memory is not empty
	WM4=nil	slot 4 of working memory is empty
	WM5=nil	slot 5 of working memory is empty
	==>	
	WM1->WM5	copy the information in slot 1 of working memory into slot 5 of working memory
		(Note: information about the subject is stored in WM1)
predict-obj	WM3<>nil	slot 3 of working memory is not empty
	WM4=nil	slot 4 of working memory is empty
	WM5=nil	slot 5 of working memory is empty
	==>	
	WM3->WM5	copy the information in slot 3 of working memory into slot 5 of working memory
		(Note: information about the object is stored in WM3)

Table 3: Operators responsible for predicting reference form.

Operator	PRIMs	Description
predict-subj-name	WM4=nil	slot 4 of working memory is empty
	WM1=WM5	slot 1 and slot 5 of working memory are the same
	RT1=nil	the retrieval buffer is empty
	==>	
	WM5->WM4	copy the information in slot 5 of working memory into slot 4 of working memory
	read-next->AC1	perform 'read-next' action
predict-obj-name	WM4=nil	slot 4 of working memory is empty
	WM3=WM5	slot 3 and slot 5 of working memory are the same
	RT1=nil	the retrieval buffer is empty
	==>	
	WM5->WM4	copy the information in slot 5 of working memory into slot 4 of working memory
	read-next->AC1	perform 'read-next' action
predict-PRO	WM4=nil	slot 4 of working memory is empty
	WM5<>nil	slot 5 of working memory is not empty
	RT1=nil	the retrieval buffer is empty
	==>	
	lexical-entry->RT1	retrieve a 'lexical-entry' chunk from the declarative-
	pronoun->RT3	where slot 3 of the chunk is 'pronoun'
retrieve-PRO	WM4=nil	slot 4 of working memory is empty
	WM5<>nil	slot 5 of working memory is not empty
	RT1<>nil	the retrieval buffer is not empty
	==>	
	RT2->WM4	store the information in slot 2 of the retrieval into slot 4 of working memory
	read-next -> AC1	perform 'read-next' action

this information matches the model's predictions, a reward is issued.

During the initial trials, when the model is naive, various operators are just as likely to fire. For example, predicting an subject versus object continuation is just likely even for items containing a subject-biased IC verb. However, by utilizing PRIMs' context-operator learning, the model is able to learn which combination of operators is most likely to lead to a reward given the current context. As mentioned, operators are retrieved based on their activation, which is primarily influenced by spreading activation. The associated strengths for spreading activation to operators are learned via reinforcement learning. Thus, whenever the model is issued a reward, it increases the association between the current context and all of the operators that lead to the reward being issued. In PRIMs 'context' can be used to refer to the entire global workspace (i.e., all the buffers), however, for this particular model we were only interested in spreading activation from the WM buffer. Thus, when certain information is held in working memory (e.g., 'angered'), specific operators (e.g., predict-subj) are more likely to fire given their increased activation. We also utilized PRIMs' operator compilation, which allows the model to create new operators by combining pairs of operators that resulted in a reward being issued. In our model predicting the 1) referent and 2) form of reference is initially a two-step process. However, we expect the model to compile the operators responsible for predicting the referent (e.g., predict-subj-continuation) and form of reference (e.g., predict-PRO), given that in the input data subjects are most often referred to using a pronoun and objects are most often referred to using a name, which the model should pick up on. We ran the model 100 times, where a single run consisted of the model being presented all 1000 items from the input dataset in a completely randomized order. This eliminates any order effects and allows for us to analyze 'average' behavior.

Results

With respect to the continued referent asymmetry, Figure 1 shows the proportion of subject (dark gray) versus object (light gray) continuations across the three different verb

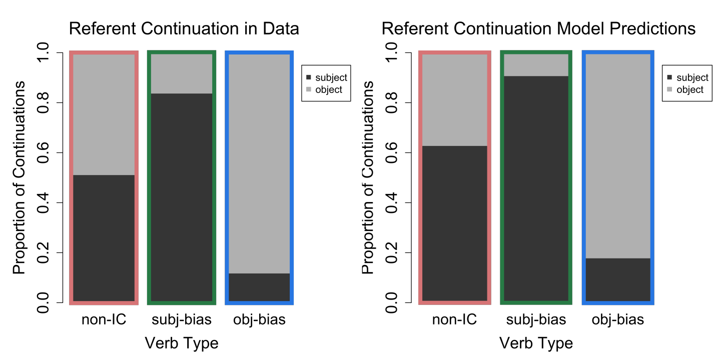

Figure 1: Referent continuations (subject vs. object) by verb type.

types. The left panel is the actual input data and the right panel is the model predictions. As can be seen, the model predictions mirror the input data, primarily predicting subject continuations for items with a subject-biased IC verb and object continuations for items with an object-biased IC verb. With respect to the form of reference asymmetry, Figure 2 shows the proportion of using a name (dark gray) versus pronoun (light gray) when referring to subjects (top panels) and objects (bottom panels) across the three different verb types. The left panels are again the input data and the right panels are the model predictions. When the continued referent was predicted to be the subject, the model largely predicted that the reference would be in the form of a pronoun. When the continued referent was predicted to be the object, the model was more likely to predict the reference would be in the form of a name (except in the case of subject-biased verbs, where it was 50/50 for names and pronouns). These form of reference predictions are in line with the input data (i.e., subjects primarily get referred to with pronouns and objects primarily get referred to with names), however, the model overpredicted pronouns for both subjects and objects.

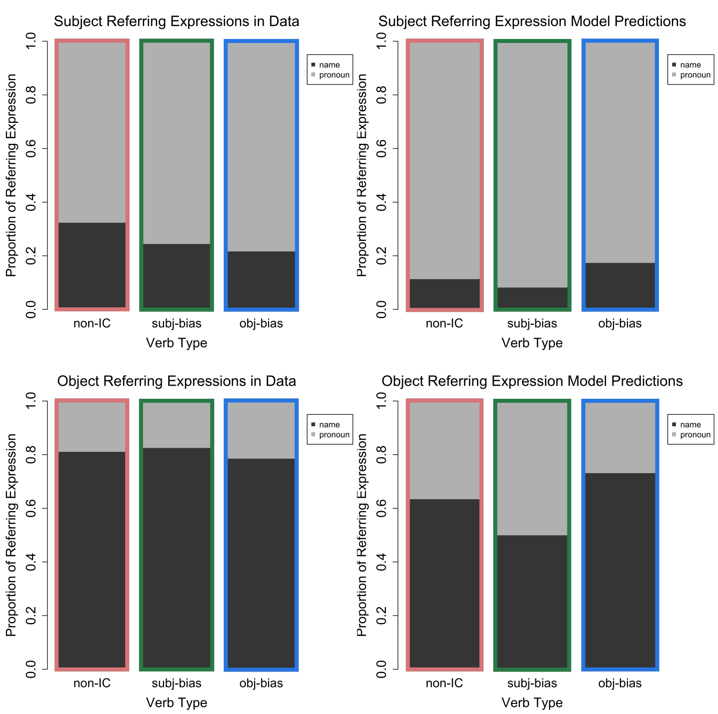

Figure 2: Form of reference continuations (name vs. pronoun) by referent and verb type.

We were also interested in examining learning over trials. Figure 3 illustrates how referent predictions developed across trials (averaged over the 100 model runs). The y-axis represents the proportion of predicting a subject continuation for each of the three verb types. During the initial trials the model predicted subject continuations at chance level across all three verb types. However, as trials unfolded the proportion of predicting a subject continuation increased for items containing a subject-biased IC verb and decreased for items containing an object-biased IC verb. For the non-IC verbs a gradual increase in subject predictions is seen.

Figure 4 illustrates how form of reference predictions developed across trials. The y-axis represents the proportion of pronoun predictions, for both predicted subject continuations (dark gray) and predicted object continuations (light gray). Here we collapsed over verb type, as the verb itself does not influence the form of reference. During the initial trials the model predicted a pronoun at chance level for both subject and object continuations. However, as trials unfolded the proportion of predicting a pronoun increased in cases where the model predicted a subject continuation and decreased in cases where the model predicted an object continuation. The proportion of predicting a name can be calculated by subtracting the pronoun proportion from 1.

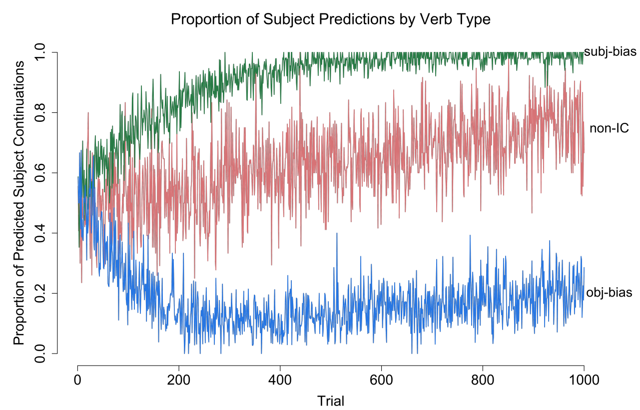

Figure 3: Proportion of subject continuation predictions across trials by verb type.

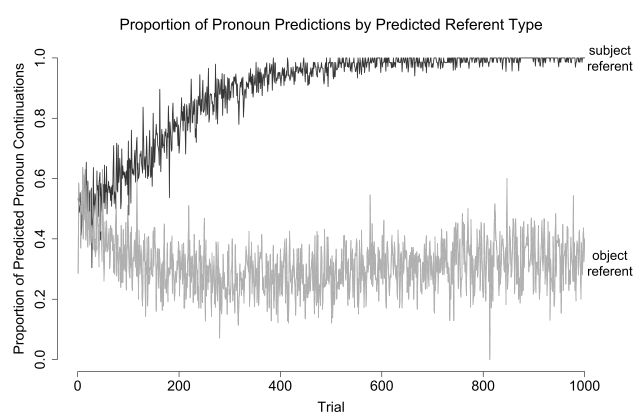

Figure 4: Proportion of pronoun continuation predictions across trials by predicted referent type.

With respect to operator compilation, we expected the model to compile the operators responsible for predicting the referent (e.g., predict-subj) and form of reference (e.g., predict-PRO). However, the model instead combined the operators responsible for predicting the referent (predict-subj and predict-object), with the previous operator that stores lexical information about the object in working memory (i.e., store-v3). So rather than processing the object of the first sentence (i.e., the Marie in 'Kaitlyn angered Marie') and then predicting the subsequent referent, the model combined the two steps. Although this is not what we expected, it also makes sense given that as soon as the verb is stored in working memory, the model can already predict which referent will be referred to next.

Discussion

We constructed a model in the cognitive architecture PRIMs (Taatgen, 2013, 2014) and presented it with input data that contained reference asymmetries. More specifically, in the input dataset we manipulated the proportion of subject versus object continuations by using verbs differing in their associated implicit causality (subject-biased, object-biased and non-IC). Furthermore, we manipulated the likelihood that names versus pronouns would be used to refer to the different referents, such that there was a greater likelihood of using a pronoun when referring back to subjects and a name when referring back to objects. Initially the model was unaware of the asymmetries present in the input. However, by utilizing PRIMs' context-operator learning (reinforcement learning) the model was able to pick up on the asymmetries present in the data, as reflected by its predictions about subsequent reference.

Our model processed sentences like 'Kaitlyn angered Marie' and then made predictions about the subsequent referent, as well as the form of reference. With respect to referent predictions, during the initial trials the model equally predicted subject and object continuations across the three verb conditions. However, as the model was presented with more input this pattern uniquely changed across the three verb conditions, as to mirror the input data. With respect to form of reference predictions, the model initially predicted an equal number of names and pronoun for both subjects and objects. However, as the model was presented with more input, the proportion of predicting a pronoun increased for subjects and decreased for objects (mirroring the input data). Assuming that humans make such predictions of course is not trivial. However, predictive processing is widely assumed to be a core aspect of cognition, especially when it comes to the processing of serial order information, such as language. Furthermore, it is assumed that the reason humans make predictions is to save future processing costs (see Bubic, Von Cramon, & Schubotz, 2010, for a review of predicition). The primary evidence of predictive language processing comes from ERP and visual world eye-tracking studies, which show that people anticipate upcoming arguments following specific verbs, for example, anticipating to hear about something edible, following the verb *ate* (e.g., Altmann & Kamide, 1999; Nieuwland & Van Berkum, 2006).

Our model's predictions can help explain Koornneef and Van Berkum (2006)'s finding that reading times are slower when a pronoun is inconsistent with the bias setup by the verb preceding it. For example, following an object-biased IC verb our model most often predicted an object continuation. However, when the continuation ended up being about the subject (i.e., inconsistent with the prediction), the model had to revise the contents of working memory to accurately represent the second sentence, which took additional time. It is difficult to say how exactly the reward issued to our model relates to a reward in the real world. However, one possibility is that the reward in the real world is a successful saving of processing time (or effort).

Similarly, our findings provide insight into why visual world eye-tracking studies find an immediate effect of implicit causality on pronoun resolution (e.g., Järvikivi et al., 2017). One possibility is that following a sentence containing an IC verb, listeners expect to hear about a specific referent. In cases where listeners expected to hear about the subject, they may have also expected a pronoun, whereas in cases they expected to hear about the object, expecting a pronoun would be less likely. Nevertheless, in both cases listeners are presented with an ambiguous pronoun and have no choice but to incorporate it into their discourse representation, which leads to the different gaze patterns. Unfortunately, studies specifically interested in pronoun resolution only report on time windows starting at the pronoun onset. Looking at earlier times (e.g., starting at the verb onset) may actually be crucial for understanding previously reported 'pronoun' effects. An empirical prediction of our model, is that that the previous findings are not necessarily about pronouns and how to interpret them, but about how verb biases affect predictions about next referents and their forms.

With respect to future directions, our model was always given enough time to make predictions about upcoming referents and their forms. This is not ideal given that in the real world there is a continuous stream of input that cannot be controlled by the comprehender. This will be addressed in future studies so that the model will only make predictions when it has enough time to do so. Furthermore, we determined our input frequencies based on reasonable assumptions given the current lack of an annotated corpus. In the future it would be informative to explore various relative frequencies and see what effect it has on the model's predictions. One way to address this would be to include a wide range of verbs, as evidence suggests that implicit causality actually lies on a continuum. Investigating how a continuous measure of IC influences the type of predictions may be fruitful for understanding people's behavior in previous experimental studies. Finally, although our model predictions can provide insight into some of the previous experimental findings, in none of those studies were participants asked to make explicit predictions. It would be informative to carry out a prediction study so that we could directly compare the model to human prediction data. Although our approach was quite simple, it highlights the fact that seemingly complex behavior can often be explained by simple learning mechanisms.

References

Altmann, G. T., & Kamide, Y. (1999). Incremental interpretation at verbs: restricting the domain of subsequent reference. *Cognition*, *73*(3), 247-264.

Anderson, J. (2007). *How can the human mind occur in the physical universe?* 198 Madison Avenue, New York, New York: Oxford University Press.

Anderson, J., Bothell, D., Lebiere, C., & Matessa, M. (1998). An integrated theory of list memory. *Journal of Memory and Language*, *38*, 341-380.

Arnold, J. E. (1998). *Reference form and discourse patterns*. Doctoral dissertation, Stanford University.

Brown, R., & Fish, D. (1983). The psychological causality implicit in language. *Cognition*, *14*(3), 237-273.

Bubic, A., Von Cramon, D. Y., & Schubotz, R. (2010). Prediction, cognition and the brain. *Frontiers in Human Neuroscience*, *4*(25), 1-15.

Ferstl, E. C., Garnham, A., & Manouilidou, C. (2011). Implicit causality bias in english: a corpus of 300 verbs. *Behavior Research Methods*, *43*, 124-135.

Garnham, A., Traxler, M., Oakhill, J., & Gernsbacher, M. A. (1996). The locus of implicit causality effects in comprehension. *Journal of Memory and Language*, *35*(4), 517-543.

Gernsbacher, M. A. (1989). Mechanisms that improve referential access. *Cognition*, *32*(2), 99-156.

Järvikivi, J., Van Gompel, R. P. G., & Hyönä, J. (2017). The interplay of implicit causality, structural heuristics, and anaphor type in ambiguous pronoun resolution. *Journal of Psycholinguistic Research*, *46*, 525-550.

Järvikivi, J., Van Gompel, R. P. G., Hyönä, J., & Bertram, R. (2005). Ambiguous pronoun resolution: contrasting the first-mention and subject-preference accounts. *Psychological Science*, *16*, 260-264.

Kehler, A., Kertz, L., Rohde, H., & Elman, J. L. (2008). Coherence and coreference revisited. *Journal of Semantics*, *25*(1), 1-44.

Koornneef, A. W., & Van Berkum, J. J. A. (2006). On the use of verb-based implicit causality in sentence comprehension: Evidence from self-paced reading and eye tracking. *Journal of Memory and Language*, *54*(4), 445-465.

Nieuwland, M. S., & Van Berkum, J. J. A. (2006). When peanuts fall in love: N400 evidence for the power of discourse. *Journal of cognitive neuroscience*, *18*(7), 1098-1111.

Rohde, H., & Kehler, A. (2014). Grammatical and information-structural influences on pronoun production. *Language, Cognition and Neuroscience*, *29*(8), 912-927.

Rudolph, U., & Forsterling, F. (1997). The psychological causality implicit in verbs: A review. *Psychological Bulletin*, *121*, 192-218.

Stevenson, R. J., Crawley, R. A., & Kleinman, D. (1994). Thematic roles, focus and the representation of events. *Language and Cognitive Processes*, *9*(4), 519-548.

Sukthanker, R., Poria, S., Cambria, E., & Thirunavukarasu, R. (2020). Anaphora and coreference resolution: A review. *Information Fusion*, *59*, 139-162.

Taatgen, N. A. (2013). The nature and transfer of cognitive skill. *Psychological Review*, *120*(3), 439-471.

Taatgen, N. A. (2014). Between architecture and model: Strategies for cognitive control. *Biologically Inspired Cognitive Architectures*, *8*, 132-139.