So at this point I think it will come as a shock to absolutely nobody in this room that competent language users have strong intuitions about who a pronoun refers to even given minimal context. Here, for instance, are a few fairly minimal sentences with pronouns, which I've truncated right after the pronoun
Re-Mention Biases

“First-Mention Bias”

Sally ate lunch with Mary. She...
Sally went to the store with Mary. She...

[READ] Even though you don't know how this sentence ends, you probably suspect that the pronoun refers to "she" in both cases."
In fact, in a variety of contexts, people tend to interpret pronouns as coreferencing with the subject of the previous sentence. Whether this is really a subject bias, a topic bias, or a first-mention bias is something that can be debated, but by convention it's usually called the "first-mention bias".

[PAUSE] By the way, you might wonder why I call this a "re-mention bias" rather than pronoun bias, etc. I'll answer that question in a moment.

Here are a couple more examples
Re-Mention Biases

“First-Mention Bias”
- Sally ate lunch with Mary. She...
- Sally went to the store with Mary. She...

“Implicit Causality Bias”
- Sally surprised Mary because she...
Re-Mention Biases

“First-Mention Bias”

Sally ate lunch with Mary. She...
Sally went to the store with Mary. She...

“Implicit Causality Bias”

Sally surprised Mary because she...
Re-Mention Biases

“First-Mention Bias”
Sally ate lunch with Mary. She...
Sally went to the store with Mary. She...

“Implicit Causality Bias”
Sally surprised Mary because she...
Sally loved Mary because she...
Re-Mention Biases

“First-Mention Bias”
Sally ate lunch with Mary. She...
Sally went to the store with Mary. She...

“Implicit Causality Bias”
Sally surprised Mary because she...
Sally loved Mary because she...
Re-Mention Biases

“First-Mention Bias”
Sally ate lunch with Mary. She...
Sally went to the store with Mary. She...

“Implicit Causality Bias”
Sally surprised Mary because she...
Sally loved Mary because she...

“Implicit Consequentiality Bias”
Sally surprised Mary so she...
**Re-Mention Biases**

**“First-Mention Bias”**
- Sally ate lunch with Mary. She...
- Sally went to the store with Mary. She...

**“Implicit Causality Bias”**
- Sally surprised Mary because she...
- Sally loved Mary because she...

**“Implicit Consequentiality Bias”**
- Sally surprised Mary so she...
Re-Mention Biases

“First-Mention Bias”
Sally ate lunch with Mary. She...
Sally went to the store with Mary. She...

“Implicit Causality Bias”
Sally surprised Mary because she...
Sally loved Mary because she...

“Implicit Consequentiality Bias”
Sally surprised Mary so she...
Sally loved Mary so she...
So why call these re-mention biases? [PAUSE] It turns out that if you take these same sentences
Sally surprised Mary so she...
Sally loved Mary so she...
Sally surprised Mary because she...
Sally loved Mary because she...
Sally ate lunch with Mary. She...
Sally went to the store with Mary. She...
“First-Mention Bias”

“Implicit Causality Bias”

“Implicit Consequentiality Bias”

[PAUSE] and truncate the sentence before the pronoun
Re-Mention Biases

“First-Mention Bias”
Sally ate lunch with Mary....
Sally went to the store with Mary....

“Implicit Causality Bias”
Sally surprised Mary because...
Sally loved Mary because...

“Implicit Consequentiality Bias”
Sally surprised Mary so...
Sally loved Mary so...

and ask people to write continuations for each sentence ... people are more likely to continue the stories by talking about one character than the other, and that pattern of who they talk about matches the pronoun comprehension biases almost exactly.
This has been shown a few times, and I'll show you some more evidence towards the end of the talk. [PAUSE] So it looks like speakers have a bias to talk about certain people in certain contexts, and listeners have a bias to expect speakers to talk about those people in those contexts. It's probably a good thing that these biases match up -- it's hard to imagine how communication would be possible if speakers had one set of biases and comprehenders another! In any case, I'll use the term "re-mention bias" to refer to *both* these production biases and comprehension biases.
So like many people here, I'd like to understand these biases, understand how they're acquired, what representations they depend on etc. For this, it's helpful to have some empirical facts on the table that you can build your theory around. My reading of the literature is that in recent years, there is increasingly a consensus about these facts, which looks something like this.
Re-Mention Biases: Current Consensus

1. Rapid effects on pronoun interpretation
   (Cozijn, Commandeur-Vonk & Noordman, 2011
    Featherstone & Sturt, 2010
    Koornneef & Sanders, 2012
    Koornneef & van Berkum, 2006
    Pyykkonen & Jarvikivi, 2010
    van Berkum, Koornneef, Otten & Nieuwland, 2007
    etc.)

First, re-mention biases work very rapidly to constrain pronoun interpretation. There was some debate about this in the 90s, but in the last decade there's been a slew of studies to support this claim.
Re-Mention Biases: Current Consensus

1. Rapid effects on pronoun interpretation
   (Cozijn, Commandeur-Vonk & Noordman, 2011
    Featherstone & Sturt, 2010
    Koornneef & Sanders, 2012
    Koornneef & van Berkum, 2006
    Pykkonen & Jarvikivi, 2010
    van Berkum, Koornneef, Otten & Nieuwland, 2007
    etc.)

2. Not predictable from linguistic structure
    Garvey & Caramazza, 1974
    Koornneef & Sanders, 2012
    Pickering & Majid, 2007
    etc.)

Second, re-mention biases cannot be explained by traditional theories of linguistic structure and rather seem to be the result of pragmatics.
This is because re–mention biases can be shown to be driven primarily by world knowledge. [PAUSE] So my talk has this cheeky title "three myths about re–mention biases" so you probably have guessed where I'm going -- you're building discourse expectations! I'm going to try to convince you that all three of these claims are false. This is really important. I got into this work because, like many of us, I was interested in what the fact that re–mention biases are fast, non–linguistic, and derived from world knowledge should mean for our theories of language processing and discourse expectations. If these assumptions are wrong, we're barking up the wrong tree. But the right tree turns out to be a really interesting, fun tree to bark up, as I'll discuss at the end. So this talk won't be all doom and gloom. But it is doom & gloom for these three claims.
### Re-Mention Biases: Current Consensus

1. **Rapid effects on pronoun interpretation**
   - Cozijn, Commandeur-Vonk & Noordman, 2011
   - Featherstone & Sturt, 2010
   - Koornneef & Sanders, 2012
   - Koornneef & van Berkum, 2006
   - Pyykkonen & Jarvikivi, 2010
   - van Berkum, Koornneef, Otten & Nieuwland, 2007 etc.

2. **Not predictable from linguistic structure**
   - Garvey & Caramazza, 1974
   - Koornneef & Sanders, 2012
   - Pickering & Majid, 2007 etc.

3. **Largely derived from world knowledge**
   - Pickering & Majid, 2007 etc.

---

First, looking at the claim that re-mention biases operate rapidly... [PAUSE] So I've actually got a poster going into this issue in more depth,
It turns out that while authors have typically concluded that re-mention biases are fast, that's not what the data actually show. The issue has been studied most extensively in terms of implicit causality, and it turns out that when you look at subject–biased and object–biased sentences separately, you usually find that pronouns in subject–biased sentences are resolved quickly, but object–biased sentences are resolved more slowly.
Here I have just replicated that. This is a visual-world eyetracking study. So participants are hearing implicit causality sentences with pronouns in them [READ] while looking at a picture like this one, and we're just keeping track of who they look at. I'm graphing looks to the subject of the sentence on the y-axis. The blue line is subject-biased sentences. You will see that right after participants hear the pronoun, they look to the subject of the sentence. [POINT] This replicates this rapid effect of implicit causality bias. But when you look at the red line, which is the object-biased sentences, pretty much nothing happens at the pronoun, and participants looking at the subject half the time and the pronoun half the time.

[PAUSE] This is what nearly every study that has broken down their results into subject-bias and object-bias has shown. Usually the slow resolution for object-biased sentences is explained away as a confound. For instance, maybe it's a confound of order-of-mention. Maybe it's just easier to resolve pronouns to previous subjects, and that interacts with the implicit causality bias during online processing. Well, you can control for these confounds, which I've done.
You'll have to come to the poster to see the details, but it turns out that they can't explain these effects. There are cases in which re-mention biases are in fact quite slow, and those cases may be more common than the cases in which they are fast.
<table>
<thead>
<tr>
<th>Re-Mention Biases: Current Consensus</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>1. Rapid effects on pronoun interpretation</strong></td>
</tr>
</tbody>
</table>
| (Coqjin, Commandeur-Vonk & Noordman, 2011  
  Featherstone & Sturt, 2010  
  Koornneef & Sanders, 2012  
  Koornneef & van Berkum, 2006  
  Pyykonen & Jarvikivi, 2010  
  van Berkum, Koornneef, Otten & Nieuwland, 2007  
  etc.) |
| **2. Not predictable from linguistic structure** |
  Garvey & Caramazza, 1974  
  Koornneef & Sanders, 2012  
  Pickering & Majid, 2007  
  etc.) |
| **3. Largely derived from world knowledge** |
  Pickering & Majid, 2007  
  etc.) |

So we're going to need to revise conventional wisdom.
Re-Mention Biases: Current Consensus

<table>
<thead>
<tr>
<th>Occasional</th>
<th>1. Rapid effects on pronoun interpretation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(Coijn, Commandeur-Vonk &amp; Noordman, 2011</td>
</tr>
<tr>
<td></td>
<td>Featherstone &amp; Sturt, 2010</td>
</tr>
<tr>
<td></td>
<td>Koornneef &amp; Sanders, 2012</td>
</tr>
<tr>
<td></td>
<td>Koornneef &amp; van Berkum, 2006</td>
</tr>
<tr>
<td></td>
<td>Pyykkonen &amp; Jarvikivi, 2010</td>
</tr>
<tr>
<td></td>
<td>van Berkum, Koornneef, Otten &amp; Nieuwland, 2007 etc.)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>2. Not predictable from linguistic structure</th>
</tr>
</thead>
<tbody>
<tr>
<td>Garvey &amp; Caramazza, 1974</td>
</tr>
<tr>
<td>Koornneef &amp; Sanders, 2012</td>
</tr>
<tr>
<td>Pickering &amp; Majid, 2007</td>
</tr>
<tr>
<td>etc.)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>3. Largely derived from world knowledge</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pickering &amp; Majid, 2007</td>
</tr>
<tr>
<td>etc.)</td>
</tr>
</tbody>
</table>

[PAUSE] MOving on
Re-Mention Biases: Current Consensus

1. Rapid effects on pronoun interpretation
   - Cozijn, Commandeur, Vonk & Noordman, 2011
   - Featherstone & Sturt, 2010
   - Koornneef & Sanders, 2012
   - Koornneef & van Berkum, 2006
   - Pyykkonen & Jarvikivi, 2010
   - van Berkum, Koornneef, Otten & Nieuwland, 2007
   - etc.

2. Not predictable from linguistic structure
   - Garvey & Caramazza, 1974
   - Koornneef & Sanders, 2012
   - Pickering & Majid, 2007
   - etc.

3. Largely derived from world knowledge
   - Pickering & Majid, 2007
   - etc.

As I noted, a consensus seems to be emerging that you can't predict implicit causality biases from the structure of the verb.
Re-Mention Biases

“First-Mention Bias”
Sally ate lunch with Mary. She...
Sally went to the store with Mary. She...

“Implicit Causality Bias”
Sally surprised Mary because she...
Sally loved Mary because she...

“Implicit Consequentiality Bias”
Sally surprised Mary so she...
Sally loved Mary because she...

So far as we know, elaboration and serial biases depend very little on semantic content, since they strongly prefer that topics co-refer. However, in explanation or result discourses, as these examples show, listeners build co-reference expectations in part based on the verb. A number of researchers have tried to classify verbs such that their implicit causality and consequentiality biases would be explicable, with most of this work focusing on implicit causality. Unfortunately, almost as soon as a new taxonomy was proposed, exceptions were found. The most recent and best-known attempt comes from Rudolph & Forsterling, who in 1997 proposed grouping verbs into four semantic classes.
<table>
<thead>
<tr>
<th>Direction of Bias</th>
<th>Implicit Causality</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stimulus-Experiencer</td>
<td>Subject</td>
</tr>
<tr>
<td>(frighten, confuse, delight)</td>
<td></td>
</tr>
<tr>
<td>Agent-Patient</td>
<td>Subject</td>
</tr>
<tr>
<td>(call, question, telephone)</td>
<td></td>
</tr>
<tr>
<td>Experiencer-Stimulus</td>
<td>Object</td>
</tr>
<tr>
<td>(like, know, see)</td>
<td></td>
</tr>
<tr>
<td>Agent-Evocator</td>
<td>Object</td>
</tr>
<tr>
<td>(blame, praise, punish)</td>
<td></td>
</tr>
</tbody>
</table>

(Read & Explain) (Consequentiality is on next page)
Now, there are a couple different problems with this hypothesis. You might have noticed that on this hypothesis, all verbs are either subject-biased or object-biased. In fact, as a number of authors have noticed, most verbs are neither.
So here are representative sample of 328 high-frequency verbs, tested by myself and Jesse Snedeker. You can see that most of them show no bias at all.

(Hartshorne & Snedeker, in press, Verb argument structure predicts implicit causality)
A more serious problem, noted by Rudolph & Forsterling in their original article, is that the definitions of the agent–patient and agent–evocator classes

<table>
<thead>
<tr>
<th></th>
<th>implicit causality</th>
<th>implicit consequentiality</th>
</tr>
</thead>
<tbody>
<tr>
<td>stimulus-experiencer</td>
<td>Subject</td>
<td>Object</td>
</tr>
<tr>
<td>(frighten, confuse, delight)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>agent-patient</td>
<td>Subject</td>
<td>Object</td>
</tr>
<tr>
<td>(call, question, telephone)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>experiencer-stimulus</td>
<td>Object</td>
<td>Subject</td>
</tr>
<tr>
<td>(like, know, see)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>agent-evocator</td>
<td>Object</td>
<td>Object</td>
</tr>
<tr>
<td>(blame, praise, punish)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(Rudolph & Forsterling, 1997; Crinean & Garnham, 2006)
are circular. Agent–patient verbs are all and only the action verbs which give rise to subject implicit causality biases, whereas agent–evocator verbs are all and only the action verbs which give rise to object implicit causality biases. Rudolph & Forsterling express hope in their 1997 paper that someday an independent definition would be found, but none was, and after a decade, a number of researchers were ready to call it quits on the whole project, the most emphatic probably being Pickering & Majid 2007.

[PAUSE] Now these taxonomies were largely inspired by theories from social psychology. My graduate advisor Jesse Snedeker and I wondered if the linguistic literature might not be any help in finding ways to classify verbs. In particular, we looked at syntactic verb classes. The jump to syntax may not seem obvious at the moment, but bear with me.

<table>
<thead>
<tr>
<th></th>
<th>direction of bias</th>
<th>implicit causality</th>
<th>implicit consequentiality</th>
</tr>
</thead>
<tbody>
<tr>
<td>stimulus-experiencer</td>
<td></td>
<td>Subject</td>
<td>Object</td>
</tr>
<tr>
<td>(frighten, confuse, delight)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>agent-patient</td>
<td></td>
<td>Subject</td>
<td>Object</td>
</tr>
<tr>
<td>(call, question, telephone)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>experiencer-stimulus</td>
<td></td>
<td>Object</td>
<td>Subject</td>
</tr>
<tr>
<td>(like, know, see)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>agent-evocator</td>
<td></td>
<td>Object</td>
<td>Object</td>
</tr>
<tr>
<td>(blame, praise, punish)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(Rudolph & Forsterling, 1997; Crinean & Garnham, 2006)
There is a large body of work, both in linguistics and psychology, that is focused on that fact that a given verb can appear in some kinds of sentences but not others. Consider the following [CLICK & READ]. Each of these verbs can appear in certain sentence types -- certain syntactic constructions, to use the technical term -- and not others. [PAUSE] There are many dozens of types of sentences; I’m just showing you a smattering here. An important discovery is that when you classify verbs according to the types of sentences they can appear in, you get these classes of verbs, anywhere from a handful to a few hundred verbs, and the verbs in each of these syntactically-defined classes tend to have obvious similarities in terms of meaning. The verbs that can appear in the same sentence types as “hit” are all verbs describing physical contact.
Syntactic Verb Classes

Sally hit Mary.
Sally is hitting Mary.
Sally hit Mary with a stick.
Sally hit that Mary is nice.

Sally frightened Mary.
Sally is frightening Mary.
Sally frightened Mary with a stick.
Sally frightened that Mary is nice.

Sally liked Mary.
Sally is liking Mary.
Sally liked Mary with a stick.
Sally liked that Mary is nice.

Sally hit Mary.
Sally is hitting Mary.
Sally hit that Mary is nice.

see also

touch
touch
knight
love

kick
tap
hate

etc.

love
fear
delight

etc.

surprise
confuse
etc.
As I mentioned earlier, Jesse and I collected implicit causality biases for 328 high-frequency verbs, and got this distribution. Now, it turns out that there are about 280 syntactic verb classes, so in most cases, we didn’t have many verbs in any particular class, but by luck, we had a few classes with a reasonable number of verbs. When we looked at those, each one had systematic effects.
Here are the results for five of those classes. Here, again, I’m plotting data by-verb, so these are box plots of the results for each verb. You can see we have two classes that are systematically object-biased, one that is strongly subject-biased, and two that have no strong bias. I’ve given examples of each class here, but you can look them up by number on VerbNet, which is a free Internet database of syntactic verb classes.

Of course, this was a post hoc analysis. Maybe we just lucky with the verbs we tested. [PAUSE] We took two of these classes and tested all of the verbs in them.
Since there are only two classes, I’m back to using histograms. You can see that the effects are pretty systematic and there is basically no overlap. BTW I haven’t checked systematically, but most of the spread you see here is due to noise. There is little evidence of systematic variation within class, though I can’t rule it out at this time. Certainly, most of the variance is accounted for by verb class and noise. [PAUSE] That was implicit causality. What about implicit consequentiality? I recently collected implicit consequentiality biases for all the monosemic verbs in the syntactic birth classes for just described as well as a couple more, for 502 for total.
So it looks like re-mention biases pattern with syntactic verb class. Can we make any sense of why some classes go one way and others go another way?
It is called implicit causality, but really these sentences are about explanations. Explanations usually refer to causes. Most researchers agree that whether how a verb marks causation plays a central role in which syntactic verb class a verb appears in.
Looking at the implicit causality biases, the one subject-biased class, 31.1, which contains emotion verbs like surprise, delight, and confuse, is argued by many researchers to entail that the subject of the verb caused the emotion: frighten means “cause to feel fright”. Now, it makes sense to explain things in terms of their causes, and since the subject of the verb has just been explicitly marked as playing a causal role, a reasonable listener might expect that the explanation will also refer to that causal agent, which gives us the observed implicit causality bias.

In contrast, there is good reason to believe that the two object-biased classes -- 31.2, which contains emotion verbs like fear, like, and love, and 33, which contains verbs of judgment like blame and praise, describe events which happened in reaction to some property of or previous action by the verb’s object. You blame someone or fear them because of something they did. So this actually fits the implicit causality biases. As far as the remaining two classes ... they aren’t discussed much in the literature. We’re currently working on determining their causal structure; ask me in a year.
Looking at implicit consequentiality, the second sentence describes some consequence of the event in the first sentence. Many researchers believe that one of the semantic features that guides syntactic verb class is the notion of affectedness. Some verbs describe an event in which one of the participants is affected by the event in some way. And if we were talk about the consequences of an event, it’s not unreasonable to expect those consequences to have something to do with the affected entity.
Of our seven classes, five of them were object-biased, and all of them describe events in which the object of the verb was clearly the one most affected. If Sally dared Mary, weakened Mary, blamed Mary, appealed to Mary, or surprised Mary, Mary is in each case changed by the event. And so the implicit consequentiality bias follows. In contrast, if Sally marvels at Mary or fears Mary, it’s Sally that’s having the emotional state, she’s the one affected, and so it makes sense to talk about the consequences to Sally.
Re-Mention Biases: Current Consensus

1. Rapid effects on pronoun interpretation
   (Cozjin, Commandeur, Vonk & Noordman, 2011
    Featherstone & Sturt, 2010
    Koornneef & Sanders, 2012
    Koornneef & van Berkum, 2006
    Pyylkonen & Jarvikivi, 2010
    van Berkum, Koornneef, Otten & Nieuwland, 2007
    etc.)

2. Not predictable from linguistic structure
    Garvey & Caramazza, 1974
    Koornneef & Sanders, 2012
    Pickering & Majid, 2007
    etc.)

3. Largely derived from world knowledge
    Pickering & Majid, 2007
    etc.)

So far, so good. Now, there are 280 verb classes, and we’ve only looked at a handful. It’s an important handful, covering the vast majority of the verbs in every previous implicit causality or consequentiality study in English. Nonetheless, it remains to be seen whether the pattern of results generalizes and whether the theoretical account I just sketched out will account for future data. Still, these are data that any theory will need to account for.
Re-Mention Biases: Current Consensus

1. Rapid effects on pronoun interpretation
   (Cozijn, Commandeur-Vonk & Noordman, 2011
   Featherstone & Sturt, 2010
   Koornneef & Sanders, 2012
   Koornneef & van Berkum, 2006
   Pyykkonen & Jarvikivi, 2010
   van Berkum, Koornneef, Otten & Nieuwland, 2007
   etc.)

2. Not predictable from linguistic structure
   Garvey & Caramazza, 1974
   Koornneef & Sanders, 2012
   Pickering & Majid, 2007
   etc.)

3. Largely derived from world knowledge
   Pickering & Majid, 2007
   etc.)

[PAUSE] Now moving on to the last alleged fact
Re-Mention Biases: Current Consensus

1. **Rapid effects on pronoun interpretation**
   (Cozijn, Commandeur, Vonk & Noordman, 2011
   Featherstone & Sturt, 2010
   Koornneef & Sanders, 2012
   Koornneef & van Berkum, 2006
   Pyykkönen & Jarvikivi, 2010
   van Berkum, Koornneef, Otten & Nieuwland, 2007
   etc.)

2. **Not predictable from linguistic structure**
   Garvey & Caramazza, 1974
   Koornneef & Sanders, 2012
   Pickering & Majid, 2007
   etc.)

3. **Largely derived from world knowledge**
   Pickering & Majid, 2007
   etc.)

By the way, most of this work is now in press at LCP, and you can find the paper on my website in the meantime. So just to be clear what I mean about effects of world knowledge.
Sally frightened Mary so she ran away.

[READ] [VOTE] Here, I take "she" to refer to Mary
Sally frightened Mary so she ran away.

Whereas if
Sally frightened Mary so she ran away.

Sally frightened Mary so she got in trouble.
Sally frightened Mary so she ran away.

Sally frightened Mary so she got in trouble

[PAUSE] I take it as uncontroversial that world knowledge is playing a role here, that we think it's more likely that a [PAUSE] There is a separate question about whether the discourse expectations
Remember this slide? [PAUSE] I take it as uncontroversial that world knowledge is playing a role here, that we think that frightenees are more likely to run away than frighteners, and that you get in trouble for frightening someone, not for being frightened. [PAUSE] There is a separate question about whether the discourse expectations
we’ve been focusing on depend on world knowledge. A lot of authors seem to think so. One important piece in the argument was that implicit causality and consequentiality biases couldn't be predicted by linguistic structure. The data I have just shown you I think weakens this argument. But even if verb argument structure influences re-mention biases, that does not mean they are sufficient. Pickering & Majid (2007) make the most compelling case. Here's what they actually say
How then do people compute the inference of implicit causality? Various components of the verb’s meaning are of course important [including] how enduring the event is, how concrete it is, whether it is telic or not (Semin & Fiedler, 1988; Rudolph & Forsterling, 1997), and how negative its connotative meaning is (Semin & Marsman, 1994). In addition, properties of the participants affect implicit causality. Changing the gender (Lafrance, Brownell, & Hahn, 1997), animacy (Corrigan, 1988, 1992), or typicality (Corrigan, 1992; Garvey et al., 1976 [sic]) of the participants changes the implicit-causality bias, as do contextual factors that affect focus (Majid, Sanford, & Pickering, 2006) … All of these factors affect the construction of the event representation, and it is this event representation that is used to infer the cause. (Pickering & Majid, 2007, pp. 785-86, emphasis added).

[READ] So this seems very compelling, except for one very important limitation: most of these studies didn’t explicitly test re-mention biases. These data mostly come from a task pioneered in 1983 by Roger Brown and Deborah Fish.
They were looking at sentences like “Sally frightened Mary”. On reflection, it seemed to them that while many people are capable of being frightened, only a select few are scary, so really this sentence probably tells you more about Sally than about Mary. They then designed a task to see whether this is really the case.
Brown & Fish Causal Attribution Task

Sally frightened Mary
How likely is it that this is because:
A. Sally is the kind of person that frightens people.  
   UNLIKELY 1 2 3 4 5 6 7 8 9 LIKELY

(Brown & Fish, 1983)
Brown & Fish Causal Attribution Task

*Sally frightened Mary*

How likely is it that this is because:

A. Sally is the kind of person that frightens people.
   - UNLIKELY 1 2 3 4 5 6 7 8 9 LIKELY

B. Mary is the kind of person that people frighten.
   - UNLIKELY 1 2 3 4 5 6 7 8 9 LIKELY

[READ & EXPLAIN] [CLICK] The typical analysis is to take the answer given for the subject and subtract the answer given for the object. So here we have a subject–bias of 8 – 2 = 6.

[PAUSE]

If you try other verbs, you get other results. They argued that this is actually the notion of causality that underlies the pronoun interpretation effect — that Sally’s participation in this event tells you more about Sally than Mary’s participation tells you about Mary — and so they named their effect “implicit causality” as well. Although the claim has been more or less universally adopted, Brown and Fish never provided any evidence that this task really *does* tap the same representations as the pronoun resolution phenomenon, and the little evidence that has been published has been mixed. And when you think about it, while this is a really interesting task, it’s not obvious how much it has to do with causality.
Suppose Sally kicked a soccer ball. Is that because she is the kind of person who kicks things? [CLICK] I’m not sure. Is it because the soccer ball is the kind of thing people kick? Well, that is part of what makes a soccer ball a soccer ball, [CLICK] so I suppose so. So Brown and Fish would conclude that the ball caused Sally to kick it, but I’m not sure that’s really the conclusion we want to make.

[PAUSE] Most of the work suggesting that your world knowledge about the event participants affects your re-mention biases involves manipulation of gender or social status, and most of it was done using the Brown & Fish causal attribution task. Of the pronoun studies that have been run, one found an effect of gender but two did not. There haven’t been any re-mention studies that manipulated social status other than one preliminary study by Garvey & Caramazza involving only 5 items and mixed results. So I set out both to test whether manipulations of gender and social role affect pronoun interpretation AND to see what the relationship between the pronoun task and the causal attribution task is.
The first experiment looked at a social hierarchy manipulation. I tested 12 subject-biased and 12 object-biased verbs. Social rank of even participants was manipulated within and between subjects -- that is, some participants would see “the duke frightened the butler” and some “the butler frightened the duke”. Half the participants were tested using the causal attribution task that I just described. The other half completed a pronoun resolution task. The stimulus lists were otherwise the same across the 2 tasks, except that in the pronoun task the sentences were continued with the phrase “because he is a dax”, and participants’ pronoun interpretation was assessed by asking who they thought the “dax” was, just as in Paper 1. [PAUSE] Now if we get the kind of results previously reported in the literature, this is what we should see
What I’m doing here is calculating the “high-status bias” for each verb: the degree to which it is more subject-biased when the subject is high-status versus low-status. [PAUSE] This is scaled so that +100 is the strongest measurable high-status bias and −100 is actually the strongest low-status bias (that is, the exact opposite of the predictions). This plot shows the imaginary results for individual verbs, and you can see all of them show a high-status bias in both tasks.
Here are the actual results. [PAUSE] The prediction was born out for the causal attribution task -- people did assign more causality -- as defined in that task -- to the subject when the subject was high-status as opposed to low-status. However, this same manipulation did not produce an overall effect on pronoun interpretation. [PAUSE] You might think that perhaps the distribution is just shifted for the pronoun task, but actually when you compare the effect of the status manipulation verb-by-verb across the two tasks, there is no correlation. [PAUSE] So that’s the social rank manipulation. What about gender?
Gender Manipulation

12 subject-biased verbs
12 object-biased verbs

Gender of event participants manipulated within & between subjects
(John frightened Sally vs. Sally frightened John)

A. Causal attribution task (N=96)

Sally frightened John.
How likely is it that this is because
1) Sally is the kind of person who frightens people
   UNLIKELY 1 2 3 4 5 6 7 8 9 LIKELY
2) John is the kind of person whom people frighten
   UNLIKELY 1 2 3 4 5 6 7 8 9 LIKELY

-MANIPULATING GENDER-

B. Pronoun resolution task (N=96)

-Sally frightened John because... What is the next word?: she he

I used the same verbs, but this time I manipulated gender instead of social rank [PAUSE] The causal attribution task was otherwise the same. The pronoun interpretation task had to be a bit different, because of course the gender of the pronoun forces you to interpret pronouns one way or another. The way studies have gotten around this is by presenting sentence fragments up through the word because -- “Sally frightened John because...” and asking participants to fill in the blank. I adapted this method by asking the participant just to pick the next word, forcing them to choose between “he” and “she”. We’ll see that this in fact gives very similar results to those of the other pronoun resolution task. [PAUSE]
Again, here is what the predicted results look like. Again, I’m computing for each verb the degree to which it is male-biased: the degree to which it is more subject-biased when the subject is male relative to when the subject is female. According to the standard prediction, this is what we should see across verbs and in both tasks.
Here is what we actually see. It is not the case that, on average, there is a male bias. Could it be that some verbs are male-biased, others female-biased, and others not? Perhaps, though as you can see the effect size here is small even for the outliers. Moreover, there is no correlation in the degree of male bias for each verb across the two tasks, so if gender if affecting these tasks, it’s doing so differently for each. 

Incidentally: what is the relationship between causal attribution bias and pronoun interpretation bias? I mentioned that it hadn't been studied that systematically. I correlated the biases for each verb across the two tasks.
Here is what the predicted results look like if there was a perfect correlation. I’ve got the two experiments graphed separately because of course their stimuli were quite different. [PAUSE] On the x-axis, I've got subject bias in the pronoun tasks, whereas on the y-axis, I've got subject bias for the Brown & Fish causal attribution task. [PAUSE] Here are the actual results.
Do both tasks measure the same bias?

[PAUSE] There actually is some relationship in the results between the two tasks, but it’s not very strong. [PAUSE] It looks particularly bad if you classify these verbs into subject- and object-biased using a binary division.
Do both tasks measure the same bias?

social rank experiment

gender experiment

same bias on both tasks: 71% 58%

[PAUSE] You’ll notice in particular that there are a lot of items in the bottom right corner [CLICK]. These are items which are subject–biased on causal attribution but object–biased on pronoun resolution. When you look at how many verbs show the same bias on both tasks [CLICK] it’s 71% for the social rank experiment and 58% for the gender experiment. And chance is at 50%, so that's not great. [PAUSE] this is very different that what you get if you compare the biases across the two pronoun tasks.
Subject Bias: Production vs. Comprehension

production: Sally frightened John because...she/he
comprehension: Sally frightened Mary because she was a froom

keep in mind, one of them was comprehension and the other was production, so there are a lot of differences between these tasks [PAUSE]
Subject Bias: Production vs. Comprehension

production: Sally frightened John because...she/he
comprehension: Sally frightened Mary because she was a froom

Hartshorne (in press) What is implicit causality?

[PAUSE] these correlations were at about .9, which is near ceiling as far as correlations go.

note that the paper this graph comes from was also interested in verb classes and whether these effects were different across four verb class. So that's why I have four verb classes there. You can ask me about that later or look up the paper if you are interested, in the meantime [FILL IN]
How then do people compute the inference of implicit causality? Various components of the verb’s meaning are of course important [including] how enduring the event is, how concrete it is, whether it is telic or not (Semin & Fiedler, 1988; Rudolph & Forsterling, 1997), and how negative its connotative meaning is (Semin & Marsman, 1994). In addition, properties of the participants affect implicit causality. Changing the gender (Lafrance, Brownell, & Hahn, 1997), animacy (Corrigan, 1988, 1992), or typicality (Corrigan, 1992; Garvey et al., 1976 [sic]) of the participants changes the implicit-causality bias, as do contextual factors that affect focus (Majid, Sanford, & Pickering, 2006) … All of these factors affect the construction of the event representation, and it is this event representation that is used to infer the cause. (Pickering & Majid, 2007, pp. 785-86, emphasis added).

So in the paper we go through every finding cited in this quote, as well as some not cited in the quote, and show that every one of them doesn’t actually apply to re-mention biases. Again, the paper is in press in LCP and can be found on my website, so you can look it up there.
Re-Mention Biases: Current Consensus

1. Rapid effects on pronoun interpretation
   (Cozijn, Commandeur-Vonk & Noordman, 2011
    Featherstone & Sturt, 2010
    Koornneef & Sanders, 2012
    Koornneef & van Berkum, 2006
    Pyykkonen & Jarvikivi, 2010
    van Berkum, Koornneef, Otten & Nieuwland, 2007
    etc.)

2. Not predictable from linguistic structure
    Garvey & Caramazza, 1974
    Koornneef & Sanders, 2012
    Pickering & Majid, 2007
    etc.)

3. Largely derived from world knowledge
    Pickering & Majid, 2007
    etc.)

and so so much for the third myth
Re-Mention Biases: Current Consensus

1. *Rapid effects on pronoun interpretation*  
   (Cozijn, Commandeur, Vonk & Noordman, 2011  
    Featherstone & Sturt, 2010  
    Koornneef & Sanders, 2012  
    Koornneef & van Berkum, 2006  
    Pyykkonen & Jarvikivi, 2010  
    van Berkum, Koornneef, Otten & Nieuwland, 2007  
    etc.)

2. *Not predictable from linguistic structure*  
    Garvey & Caramazza, 1974  
    Koornneef & Sanders, 2012  
    Pickering & Majid, 2007  
    etc.)

3. *Limited effects of world knowledge*  
    Pickering & Majid, 2007  
    etc.)

[continue]
Re-Mention Biases: Current Consensus

1. Rapid effects on pronoun interpretation
   (Cozjin, Commandeur-Vonk & Noordman, 2011
   Featherstone & Sturt, 2010
   Koornneef & Sanders, 2012
   Koornneef & van Berkum, 2006
   Pyykkonen & Jarvikivi, 2010
   van Berkum, Koornneef, Otten & Nieuwland, 2007
   etc.)

2. Not predictable from linguistic structure
   Garvey & Caramazza, 1974
   Koornneef & Sanders, 2012
   Pickering & Majid, 2007
   etc.)

3. Largely derived from world knowledge
   Pickering & Majid, 2007
   etc.)

[PAUSE] So all this disappointing if you liked the myths. Which I did! I got into this line of research because I was excited about understanding how world knowledge rapidly guided linguistic interpretation. I know I just presented a lot of data, but that was only a fraction of the studies we ran, trying to make my findings go away, trying to preserve the myths. We live in the world we have, not the one we want. So it was sad...except that, on thinking about it more, the world we have is actually pretty exciting, too. I mean, we really don’t understand why some re-mention biases are fast and others slow. On many of the currently dominant accounts of sentence processing, that really shouldn’t be possible. I can explain why in the Q&A. In any case, it suggests there is something big out there that we currently know nothing about. Which is exciting.

I guess massive effects of world knowledge on these discourse expectation biases would have been exciting, but they had their drawbacks. Think about that Pickering & Majid quote. It was just a list of effects. There hadn’t been any real progress in developing a coherent, predictive theory, and it’s hard to even imagine how you’d go about doing that.

In contrast, I think now that by relating implicit causality and consequentiality biases to the massive verb semantics literature, I think we’re making real progress and that there are clear avenues for further development of the theory. And I think whenever you have two well-developed literatures meet up, where there’s real synergistic potential, I think that’s exciting, I think it suggests we, as a field, are really making some progress. If our theories of re-mention biases and of verb argument structure were completely off track, there would be no synergy.
*Huge debt to Hannah & Andy*

With that, I’d like to thank you for listening, and I look forward to your questions & comments.