It's worth doing a bit of a deeper dive on the psycholinguistic models that have been developed to account for the repeated name penalty, to see how they make contact with, or differ from, the model we are developing here. Perhaps most prominent example is Gordon and Hendrick's Discourse Prominence Theory.

DPT attributes the repeated name penalty to the processes of incremental referential and discourse processing that are automatically engaged as we process sentences in context. Importantly, this model offers an explanation of the RNP that has nothing to do with pragmatics at all. Instead, it locates the difficulty in the RNP squarely at the syntax-discourse interference, a side-effect of how different types of DPs interact with the discourse. So it'll be useful to see what it's all about.

0.1 DRT

DPT takes as its jump-off point Discourse Representation Theory, originally developed by Hans Kamp (e.g. Kamp & Reyle, 1993). DRT offers a formalism for modeling the mental representations associated with building a discourse incrementally as the sentence is interpreted left-to-right. There are two essential concepts we have to have in mind to understand how this works:

(1) **Discourse representation structures**: A representation of the hearer's discourse model. It consists of a set of discourse referents, and DRS conditions. DRS conditions are the information the hearer has processed about each of the discourse referents.

(2) **Construction Rules**: Rules that translate from syntactic structures to DRSs.

The construction rules are triggered by recognizing certain syntactic structures, and the application of these rules results in the creation and updating of the DRS.

Here's a sample. Let's take Gordon and Hendrick's Construction Rule for proper names (Gordon & Hendrick 1998, p. 400):

With this construction rule, we can sketch the incremental processing of a simple sentence like *John likes Mary*:
Here, the discourse model is incrementally elaborated as the listener recognizes Proper Name NPs; upon recognizing these items in the input string, the construction rule is automatically triggered. This has two effects. It introduces a new discourse entity into the discourse model, and it introduces a new discourse condition. For example, encountering John in the input results in the creation of a new discourse entity in a discourse model that has the condition $\text{John}(x)$. The processing of the rest of the sentence proceeds similarly (setting aside how the verb is parsed), and results in the creation of the following DRS:

\[
\begin{array}{c}
x & y \\
\text{John (x)} \\
\text{Mary (y)} \\
x \text{ sees } y
\end{array}
\]

So, pronouns are primarily anaphoric in that their construction rule commits them to an immediate instruction to choose a suitable antecedent in the DRS. What makes a suitable antecedent is that the presuppositional content of the pronoun (e.g. that its referent has such and such gender features, or animacy features) is satisfied.

Consider the sentence Jane thinks that she is sick. The application of the Proper Name construction rule, followed by the Pronoun construction rule, would yield the following DRS:

\[
\begin{array}{c}
x \\
\text{Jane (x)} \\
x \text{ thinks } x \text{ is sick}
\end{array}
\]

All normal.

0.3 Names where they shouldn’t be

Now let’s look at what happens when a name is repeated, as in Jane thinks Jane is sick. Since both DPs are names, they both trigger the same construction rule, which yields the following DRS:

\[
\begin{array}{c}
x \\
\text{Jane (x)} \\
x \text{ thinks } x \text{ is sick}
\end{array}
\]

All normal.
This DRS represents a model of the discourse where there are two different individuals, both of whom share the same properties. The second name has been interpreted as disjoint with the first, as there are two non-identical discourse referents. However, the situation where there are two discourse referents that share discourse conditions within a given span of the input triggers another construction rule. This construction rule modifies the DRS and equates the two discourse referents.

<table>
<thead>
<tr>
<th>x</th>
<th>y</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jane (x)</td>
<td>Jane (y)</td>
</tr>
<tr>
<td>x thinks y is sick</td>
<td></td>
</tr>
</tbody>
</table>

Now, we should probably be a little skeptical of the condition that triggers this rule. Depending on how we elaborate the discourse conditions that are entered into the DRS, this specific formulation of the triggering conditions could be deleterious. For example, one might worry that this could get triggered if two antecedents share only one property that is represented in the DRS (e.g. suppose that both antecedents are associated with presuppositions such as `female(x)`; this would trigger this rule, inappropriately); something more might need to be said once we expand this out beyond Proper Names.

The application of this rule leads to the updated DRS:

<table>
<thead>
<tr>
<th>x</th>
<th>y</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jane (x)</td>
<td>Jane (y)</td>
</tr>
<tr>
<td>x = y</td>
<td></td>
</tr>
<tr>
<td>x thinks y is sick</td>
<td></td>
</tr>
</tbody>
</table>

(I believe, but I am not sure, that there is a mistake in the above DRS, and that Jane(y) should not be there.)

The reason, according to DPT, that Jane thinks Jane is sick is degraded is because of the additional construction rules that are applied to arrive at the coreferential interpretation.

Now, consider She thinks Jane is sick: here, coreference is very difficult to achieve. The linear order of the pronoun and name, coupled with the construction rules we have on the table, lead to a very different outcome. First, the pronoun has no suitable antecedent in the DRS, and so it can instantiate a new discourse referent. The name is processed as usual. The following DRS results:

<table>
<thead>
<tr>
<th>x</th>
<th>y</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jane (y)</td>
<td></td>
</tr>
<tr>
<td>x thinks y is sick</td>
<td></td>
</tr>
</tbody>
</table>

In this way, disjoint reference is automatically and immediately computed as a consequence of the construction rules already at play. Notice now that at this point there is no mechanism for achieving coreference, which is DPT’s explanation for why coreference is so difficult to achieve in these examples: not Principle C.

So there we have it: DPT’s explanation for why Jane thinks Jane is sick often yields a perception of coreference and feeling of unacceptability,
and why *She thinks that Jane is sick* instead yields a perception of disjoint reference and no acceptability penalty.

Their model predicts a cline of ‘coreferential possibility’ across the three examples we’ve seen:

(4) Name - pronoun: Jane thinks that she is sick.
(5) Name - Name: Jane thinks that Jane is sick.
(6) Pronoun - name: She thinks that Jane is sick.

That matches both intuition, as well as experimentally collected judgments of coreference on similar examples (Gordon & Hendrick, 1997):

<p>| Sample stimuli for Experiment 2 |</p>
<table>
<thead>
<tr>
<th>NP₁</th>
<th>NP₂</th>
<th>C-command</th>
<th>Sample sentence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Name</td>
<td>Pronoun</td>
<td>No</td>
<td>Before Susan began to sing, she stood up.</td>
</tr>
<tr>
<td>Name</td>
<td>Pronoun</td>
<td>Yes</td>
<td>Susan stood up before she began to sing.</td>
</tr>
<tr>
<td>Name</td>
<td>Name</td>
<td>No</td>
<td>Before Susan began to sing, Susan stood up.</td>
</tr>
<tr>
<td>Name</td>
<td>Name</td>
<td>Yes</td>
<td>Susan stood up before Susan began to sing.</td>
</tr>
<tr>
<td>Pronoun</td>
<td>Name</td>
<td>No</td>
<td>Before she began to sing, Susan stood up.</td>
</tr>
<tr>
<td>Pronoun</td>
<td>Name</td>
<td>Yes</td>
<td>She stood up before Susan began to sing.</td>
</tr>
</tbody>
</table>

Table 3

| Overall results for Experiment 2. The table shows the proportion of acceptable sentences as a function of c-command relation and type of NP sequence |
|-----------------|-----------------|
| No c-command | C-command |
| NP₁ (name) | NP₂ (pronoun) | NP₁ (name) | NP₂ (pronoun) |
| NP₁ (name) | 49 | 88 | .31 | .04 |
| NP₁ (pronoun) | 90 | | 99 | |

Where the probability of coreference (here labeled ‘acceptability’) tracks intuition. Interestingly, this did not hold for examples like *He asked John’s roommates about the game*, which tolerated coreference more than did *John asked John’s roommates about the game*.

0.4 Getting prominence in there

The antecedent’s syntactic position vis-a-vis the anaphor evidently matters. DPT accounts for this by positing that the list of discourse referents in the DRS are represented as an ordered set \{\(\alpha_1, \alpha_2, \ldots, \alpha_n\)\}, such that for each \(i < j\), \(\alpha_i\) is more accessible than \(\alpha_j\). This is an idea they inherit from centering theory, which holds that discourse referents are entered into such a ranked list, as a function of where they occur in the surface syntax.

Given such an accessibility ordering over the list of discourse referents, Gordon and Hendrick propose that the pronoun construction rule is modified as follows:

This rule proposes that the search for a discourse referent is ranked by accessibility, and you choose the most accessible one. There is controversy over how literally we should take ‘search’ as a psychological claim (see Foraker & McElree, 2007; Kush, Johns & van Dyke, 2019), but Gordon and Hendrick note that even if we reinterpret their search clause as probabilistic random access gated by the accessibility ranking and match to features of the pronoun, as in the Parker & Phillips model. They argue that the two are not empirically indistinguishable, which is not accurate. They may be distinguished by measuring the fine time course of how an antecedent is accessed (see McElree, 2006), although Gordon and Hendrick’s point that the distinction is not entirely critical to their theoretical claims is well taken.

The construction rule that equates elements that are coreferent in the DRS is also sensitive to the discourse referent ranking, but in the opposite way from the pronoun rule. Here, the search for a matching discourse referent proceeds in the opposite manner. It proceeds bottom to top in the referent ranking:

(41) CR.EQ (revised)
This is meant to account for the finding that the repeated name penalty is more severe for more prominent antecedents. On this model, it is because more searching is needed to find the more prominent antecedent in the discourse model, which means that the probability of failure is higher.

The DPT has some clear strengths as a model of the repeated name penalty: it’s fairly explicit, and it pins the RNP to normal processes involved in incrementally establishing (co-)reference across a discourse. It makes predictions both about the probability of achieving a disjoint interpretation of a referring expression, but it also yields directly yields a prediction about reading time and other processing measures: it is important to underline this, because the name ‘repeated name penalty’ seems to me to be vague with respect to the type of penalty we have in mind. There could be a penalty against coreference, in which case disjoint reference arises. There could also be a ‘penalty’ in the sense of processing difficulty.

DPT predicts both. It predicts that there should be localized processing difficulty at the embedded Jane in Jane thinks that Jane is sick, but not in She thinks that Jane is sick; this is a prediction that is not very clearly met in the empirical literature. The effect is present in on-line eye-tracking measures but not very precisely localized (Ledoux et al., 2007; Kennison & Gordon, 1997), and as we’ve seen, there is some controversy about whether we see any processing effect in She thinks that Jane is sick. If anything (c.f. the Drummer and Felser paper!), it seems like we might even expect that example to be processed more quickly than a control where gender features rule out a match between the pronoun and the embedded name (e.g. He thinks that Jane is sick).

But it’s got some drawbacks. Chief among them, I think, is that it does not give an adequate explanation for definite descriptions; it is stipulated that they are treated like proper names in that they automatically introduce new entities into the DRS. That works, but at the cost of us taking on board some pretty unhappy assumptions about what we do with definite descriptions during routine comprehension!

Relatedly, it does not predict any differences beyond a coarse cut between pronouns and other types of referring expressions. However, the repeated name penalty, as Kyle showed in the first class, has little to do with repeated names: it occurs (sometimes even more acutely!) when there is non-repeated material, and it occurs for definite descriptions. I’m not sure we have that much to say about definite descriptions yet in this system. We do have a way of capturing some of the variable force of the RNP, however. Recall some of Kyle’s examples, slightly modded:

(7) ? A linguist walked into a bar. The linguist ordered a Corpse Reviver
(8) * A linguist walked into a bar. Brian ordered a Corpse Reviver

The construction rule that equates two discourse entities will only be triggered in the first example, because only in that case is there identity in the predicates that hold of each discourse referent introduced by the antecedent and its anaphor: linguist(x). That doesn’t hold in the second case; this means that this construction rule is not predicted to be triggered. This means that i) it is predicted to have a strong preference for disjoint reference and ii) there should be no processing difficulty associated with Brian.

References