

# Syntactic Biases in Intentionality Judgments

Brent Strickland (brent.strickland@yale.edu)  
Department of Psychology, Yale University

Matt C. Fischer (matthew.c.fisher@biola.edu)  
Biola University

Elodie Peyroux (elodie.peyroux@hotmail.fr)  
Universite Lumiere Lyon 2

Frank Keil (frank.keil@yale.edu)  
Department of Psychology, Yale University

**Abstract:** The notion of a “thematic role” has been an important one in linguistic theories concerning the syntax/semantics interface (Fillmore, 1968), but their effects on high-level cognition remain relatively unexplored. Thematic roles, like AGENT and PATIENT, are said to be linked with specific grammatical positions like that of “subject” and “direct object”. Here we ask if the link between grammatical subjects and the thematic role AGENT might create an intentionality bias for subjects. We tested this in a series of studies examining both response times and accuracy. These studies reveal a quick and reflexive bias to treat grammatical subjects (but not prepositional objects or direct objects) as being more intentional than they actually are. However, this bias may be overcome when people are encouraged to reflect on their true knowledge. Broader implications of our findings for research in psychology and linguistics are discussed.

**Keywords:** Intentionality Bias, Theory of Mind, Dual Systems, Thematic Roles

## Introduction

Linguists have hypothesized the existence of “thematic roles” in order to explain semantic similarities between actors mentioned in specific syntactic positions (Dowty, 1991; Carlson & Tanenhaus, 1988; Fillmore, 1968). The most relevant example here is that syntactic subjects are said to be associated with the AGENT thematic role, where semantic agents are those things which prototypically cause events to happen, intend them to happen, and initiate their occurrence (Dowty, 1991).

Given their association with the AGENT thematic role, syntactic subjects in English are often understood to have acted intentionally. Take the

following four verbs as an example: “eat,” “help,” “deceive,” and “construct.” Despite large differences in meaning, for each verb it is nearly always the case that the syntactic subject denotes an agent who intentionally brought about some state of affairs. Of course, there are syntactic subjects who do not act intentionally but do other things like cause or initiate an event (as in “the tornado destroyed the power lines”). Nevertheless, the notion of intentionality does seem to be strongly associated with the syntactic subject position across many verbs and many verb classes (Sorace, 2000).

The computational logic of on-line language processing suggests that inferring others’ intentions is likely to reflexively influence decisions regarding syntactic subjects. Imagine someone producing the sentence “Man bites dog” in context. When the speaker decides to put “man” as the subject instead of the object, he/she must have identified something about the role the man played in the event in order to decide that “man” is a grammatical subject and not an object. In this case, intentionality is a pretty good indicator of which actor goes in the subject position of the sentence.

In comprehension contexts, there is also some recent empirical evidence suggesting that the representation of others’ intentions interacts with one’s grammatical knowledge. Childers & Echols (2004) showed that when learning the meaning of a new word, children have a strong expectation for grammatical subjects to be intentional, animate actors. While this bias could interfere with learning in some cases, it improves performance in prototypical events where an animate agent acts on an inanimate patient, and that event is described with by a sentence placing the animate actor in the grammatical subject position.

Thus, both empirical studies and common sense suggest that the theory of mind concept of “intention”

could influence language processing. But here we want to ask the novel, opposite question. Could a noun's syntactic position relative to a verb influence how that noun's role in an event is conceptualized?

More specifically, perhaps the grammatical subject of a sentence undergoes more of an intentionality bias than other parts of speech due to its association with the AGENT thematic role. So while people can reflect on the intentionality of the actors described in a sentence in a deeper, more reflective way, there may also exist a simple heuristic like, "Syntactic Subject  $\rightarrow$  Intentional," which operates in quick, reflexive manner and trumps deeper knowledge in some cases.

## Experiment 1

While the notion of a general intentionality bias has recently become a hot topic in cognitive science (see Rosset, 2008), the question of how syntactic heuristics may influence any putative intentionality bias has not, to our knowledge, been explored.

In order to determine whether there may be more of an intentionality bias for subjects compared to other parts of speech, we employed logically equivalent sentences like "Frank exchanged products with Steven" and "Steven exchanged products with Frank." (see Gleitman et al, 1996). In such sentences, each sentence is both necessary and sufficient for it's reverse to be true. Given their logical equivalence, both sentences necessarily describe an event (or set of possible events) in which Frank must be acting equally intentionally in both cases. However, we predicted that "Frank" would be judged as more intentional when he was described as the grammatical subject than when he was described as the object of the preposition.

In order to test for any effects of word order we ran control Experiment 1b. In order to test for any effects of distance from the beginning or end of the sentence, we also ran control Experiment 1c.

### Method

1000 paid online participants from Amazon's Mechanical Turk.

**Materials and Procedure:** In Experiment 1a, participants saw a single sentence from a pair of logically equivalent sentences like the "exchange" example above. The other verbs were: "swap", "trade", "loan/borrow", and "buy/sell". All verbs contained "Frank" and "Steven" as the proper nouns designating the actors in the event.

Each participant read only one sentence and rated on a scale of 1-7 how intentionally they thought either the subject (e.g. "Frank") or the prepositional object acted (e.g. "Steven").

Experiment 1b employed used the same method

except that list of verbs was limited to exchange, trade and swap. Both nouns were the grammatical subject in this condition. An example sentence is: "Frank and Steven exchanged books."

Experiment 1c was identical to 1a with the following exceptions. The verbs employed were: "exchange", "swap", "loan", "borrow", "sell", "buy". Each verb had two constructs: a clefted subject ("It was Frank that exchanged books with Steven") or a clefted object ("It was Frank that Steven exchanged books with").

## Results

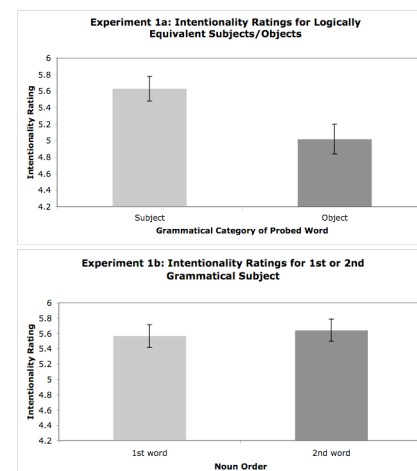
The results of Experiments 1a-1c matched our predictions (see Figure 1 below). All subsequent analyses below are by item analyses unless otherwise stated. In Experiment 1a, participants judged the grammatical subject as having acted more intentionally ( $M = 5.63$ ) than the prepositional object ( $M = 5.01$ ),  $t(6) = 3.7$ ,  $p < .05$ .

In Experiment 1b where both actors were the grammatical subject, both the first noun ( $M = 5.56$ ) and the second noun ( $M = 5.64$ ) were judged to be equally intentional,  $p = .71$ .

In Experiment 1c the clefted subject ( $M = 5.68$ ) was rated as being more intentional than the clefted object ( $M = 4.94$ ),  $t(5) = 4.02$ ,  $p < .05$ .

## Discussion

Experiment 1 shows that participants use grammatical position as a heuristic for intentionality judgments. Participants judged the grammatical subject of logically reversible sentences as acting more intentionally than the prepositional object. Thus the same actor from the exact same event will be treated differently depending on how that event is described. Experiments 1b and 1c rule out the possibility that these effects are due to simple word order or recency effects.



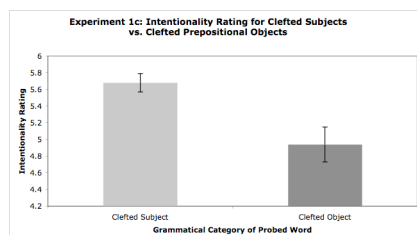


Figure 1: Results from Experiments 1a – 1c. The grammatical subject of the sentence is reliably rated as more intentional than the prepositional object in logically reversible sentences.

## Experiment 2

In Experiment 2 we ask two related questions. (1) Can people hold this intentionality bias for syntactic subjects while still believing in the logical equivalence of our sentences? (2) Can people overcome this bias by shifting to a more logical mode of thought?

Here, we employed a slightly more complicated design than in the previous experiment. The experiment was divided into three parts. In the pre-prime phase, we probed participants on their intentionality judgments for grammatical subjects *and* objects in logically reversible sentences as well as non-logically reversible sentences. We then “primed” them to think logically by asking them if they agreed that the logically reversible sentences from Experiment 1 were in fact logically reversible. Then we re-tested them as in the pre-prime phase.

First, we predicted that people would show an intentionality bias even if they would later agree that our “logically reversible” sentences were in fact logically reversible. In other words, we predicted that participants would contradict themselves. Secondly, we predicted that participants would eliminate or reduce their intentionality bias for grammatical subjects after being encouraged to think about the logic of our sentences.

### Method

56 paid online participants from Amazon’s Mechanical Turk.

**Procedure and Stimuli:** In the “pre-prime” phase, subjects first saw nine sentences, three of which were logically reversible and six of which were not. They rated both the subject and object on how intentionally they acted.

During the priming phase, participants were then shown nine different sentences, three of which were logically reversible and six of which were not. For each sentence, participants were asked whether the

sentence entails the same sentence with the actors reversed. For example, participants might be asked “If Bill swapped books with Susan, is it necessarily the case that Susan swapped books with Bill?” Participants responded “Yes” or “No.”

After making each response they then rated the intentionality of both the grammatical subject and grammatical object on a seven-point scale.

### Results

Participants “qualified” for data analysis if they indicated that they believed *all* logically reversible sentences were in fact logically reversible (45 out of 56 subjects). We selected this group because we were mainly interested in those participants who were reasoning in the relevant way.

These participants indeed showed less of an intentionality bias for the subject after logical priming than before. The mean ratings are as follows: pre-prime subject  $M=6.32$ ; pre-prime object  $M=5.1$ ; post-prime subject  $M=6.35$ ; post-prime object  $M=5.97$ . The pre-prime difference between subjects and objects was statistically significant  $t(44)=3.53$ ,  $p < .001$ , and the post-prime difference was also statistically significant  $t(44)=3.53$ ,  $p < .001$ . Crucially however, the interaction between the conditions was also statistically significant,  $p < .001$ . This indicates that intentionality bias for the subject significantly decreased after the logical prime.

For the 11 participants who did not agree with the principle of logical reversibility, they also treated grammatical subjects as being more intentional than grammatical objects both pre-prime ( $M=5.48$ ,  $M=4.06$ ,  $t(10)=3.47$ ,  $p < .01$ .) and post-prime ( $M=4.97$ ,  $M=3.97$ ,  $t(10)= 3.169$ ,  $p < .05$ ). However, the interaction between priming condition and grammatical position was not significant,  $p=.37$ .

### Discussion

Experiment 2 supported our two predictions. First, some participants held contradictory beliefs with regards to the intentionality of the actors described in our sentences. They were biased to see the grammatical subjects as more intentional than other logically equivalent actors. However, they also recognized that these actors were in fact logically equivalent. Thus, for the sentence “Bill swapped books with Susan,” they simultaneously believed that “Bill” was more intentional than “Susan” while being logically committed to their being equally intentional.

When confronted with this fact, participants adjusted their intentionality ratings for logically equivalent subjects and objects by making them more similar. Thus participants were able to reduce the influence of the heuristic “subject → intentional” by

entering a more logical mode of thought that clearly contrasted with their less reflective judgments.

It is worth pointing out that in adjusting their ratings in this way, intentionality ratings for the prepositional objects went up instead of ratings going down for the grammatical subject. This is pattern consistent with an “unintentionality bias” for non-subjects. But this pattern is also consistent with an intentionality bias for the subject since it could have resulted from simple preference to always move ratings up whenever adjusting. Experiment 3 addresses this point in a more direct fashion.

### Experiment 3

Experiments 1 and 2 addressed the nature and flexibility of the intentionality bias for the grammatical subject of a sentence. Experiment 3 addressed its expression under time pressure. We did this by gathering intentionality judgments from native French speakers for two types of verbs: unaccusative and unergative. According to linguistic theory, unaccusative verbs are verbs whose surface grammatical subject has been moved from the grammatical object position (Levin & Rappaport Hovav, 1994). Unergative verbs, on the other hand, have surface subjects who are also the deep structure subject<sup>1</sup>.

In French, unaccusative verbs take “être” as the auxiliary in the past tense while unergative verbs take “avoir” as the auxiliary in the past tense. Both verb types can express actions that are clearly intentional and actions that are clearly unintentional. Take the following examples (“trace” represents the syntactic position out of which the underlined noun is moved):

(1) Unaccusative (Etre) Intentional:

Christophe est sorti (trace) à neuf heures

*Christophe left (trace) at nine o'clock*

(2) Unergative (Avoir) Intentional:

Emilie a marché en ville

*Emilie walked around in town*

(3) Unaccusative (Etre) Unintentional:

François est tombé (trace) sur la glace

*Francois fell (trace) on the ice*

(4) Unergative (Avoir) Unintentional

Marc a glissé sur la chaussée

*Marc slipped on the walkway*

---

<sup>1</sup> Most modern syntacticians would disagree with the term “deep structure,” as it has been made somewhat obsolete by minimalism. We choose this term simply for clarity of explanation to a broad audience.

The heuristic account of intentionality judgments thus predicts for “avoir” verbs, correct “intentional” judgments should be faster than correct “unintentional” judgments since the surface subject is the deep structure subject. If time pressure exacerbates the intentionality bias for subjects, there should also be more “intentional” false alarms than “unintentional” false alarms for “avoir” verbs. In other words, people should be more likely to mistake an unintentional character as acting intentionally than the opposite. On the other hand, such biases should not exist for “être” verbs since the surface subject is actually the displaced deep structure object.

### Method

14 paid native French speakers from Brest, France participated in the experiment.

**Procedure and Stimuli:** Participants saw 72 sentences: 18 “avoir” intentional, 18 “avoir” unintentional, 18 “être” intentional and 18 “être” unintentional. Participants sat in front of a computer screen and were instructed to indicate (via keypress) “as quickly as possible” whether the person described in the sentence was acting intentionally. All sentences were made such that they only described one person.

All stimuli were normed beforehand on a different group of non-time pressured participants in order to ensure broad agreement that each sentence was either clearly intentional or clearly unintentional. We biased the stimuli against our experimental hypothesis such that norming agreement for “avoir” sentences was higher for unintentional actions ( $M=98.73\%$ ) than intentional actions ( $M=94.74$ ). On the other hand, agreement was higher for “être” intentional actions ( $M=90.46$ ) than unintentional actions ( $M=84.21$ ).

### Results

For “avoir” verbs, participants were indeed faster to correctly judge an action to be intentional ( $M = .24$  sec./syllable) than unintentional ( $M = .26$ ),  $t(13) = 2.62$ ,  $p < .05$ . However for “être” verbs there was no significant difference in correctly identifying an intentional action ( $M = .274$ ) compared to an unintentional action ( $M = .269$ ),  $p = .66$  (see figure 2 below).

Accuracy data also revealed that participants were indeed more likely to make “intentional” false alarms on “avoir” verbs but not “être” verbs. “Avoir” accuracy rates differed significantly between intentional ( $M=.94$ ) and unintentional actions ( $M=.82$ ),  $t(13)=4.22$ ,  $p<.001$ . However “Etre” accuracy rates did not differ between intentional ( $M=.88$ ) and unintentional ( $M=.88$ ). In other words, for “avoir” verbs, participants were more likely to

mistake an unintentional actor as acting intentionally than vice-versa. However, there was no difference in false alarm rates for “être” verbs.

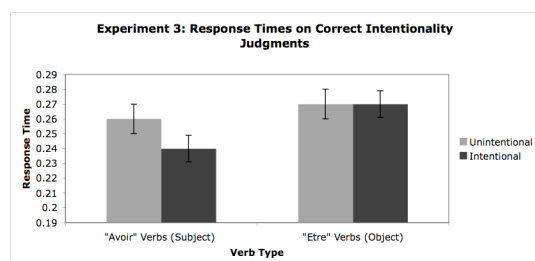


Figure 2: Response times for “avoir” (unergative) and “être” (unaccusative) verbs. Only “avoir” verbs showed evidence of an intentionality bias.

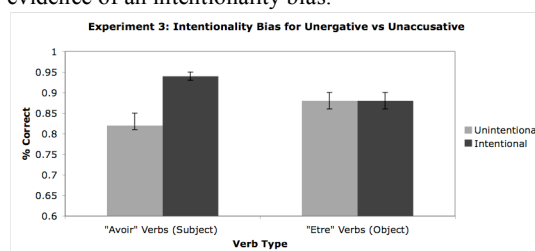


Figure 3: Accuracy data. Despite the fact that unspeeded intentionality judgments from norming experiments were biased against our hypothesis, participants displayed a bias to incorrectly label unintentional “avoir” sentences as being intentional. This did not hold for “être” verbs.

## Discussion

Experiment 3 demonstrates that under time pressure people have a bias to treat deep structure grammatical subjects (in the “avoir” verbs) as being more intentional than they should be treated. However, they have no such bias for the deep structure grammatical object that has been covertly moved to the subject position (in the “être” verbs). This bias was revealed both by “intentional false alarm rates” and response times.

Furthermore, these results suggest that there is a specific intentionality bias for grammatical subjects as opposed to an unintentionality bias for non-subjects since the main differences in response times and false alarm rates concerned syntactic subjects as opposed to transposed objects.

## Conclusion

Together, the three experiments presented here tell a coherent story. There are at least two ways of generating an intentionality judgment from verbal reports. One is a quick, heuristic judgment that is perhaps generated from an association between the syntactic subject position and the AGENT thematic role. However this bias may be overcome with more

time and a more reflective mindset. These findings have potentially interesting implications for other areas of cognitive science and linguistics.

## Two Systems?

Apperly & Butterfill (2009) have recently argued for a two systems view of belief representation whereby there exists a quick and efficient system for automatically calculating others’ belief states, along side a slower more inefficient system (see also Cohen & German, 2009).

Although we were principally interested in the representation of intentions as opposed to beliefs, the findings from the current set of studies are compatible with the two systems account. The results from Experiment 3 show that under time pressure, the intentionality bias for grammatical subjects is exaggerated. However the bias is overcome when participants have more time to reflect on the meaning sentences. The results from Experiment 2 show that participants are generally influenced by the syntactic subject bias even when they simultaneously hold a deeper commitment to the logical equivalence of subjects and objects.

One plausible explanation for these discrepancies between performance in speeded and unspeeded (or natural vs. logical) tasks is that the presence vs. absence of the bias is dictated by which “system” is being tapped to generate the relevant judgment.

However, there is another possible explanation of our effects that would be incompatible with the two systems account. Instead of multiple systems, there could be a single system that prioritizes different intentionality cues differently. Under time pressure or certain task conditions, only those cues highest on the priority list may be employed. Right now, our results cannot distinguish between these two possibilities, and this is an area for further research.

## Linguistic Theory and Thematic Roles

There are arguments in the linguistics literature about the nature and role of “thematic roles”. Notably, there have been disagreements about logically reversible verbs, with some authors arguing that logically reversible verbs like “exchange” or “buy/sell” necessarily assign identical thematic roles to all actors in the event (Dowty, 1991) while others have argued the exact opposite (Jackendoff, 1987).

The data presented here strongly suggest that in so far as thematic roles are generating any intentionality bias specific to syntactic subjects, then logically reversible verbs do not assign identical thematic roles to their actors. Instead, even when two logically reversible sentences describe the exact same situation, the way that the actors will be conceptualized depends (at least initially) on the

thematic role associated with the syntactic position they appear in.

More generally, methods like those employed in the above studies could be used in deciding between two seemingly plausible theories regarding the nature of thematic roles in an empirically grounded way. In this regard, it is at least possible that methods like ours could help advance theories dealing with the syntax/semantics interface.

Another possible theoretical advance afforded by empirical methods like these may be the ability to deduce the contents of the AGENT thematic role. One possibility is that the notion of “intention” is part of the primitive AGENT concept, and that the link from syntax (i.e. subject) to semantics (i.e. intention) is more or less direct. However, another possibility is that the link between grammatical subject and intention is less direct. For example, the AGENT concept could instead refer to event initiators, and the intentionality bias for syntactic subjects could be an indirect result of that. In our studies, people may subconsciously be reasoning that if “Frank” is the grammatical subject he must have started the event. And since he started the event, he must therefore be more intentional. Both theories are plausible, and methods like these may offer straightforward tools for making progress on this question and others like it.<sup>2</sup>

### Acknowledgements

We would like to thank Joshua Knobe, Bernard Victorri, Beth Levin and Isabelle Tapiero for useful comments on this paper. Preparation of this report and some of the research described therein was supported by NIH grant # R37HD023922 to F. Keil.

### References

- Apperly, I.A. & Butterfill, S.A., (2009). Do humans have two systems to track beliefs and belief-like states? *Psychological Review*, 116(4), 953-970.
- Baron-Cohen, S. (1995). *Mindblindness: an essay on autism and theory of mind*. Cambridge, MA: MIT Press.
- Carlson, G. and M. Tanenhaus (1988). Thematic Roles and Language Comprehension. In W. Wilkins, (ed.), *Syntax and Semantics*, Vol. 21: Thematic Relations, Academic Press, pp. 263-89.
- Childers, J. B., & Echols, C. H. (2004). Two ½ year-old children use animacy and syntax to learn a new noun. *Infancy*, 5, 109-125.

- Cohen, A.S. & German, T.C. (2009). Encoding of others' beliefs without overt instruction. *Cognition*, 111, 356-363.
- Dowty D.R (1991). Thematic Proto-Roles and Argument Selection. *Language* 67: 547-619.
- Fillmore, C. (1968). The Case for Case. In *Universals in Linguistic Theory*. Bach, E. and Harms, R. (eds.) pp. 1--90.
- Gao, T., Newman, G. E., & Scholl, B. J. (2009). The psychophysics of chasing: A case study in the perception of animacy. *Cognitive Psychology*, 59(2), 154 - 179.
- Gergely, G., Nádasdy, Z., Csibra, G., & Biró, S. (1995). Taking the intentional stance at 12 months of age. *Cognition*, 56(2), 165-193.
- Gleitman, L., Gleitman, H., Miller, C., and Ostrin, R., (1996) 'Similar' and similar concepts, *Cognition* 58, 321-376.
- Hamlin, J.K., Wynn, K., & Bloom, P. (2007). Social evaluation by preverbal infants. *Nature*, 450, 557-559.
- Jackendoff, R. 1987. The status of thematic relations in linguistic theory. *Linguistic Inquiry* 18, 369-412.
- Levin, B. & Rappaport Hovav, M. (1995) *Unaccusativity: At the Syntax-Lexical Semantics Interface*, Linguistic Inquiry Monograph 26. Cambridge, MA: MIT Press.
- Newman, G.E., Keil, F.C., Kuhlmeier, V., & Wynn, K. (2010). Sensitivity to Design: Early Understandings of the Link between Agents and Order. *Proceedings of the National Academy of Sciences*, 107, 17140-17145.
- Rosset, E. (2008). It's no accident: Our bias for intentional explanations. *Cognition*, 108, 771-780.
- Saxe, R. (2009). Theory of Mind (Neural Basis). *Encyclopedia of Consciousness*.
- Sorace, A. (2000). Gradients in auxiliary selection with intransitive verbs. *Language*, 76, 859-890.

<sup>2</sup> Indeed follow up studies in our lab actually indicate that this second possibility may in fact be the correct one.