Pronoun Resolution and The Influence of Syntactic and Semantic Information on Discourse Prominence

Ralph Rose

Gunma Prefectural Women's University
Gunma Prefecture, Japan
rose@gpwu.ac.jp

Abstract. Beginning with the observation that syntactic and semantic information often coincide (i.e., subjects are often agents, objects often patients), this study investigates the possibility that preference to resolve a sentence-initial pronoun to a syntactically prominent antecedent might actually be better explained in terms of preference for resolving to a semantically prominent antecedent. The study takes Discourse Prominence Theory (Gordon and Hendrick [11, 12]) as an underlying framework. Results of three psycholinguistic experiments using a self-paced reading task show that *both* syntactic and semantic information guide readers' pronoun resolution preferences. This suggests a revised understanding of Discourse Prominence Theory in which the prominence of discourse referents is determined through a complex process depending on multiple linguistic factors. Results further show that the relative degree of prominence among competing candidates influences resolution processes.

Key words: pronoun resolution, Discourse Prominence Theory, repeated-name penalty

1   Introduction

Most pronoun resolution algorithms incorporate some method (explicitly or implicitly) for ranking candidate antecedents with higher-ranking candidates judged more likely to be the intended antecedent. One factor which practically all of these ranking schemas share is some measure of the syntactic prominence of candidate antecedents. In Lappin and Leass' Resolution of Anaphora Procedure [28], for instance, candidates are assigned a certain index value based on their grammatical role (subject, object, etc.). Hobbs' algorithm [19], on the other hand, employs a hierarchical search of the syntactic representation, effectively ranking candidates according to the syntactic structure. A simpler procedure is proposed by Gernsbacher and Hargreaves [7] using linear order-of-mention.

However, in English, syntactic and semantic information are often conflated: Syntactic subjects are often semantic agents while syntactic objects are often semantic patients. Thus, it is conceivable that the contribution of syntactic prominence to the ranking of candidates is better explained as the contribution of

semantic prominence. The present study examines this possibility in a series of psycholinguistic experiments designed to disambiguate the effects of syntactic and semantic prominence in pronoun resolution preferences.

This paper is laid out as follows. In Section 2, I give an overview of some theoretical issues underlying the present research. Based on many existing models, I describe a general ranking schema for candidate antecedents in pronoun resolution. In Sections 3 and 4, the experiments are described. This is followed in Section 5 with some general discussion of the findings and their implications.

2   Background

2.1   Discourse Prominence Theory

In this paper, I assume as an underlying framework *Discourse Prominence Theory* (hereafter, DPT) introduced in Gordon and Hendrick [11, 12]. While cast in terms of Discourse Representation Theory (DRT) [22, 23], it is intended to be a general model of discourse processing which captures observations about how readers interpret reference and coreference in a discourse as well as the time-course of processing reference and coreference. Entities introduced in the discourse are referred to as *discourse referents* within the discourse representation, following Heim [16, 17] and Karttunen [24]. The cumulative representation of the discourse thus far—that is, the context—is then seen to contain two things: a list of discourse referents and a list of semantic conditions on those referents. In this paper, I will be centrally concerned with the list of referents and how it is utilized during discourse processing.

In DPT, each new utterance is processed and incorporated into the representation with respect to the current context [25, 27, inter alia]. As various linguistic objects or configurations are detected by the parser, corresponding operations are triggered which may access the context in order to be completed. Here I'll discuss three crucial operations in DPT. In DRT, these operations are called *constructions rules* (hereafter, CR) because they are operations that contribute to the construction of the discourse representation. First, when a proper name is encountered, a construction rule is triggered which introduces a new discourse referent into the representation (hereafter, CR.PN). Second, when a pronoun is encountered, a construction rule is triggered to search for a suitable referent in the context and then establish coreference with it (CR.PRO). Third, when it can be concluded from the semantic conditions that two independent discourse referents in the representation refer to the same real-world entity (i.e., corefer), then an operation is triggered to establish this equivalence in the representation (CR.EQ).[1]

For instance, consider the sentences in (1)-(2). Experimental evidence [11] shows that readers find it considerably easier to establish coreference between the name and pronoun in (1) than between the two names in (2) and process the

---

[1] See [12] for a formal definition of their three constructions rules: CR.PN, CR.PRO, and CR.EQ.

former sentence faster than the latter. In DPT, this is readily explained. The first occurrence of the name in both sentences triggers CR.PN which introduces a new discourse referent, say x, into the representation. Then the pronoun in (1) triggers CR.PRO which easily finds a suitable referent, x, and establishes coreference with it. However, the repeated name in (2) triggers CR.PN and introduction another new discourse referent, say y. Subsequently, the semantic information showing that x and y refer to the same real-world entity (i.e., *Jane(x)*, *Jane(y)*) trigger CR.EQ which then establishes equivalence between x and y. Thus, the additional operation accounts for the increased difficulty readers have with such a configuration.

(1)   Jane$_i$ thinks she$_i$ is sick.
(2)   Jane$_i$ thinks Jane$_i$ is sick.

In Gordon and Hendrick's description of DPT, they also imply the necessity of a further construction rule to handle cases where a pronoun has been incorrectly assigned. For instance, consider (3).

(3)   a. John sent a package to Bill.
       b. He received it two days later.

In DPT, the pronoun *he* in (3b) is initially interpreted as coreferent with the discourse referent introduce by John in the preceding sentence (because it is syntactically most prominent—discussed in further detail below). However, the more natural interpretation due to plausibility constraints is that the pronoun refers to Bill. Therefore, some sort of reanalysis operation must take place in order to 'repair' the discourse representation. I assume this operation is triggered by the recognition of some sort of inconsistency in the semantic conditions though I will leave an explicit description of this to future work. In this paper, I will refer to this construction rule as CR.RA.

In the present study, I am particularly interested in exploring how CR.PRO proceeds. Gordon and Hendrick do not discuss in great detail how the discourse processor determines what is a suitable referent, though they do seem to assume that referents introduced in syntactically more prominent positions are more suitable than those introduced in less prominent positions. In the following section I will discuss a general model of how the processor determines which referent is a suitable referent.

2.2   Pronoun Resolution

Most models of pronoun resolution incorporate two primary operations toward determining a suitable referent for a pronoun: a filtering operation and a ranking operation which take place over the referents in the context. The filtering operation removes from consideration referents which are morphosyntactically incompatible with the pronoun under consideration [2, 4, 28]. The ranking operation orders the referents with respect to some criteria. This ranking can then be seen to reflect the degree of likelihood that each referent is the suitable referent

for the current pronoun. In DPT, this ranking is referred to as the discourse prominence of a referent. In other theories and formalisms, such terms as 'focus' [35], 'givenness' [15], 'topichood' [8], 'salience' [3], and 'centering' [14, 13] describe similar or overlapping conceptualizations.

The central question surrounding the ranking operation is the actual procedure and criteria for determining this ranking. In Gordon and Hendrick's description of DPT, the discourse prominence of referents is determined entirely by syntactic information. While this might be a convenient simplification, it is surprising they do not propose a more flexible approach which depends on numerous types of information because there is much evidence that many factors (e.g., recency, parallelism, coherence relations) influence pronoun resolution preferences. Mitkov [30] provides a useful overview of these factors while Lappin and Leass [28] evaluate the relative influence of a variety of these factors in their Resolution of Anaphora Procedure. Gordon and Hendrick [12] do discuss Lappin and Leass' work, yet still seem to reject other factors, excluding them from their idea of discourse prominence.

As an adaptation of DPT, therefore, I will view the procedure for determining the discourse prominence of referents as dependent on a number of different linguistic factors in some sort of combinatorial fashion and refer to this procedure as the *multiple prominence factor method* or MPFM, for short. Exactly how the various factors in the MPFM combine I will leave to future work, but one possibility might include a simple arithmetic summation across indices determined from each prominence factor. This is the approach taken in Lappin and Leass' procedure. Another possibility might be to determine discourse prominence in a constraint-based approach with constraints derived from the various factors.

While I acknowledge that a variety of factors may play a significant role in this procedure, in this study, I look at only two factors: syntactic prominence and semantic prominence. In the next two sections I discuss these two factors in greater detail.

Syntactic Prominence. Many researchers have observed preferences for an unbound pronoun to be interpreted as coreferent with a referent previously introduced in subject position [20, 29] or in an utterance-initial position [7]. For example, the preferred interpretation of the (unaccented) pronoun in (4b) is to the referent introduced as the subject of the preceding utterance, Luke.

(4)   a. $Luke_i$ hit Max_j.
       b. $He_{i/\#j}$ ran home.

The typical account of these observations is to assume that the syntactic structure of an utterance imposes a prominence hierarchy on the referents introduced in that utterance. The exact way in which the syntactic information determines prominence varies from theory to theory—for example, grammatical role (e.g., subject, object, etc.) in the centering framework of Grosz and Sidner [13, 14], relative height in the syntactic tree in Hobbs' algorithm [18], or

linear order-of-mention [7]—but crucially it is the structural configuration of an utterance which determines the relative prominence of referents.

Semantic Prominence. One problem with the syntactic prominence account is that in English, at least, syntactic information and semantic information are often conflated. That is, for example, referents introduced as sentential subjects are often semantic agents and bear more proto-agent entailments (e.g., sentience, volition) [5] while those introduced as objects are often semantic patients and bear more proto-patient entailments (e.g., affectedness). Thus, an alternative account of the observation in (4) above is to assume that the semantic information imposes a hierarchy on discourse referents such that those introduced as agents are more prominent than those introduced as patients. As such, there is a preference to interpret the pronoun in (4b) as coreferent with the more semantically prominent referent, Luke.

Some researchers have looked at the influence of semantic information in referential processing in somewhat different ways. Prat-Sala and Branigan [31] observed that animate entities were preferred over inanimate entities as antecedents in pronoun resolution. In other work, Stevenson and colleagues [33, 34] and Arnold [1] suggest that in forward-looking discourse planning, referents introduced in certain roles (e.g., patients in agent-patient constructions, goals in source-goal constructions) are the default focal point for reference in an immediately following utterance. They suggest, however, that in pronoun resolution (a backward-looking process), only syntactic information is relevant—that the default referent of an utterance-initial pronoun is the subject of the preceding utterance.

For the present study, I will be taking a slightly different approach. I assume that the semantic prominence of discourse referents is determined by their semantic roles (e.g., agent, patient, etc.). Referents are ranked with respect to some hierarchy of semantic roles. Exactly what these roles are and how they are ranked I will leave unspecified. One possibility might include using role sets that have been proposed in the syntax-semantics literature (e.g., [6, 21, 32]). For the present study, I will assume the presence of both agent and patient semantic roles and that agent is ranked higher in the hierarchy than patient.

2.3   Summary

In this paper, I investigate semantic prominence by presenting data from a series of psycholinguistic experiments designed to evaluate and compare the effects of both syntactic and semantic prominence on pronoun resolution. Because, as noted above, syntactic and semantic information are often conflated, it is necessary to find a linguistic environment that allows the influence of each to be observed. I suggest that argument-reordering constructions are a good candidate for this. Consider the contrast between the non-*tough* and *tough* constructions in (5)-(6).

(5)   John$_i$ could easily hit Matt$_j$ in the boxing match.

(6) Matt$_j$ was easy for John$_i$ to hit $\emptyset_j$ in the boxing match.

If syntactic information is what determines discourse prominence, then the prediction would be that an immediately following pronoun (i.e., *he*) should preferentially be interpreted as coreferent with the subject: John in (5) and Matt in (6). However, if semantic information determines discourse prominence, then the preference should be for the agent in both cases: John. Thus, the experiments described in the next section make use of this contrast in a self-paced reading task to compare the influence of syntactic and semantic information on discourse prominence.

3 Experiment I

3.1 Design

The goal in the first experiment was to compare the influence of syntactic and semantic information in pronoun resolution preferences during on-line discourse processing. The experiment takes advantage of the non-*tough*/*tough* alternation discussed above and extended in (7).

(7) a. John$_i$ could easily hit Matt$_j$ in the boxing match. CONTROL
     a'. Matt$_j$ was easy for John$_i$ to hit $\emptyset_j$ in the boxing match. SPLIT
     b. He$_i$ even managed to land a knockout punch. AGENT
     b'. He$_j$ became bruised and bloodied all over. PATIENT

In the non-*tough* case, (7a), syntactic and semantic information converge to promote the same referent as more discourse prominent (i.e., John). I will refer to this case, therefore, as the CONTROL condition. In the non-*tough* case, on the other hand, syntactic and semantic information diverge and promote different referents. I will therefore refer to this case as the SPLIT condition. These two sentences, respectively for each condition, determine the context in which the continuation sentence (b/b') is processed. These continuation sentences begin with a pronoun and plausible under only one interpretation of the pronoun—coreferent with either the AGENT (John) or the PATIENT (Matt) of the preceding utterance. In terms of DPT, when the pronoun is encountered, it will be automatically assigned to the most prominent referent in the context in accordance with CR.PRO. However, if at some later point the reader realizes the assignment was incorrect, then CR.RA will be triggered costing time. Therefore, in a self-paced reading experimental task, shorter reading times will indicate which referent is perceived as more discourse-prominent. This approach was used in this experiment which was a 2×2 design pitting CONTEXT (CONTROL, SPLIT) against intended REFERENT of the pronoun (AGENT, PATIENT).

3.2 Method

Participants. The participants in this experiment were 32 undergraduate students from the Northwestern University Linguistics Department subject pool

who were native speakers of North American English. The participants received course credit in return for their participation.

Materials. A total of 32 stimulus items were prepared using six adjectives (*tough, hard, fun, easy, difficult, a cinch*) which exhibit the non-*tough*/*tough* alternation and 32 agent-patient verbs (e.g., *hit, catch, capture*). Continuation sentences were prepared so that the initial pronoun referred to either the AGENT or PATIENT. Each pair of continuation sentences was also balanced for ASCII character length. These two-sentence test sequences were then embedded in longer discourses to make a five-sentence vignette as shown in (8).

(8)   a. John and Matt took part in an important boxing match.
       b. It was twelve rounds long.
       c. John$_i$ could easily hit Matt$_j$ in the final round. CONTROL
       c'. Matt$_j$ was easy for John$_i$ to hit $\emptyset_j$ in the final round. SPLIT
       d. He$_i$ even managed to land a knockout punch. AGENT
       d'. He$_j$ became bruised and bloodied all over. PATIENT
       e. The judges had no trouble deciding the winner.

Each vignette was followed by one comprehension question. These questions focused on different parts of the vignette in order to encourage participants to read and process the entire discourse.

The 32 items were combined with 48 items from a different experiment. The items were randomized into blocks and presentation of the items in the different conditions was balanced across participants so that adjacent stimuli were not from the same experimental condition.

Procedures. The stimuli were presented one sentence at a time on a computer screen using Superlab by Cedrus Corporation. Participants were instructed to read each sentence and then press a button to continue to the next sentence. Participants were asked to read each vignette as quickly as possible, but also to concentrate on comprehension. The time between button presses was recorded as their reading time. In this study, only the reading times of the fourth sentences, (8d/d') are analyzed.

3.3   Results

The results of Experiment 1 are shown in Figure 1. There was a main effect of CONTEXT, no effect of REFERENT, but a marginally significant interaction between the two. In the CONTROL condition, participants read the AGENT continuation sentence faster suggesting they preferred the pronoun in the continuation sentence to be coreferent with John—the syntactically and semantically prominent entity in the context sentence. However, in the SPLIT condition, participants did not show any preference for either continuation sentence.
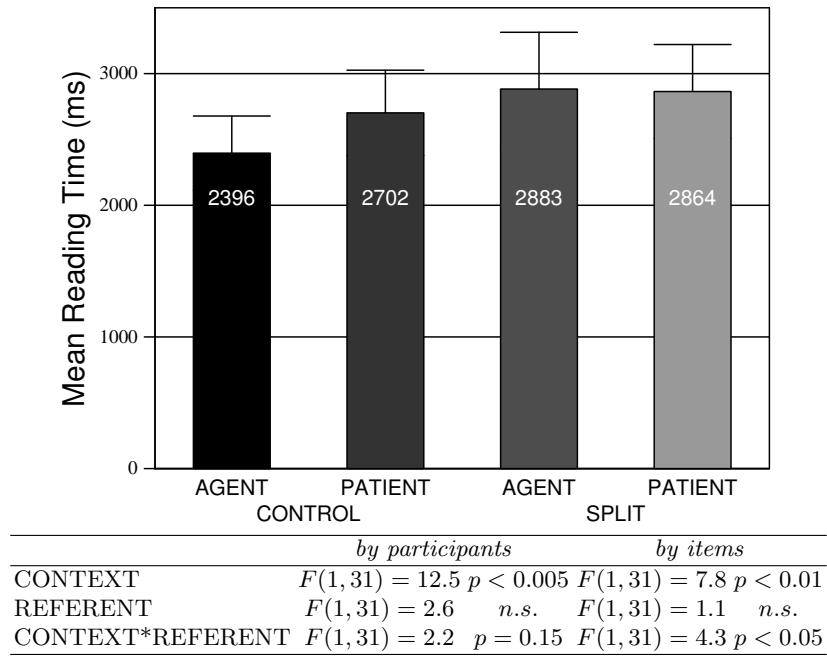
	by participants		by items	
CONTEXT	$F(1,31) = 12.5$	$p < 0.005$	$F(1,31) = 7.8$	$p < 0.01$
REFERENT	$F(1,31) = 2.6$	$n.s.$	$F(1,31) = 1.1$	$n.s.$
CONTEXT*REFERENT	$F(1,31) = 2.2$	$p = 0.15$	$F(1,31) = 4.3$	$p < 0.05$

Fig. 1. Mean reading times with 95% confidence intervals for the continuation sentences (8d/d') for participants ($n = 32$) in Experiment I. Two main factors were tested: CONTEXT (CONTROL, SPLIT) and intended REFERENT of pronoun (AGENT, PATIENT).

3.4   Discussion

The experimental results show that in the CONTROL condition, participants prefer to interpret the pronoun as coreferent with the syntactically and semantically prominent entity. This is consistent with previous experimental work described in Section 2.2 where preference is shown for a syntactically prominent entity. The current experiment thus replicates those results. However, the results in the SPLIT condition are quite interesting: Participants showed no preference for either referent.

One explanation for these results is that *both* syntactic and semantic prominence influence the ranking of candidate antecedents in an independent fashion. When syntactic and semantic prominence coincide to promote one antecedent (as in the CONTROL condition), then pronoun resolution processes can select one candidate over others. However, when syntactic and semantic prominence diverge, promoting different entities, then pronoun resolution processes do not show any preference. This could be accounted for in the MPFM in different ways: if the method uses a simple summation across prominence factors to calculate the discourse prominence of referents, then in the CONTROL condition, the syntactically and semantically prominent referent is doubly boosted and has a clearly higher total discourse prominence index than other referents. Then, in terms of DPT, the search for a suitable referent is concluded successfully and the pronoun is subsequently resolved to this referent, the AGENT. With the AGENT continuation, then, nothing more happens and the correct discourse representation is achieved. However with the PATIENT continuation, semantic information introduced later in the sentence results in an inconsistency which triggers CR.RA, leading to the increased reading times as observed.

In the SPLIT condition, however, the two different referents receive comparably-sized boosts from the different factors, respectively, such that their net discourse prominence values are essentially equal. In terms of the DPT, this would seem to be a case in which the search for a suitable referent might be unsuccessful because there is more than one such referent. DPT allows that when a search is unsuccessful, a new discourse referent is introduced. Later information, though, shows that the pronoun is coreferent with an existing referent, so CR.RA is triggered to establish equivalence between the new referent and the intended referent. In the SPLIT condition, this sequence of operations appears to have happened for both the AGENT and PATIENT continuations yielding comparable reading times in both.

Thus, the experimental results can be captured in DPT, but only with a richer conception of how suitable referents are determined—one that is based on multiple prominence factors.

4   Experiments IIa-b

One criticism that may be made of the first experiment is that reading time measurements are being compared across different sentences. While the length

of the continuation sentences was controlled, the lexical items and syntactic structure and complexity were not. This could be one source of variation. One way to overcome this problem is to take advantage of the *repeated-name penalty* experimental technique described in Gordon, et al. [10]. They observed that readers take longer to read sentences containing reference to a currently focused entity when the reference is by name (e.g., *John* as in (9b)) rather than by pronoun (e.g., *he* as in (9b')).

(9)   a. John walked to the supermarket.
      b. John bought two fish.
      b'. He bought two fish.

In DPT, this is explained in the same way as the c-commanding case discussed in Section 2.1. After the context sentence in (9a), John is the most discourse-prominent referent. Thus, when the pronoun in the continuation sentence in (9b') triggers CR.PRO, John will be judged the most suitable referent and coreference will be readily established. However, the proper name in (9b) will merely trigger CR.PN and then the introduction of a new discourse referent *different from* the existing referent of John in the context. Subsequent information indicating that these two referents point to the same entity in the real world will then trigger CR.EQ to establish equivalence between these referents. The additional operations necessary to establish coreference are presumed to lead to increased reading times and hence, the repeated-name penalty.

4.1   Design

In the present study, the repeated-name penalty experimental paradigm is a useful way to look more closely at how participants perceive the relative discourse prominence of referents in the context by comparing the repeated-name penalties across the various experimental conditions. Thus, the difference in reading times between the pronoun and name versions of (10b) can be compared to that of (10b') in both the CONTROL and SPLIT conditions. Based on the results of Experiment I, the prediction is that in the CONTROL condition, there should be a larger repeated-name penalty for the AGENT than for the PATIENT continuation sentence, but in the SPLIT condition, the repeated-name penalties should be approximately the same.

(10)   a. $John_i$ could easily hit $Matt_j$ in the boxing match. CONTROL
       a'. $Matt_j$ was easy for $John_i$ to hit $\emptyset_j$ in the boxing match. SPLIT
       b. [$John_i$ / He_i] even managed to land a knockout punch. AGENT
       b'. [$Matt_j$ / He_j] became bruised and bloodied all over. PATIENT

In order to test these predictions, two further experiments were thus performed, one looking at the CONTROL condition and the other looking at the SPLIT condition. Both experiments were a 2×2 design pitting intended REFERENT (AGENT, PATIENT) against FORM of reference (PRONOUN, NAME).

4.2   Method

Participants. 32 undergraduate students from the Northwestern University Linguistics Department subject pool who were native speakers of North American English participated in each of the two experiments reported here. None of these students had participated in Experiment I. The participants received course credit in return for their participation.

Materials. The materials for this experiment were the same as those used in Experiment I except that two versions of the continuation sentences (i.e., (8d/d'))—one beginning with a pronoun and one with a repeated name—were used. Experiment IIa used stimuli only in the CONTROL condition while Experiment IIb used stimuli only in the SPLIT condition.

Procedures. The procedures for these two experiments were exactly the same as those reported above for Experiment I.

4.3   Results

The results of Experiment IIa are shown in Figure 2. In this experiment—the CONTROL condition from Experiment 1—there was a marginal main effect of REFERENT, no effect of FORM, but a significant interaction between the factors. These results appear to be driven by an 83ms repeated-name penalty with the AGENT continuation and a 270ms repeated-name *advantage* (i.e., a negative penalty) with the PATIENT continuation as shown in Table 1. Post-hoc t-tests, however, do not show that either of these penalties is significantly different from a null hypothesis of 0ms.

Table 1. Experiment IIa: CONTROL Condition Repeated-Name Penalties

	penalty	*by participants*	*by items*
AGENT	83ms	$t(31) < 1.0$ *n.s.*	$t(31) < 1.0$ *n.s.*
PATIENT	-270ms	$t(31) = 2.5$ *n.s.*	$t(31) = 2.7$ $p = 0.07$

The main effect of REFERENT suggests that on the whole, participants prefer that the continuation contain reference (regardless of form: name or pronoun) to the most discourse-prominent entity. This is consistent with many theories of forward-looking discourse construction [33, 34, 13]. The significant interaction between REFERENT and FORM indicates that the AGENT continuation exhibited a significantly larger repeated-name penalty than the PATIENT continuation. What is interesting, though, is that—although these numbers are not
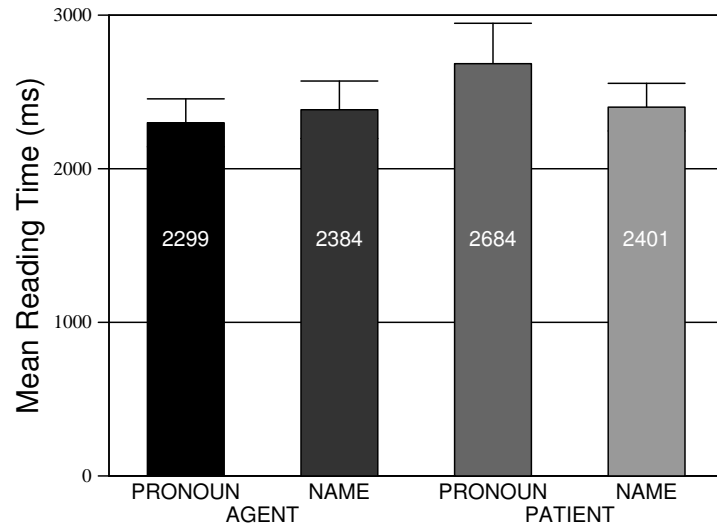
	by participants	by items
REFERENT	$F(1,31) = 7.9\ p < 0.01$	$F(1,31) = 3.6\ p = 0.067$
FORM	$F(1,31) = 1.6\quad n.s.$	$F(1,31) = 1.5\quad n.s.$
REFERENT*FORM	$F(1,31) = 6.3\ p < 0.05$	$F(1,31) = 8.5\ \ p < 0.01$

Fig. 2. Mean reading times with 95% confidence intervals for the continuation sentences (8d/d') for participants ($n = 32$) in Experiment IIa—the CONTROL condition from Experiment I. Two main factors were tested: intended REFERENT (AGENT, PATIENT) and referential FORM (PRONOUN, NAME).

statistically strong—it seems that the AGENT continuation incurs no repeated-name penalty, while the PATIENT continuation incurs a repeated-name advantage. Some implications of this will be discussed in the discussion section below.

The results of Experiment IIb using the SPLIT condition stimuli are shown in Figure 3. In contrast to Experiment IIa, there were no significant main effects and no significant interaction. There was a 168ms repeated-name advantage in the AGENT condition and a 46ms repeated-name advantage in the PATIENT condition as shown in Table 2. However, neither of these was significant.
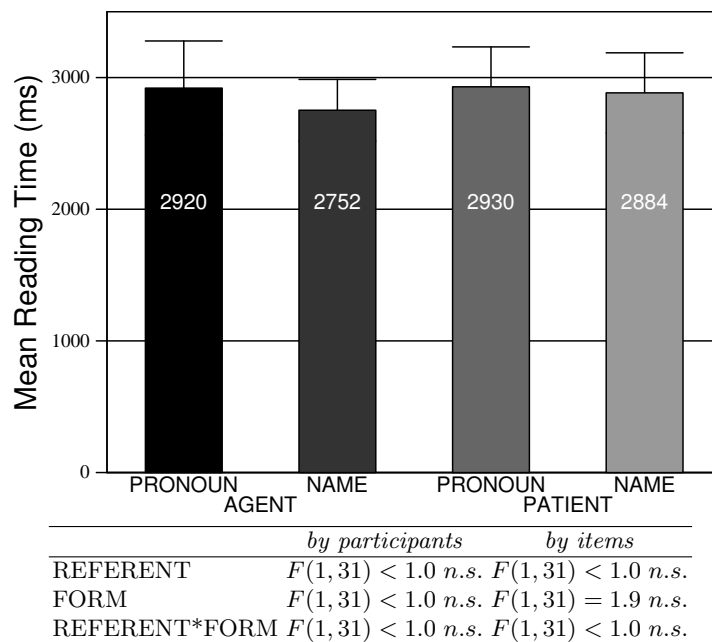

	by participants	by items
REFERENT	$F(1,31) < 1.0$ n.s.	$F(1,31) < 1.0$ n.s.
FORM	$F(1,31) < 1.0$ n.s.	$F(1,31) = 1.9$ n.s.
REFERENT*FORM	$F(1,31) < 1.0$ n.s.	$F(1,31) < 1.0$ n.s.

Fig. 3. Mean reading times with 95% confidence intervals for the continuation sentences (8d/d') for participants ($n = 32$) in Experiment IIb—the SPLIT condition from Experiment I. Two main factors were tested: intended REFERENT (AGENT, PATIENT) and referential FORM (PRONOUN, NAME).

In short, the results of Experiment IIb are basically flat-lined with participants showing no apparent preferences for any continuation across the board.

4.4   Discussion

Taken alone, the results of Experiment IIb are probably unremarkable, but taken together with the results of Experiment IIa, they reinforce the conclusion that both syntactic and semantic prominence influence the ranking of candidates for pronoun resolution: When syntactic and semantic prominence converge, then

Table 2. Experiment IIb: SPLIT Condition Repeated-Name Penalties

	penalty	*by participants*	*by items*
AGENT	-168ms	$t(31) < 1.0$ *n.s.*	$t(31) = 1.4$ *n.s.*
PATIENT	-46ms	$t(31) < 1.0$ *n.s.*	$t(31) < 1.0$ *n.s.*

pronoun resolution prefers the promoted candidate, but when syntactic and semantic diverge, then pronoun resolution shows no preference. This can be captured in the DPT as a part of the process of determining a suitable referent for a pronoun: This process takes advantage of a ranking method which depends on a number of different factors such as the MPFM described above.

A secondary implication of the results of Experiments IIa-b is that the repeated-name penalty must been seen in a new light. Ultimately, this comes down to how the search for a suitable referent proceeds. In the original experiments which established the repeated-name penalty concept [9, 10], most of the stimuli had contexts in which there was little or no chance of ambiguity because of parallelism effects, topicalization, or gender-disambiguation. If there is only one compatible referent, then the search for a suitable referent will be relatively straightforward and resolution should be quick. Similarly, if we assume that topicalizing a referent makes it very highly discourse prominent, then the search for a suitable referent may still be very easy because any competing candidates will be so low in the prominence hierarchy. Thus, in both of these cases it is not surprising that the contrasting case with repeated-name reference would take much longer because of the subsequent triggering of CR.EQ.

The present experiments indicate that the search for a suitable referent may actually be more costly when there is more than one compatible referent in the context. In the CONTROL condition of the experiment, the AGENT referent is more discourse prominent, but apparently not so much so that it is immediately deemed the most suitable referent (as it might if it were topicalized). Therefore, with the AGENT continuation, the processor must take roughly as much time in the PRONOUN condition as it must take in the NAME condition to establish equivalence among the discourse referents: In the latter condition, CR.PN and CR.EQ are triggered while in the former condition, only CR.PRO is triggered. Yet the same net time is taken in each case. This seems to be best explained by seeing the search for a suitable referent as being a more costly procedure when there are other compatible referents. With the PATIENT continuation, the same difficulty is faced by the processor except that in the PRONOUN condition, CR.RA is also triggered because the pronoun had been assigned by default to the more discourse-prominent AGENT. This leads to a large delay in this condition—thus an apparent advantage for the repeated name continuation.

In the SPLIT condition however, the search for a suitable referent is immediately concluded because no single suitable referent can be found—the two potential referents are equally ranked. Thus CR.PRO introduces a new discourse

referent and later, CR.RA is triggered to establish coreference. In the NAME conditions CR.PN introduces a new discourse referent and later, CR.EQ is triggered. Thus, in all the conditions, the same net costs are incurred: those caused by introducing a new referent and subsequently by establishing equivalence among referents.

In short, while DPT accounts nicely for the results of these experiments, it is necessary to bring in a more sophisticated conceptualization of how the process of finding a suitable referent proceeds. It is not—as originally suggested by Gordon and Hendrick [11, 12]—as simple as selecting the referent realized in the most syntactically prominent position. Rather, there is at least one other factor (perhaps many) that determine discourse prominence; namely, semantic prominence. Furthermore, the relative discourse prominence of referents influences the speed with which the search process may be concluded.

5   General Discussion

In short, the experimental evidence here, combined with evidence from the Stevenson, et al. [34] and Arnold [1] studies described briefly in Section 2.2, suggest that semantic information affects both forward-looking and backward-looking referential processes in discourse. The results of those studies, however, show some interesting contrasts with the present study. For instance, one contrast is that while the present experiment shows agent-preference for pronoun reference, [34] shows a default patient-preference for topic continuation. This is not necessarily a contradiction. While it would be theoretically convenient if the same ranking scheme affected both forward-looking and backward-looking referential processes, this does not have to be the case. Further work is clearly necessary to to understand just how these processes are related to one another.

The results of this study point toward two other areas for further study. First, while there is much work looking at ambiguous pronoun resolution, much of this work seems to be limited to cases where one candidate outranks other candidates. The present study suggests that there are cases where ranking produces ties. This is not a new notion, however. There are many models which suggest that discourse entities are only partially-ordered in their prominence (e.g., the list of forward-looking centers in classical centering theory; Grosz, et al. [13]). Yet how pronoun resolution processes actually deal with cases of equally-ranked candidates seems to be much less studied.

A second area for further research concerns the ranking scheme. The evidence here strongly suggests, as noted previously, that ranking is based on a number of factors as in the MPFM. This is not a new concept, of course, and many pronoun resolution algorithms have achieved a fairly high degree of success with such methods (e.g., [26, 28]). However, there is more work to be done on the way the ranking is actually utilized by the processor. The experimental evidence in this study suggests that the relative ranking of referents on the discourse prominence hierarchy affects how those referents are accessed during pronoun resolution processes. A referent which is ranked much higher than any other

referent seems to block, in a sense, consideration of those other referents. While on the other hand, as referents are more closely ranked in the hierarchy, more time is required to consider them. Yet when they become too closely ranked, then the search for a suitable referent fails.

6   Conclusion

In conclusion, the series of experiments presented here suggest that both syntactic and semantic prominence contribute to the ranking of candidates for pronoun resolution in a way that may result in a partially-ordered ranking. Furthermore, *tough*-constructions seem to be a useful construction for generating such partially ordered rankings and therefore may prove a useful means for studying how pronoun resolution processes deal with equally-ranked candidates. DPT provides a useful framework in which to capture the time-course of discourse comprehension and pronoun resolution, but only with a more complex conceptualization of how the discourse prominence of referents is determined and how the processor makes use of the ranked list of referents.

References

1. Arnold, J.: The effects of thematic roles on pronoun use and frequency of reference. Discourse Processes **31** (2001) 137-162
2. Arnold, J., Eisenband, J., Brown-Schmidt, S., Trueswell, J.: The rapid use of gender information: Evidence of the time course of pronoun resolution from eyetracking. Cognition **76** (2000) B13-B26
3. Arnold, J.: Marking salience: The similarity of topic and focus. Unpublished Manuscript (1998)
4. Boland, J., Acker, M., Wagner, L.: The use of gender features in the resolution of pronominal anaphora. Cognitive Science Technical Report 17, The Ohio State University Center for Cognitive Science (1998)
5. Dowty, D.: Thematic proto-roles and argument selection. Language **67** (1991) 547-619
6. Fillmore, C.: The Case for Case. In Bach, E., Harms, R., ed.: Universals in Linguistic Theory. Holt, Rhinehart, and Winston, New York (1968) 1-90
7. Gernsbacher, M.A., Hargreaves, D.: Accessing sentence participants: The advantage of first mention. Journal of Memory and Language **27** (1988) 699-717
8. Givón, T.: Topic continuity in discourse: An introduction. In Givón, T., ed.: Topic Continuity in Discourse: A Quantitative Cross-Language Study. John Benjamins, Amsterdam (1983) 1-42
9. Gordon, P., Chan, D.: Pronouns, Passives, and Discourse Coherence. Journal of Memory and Language **34** (1995) 216-231
10. Gordon, P., Grosz, B., Gilliom, L.: Pronouns, names, and the centering of attention in discourse. Cognitive Science **17** (1993) 311-347
11. Gordon, P., Hendrick, R.: Intuitive knowledge of linguistic co-reference. Cognition **62** (1997) 325-370
12. Gordon, P., Hendrick, R.: The representation and processing of coreference in discourse. Cognitive Science **22** (1998) 389-424

13. Grosz, B., Joshi, A., Weinstein, S.: Centering: A framework for modeling the local coherence of discourse. Computational Linguistics **21** (1995) 203-225
14. Grosz, B., Sidner, C.: Attention, intentions, and the structure of discourse. Computational Linguistics **12** (1986) 175-204
15. Gundel, J., Hedberg, N., Zacharski, R.: Cognitive status and the form of referring expressions. Language **69** (1993) 274-307
16. Heim, I.: The Semantics of Definite and Indefinite Noun Phrases. PhD thesis, University of Massachusetts, Amherst (1982)
17. Heim, I.: File change semantics and the familiarity theory of definiteness. In Bauerle, R., Schwarze, C., Von Stechow, A., eds.: Meaning, Use, and Interpretation. DeGruyter, Berlin (1983) 164-189
18. Hobbs, J.: Resolving pronoun references. Lingua **44** (1978) 311-338
19. Hobbs, J.: Coherence and coreference. Cognitive Science **3** (1979) 67-90
20. Hudson-D'Zmura, S., Tanenhaus, M.: Assigning antecedents to ambiguous pronouns: The role of the center of attention as the default assignment. In Walker, M., Joshi, A., Prince, E., eds.: Centering Theory in Discourse. Clarendon Press, Oxford (1997) 199-226
21. Jackendoff, R.: Semantic Interpretation in Generative Grammar. MIT Press, Cambridge, MA (1972)
22. Kamp, H.: A theory of truth and semantic representation. In Groenendijk, J., Janssen, T., Stokhof, M., eds.: Formal Methods in the Study of Language. Mathematisch Centrum, Amsterdam (1981) 277-322
23. Kamp, H., Reyle, U.: From Discourse to Logic. Kluwer Academic, Dordrecht (1993)
24. Karttunen, L.: Discourse referents. In McCawley, J., ed.: Syntax and Semantics, Vol. 7: Notes from the Linguistic Underground. Academic Press, New York (1976) 363-385
25. Kehler, A.: Coherence, Reference, and the Theory of Grammar. CSLI Publications, Stanford University, CA (2002)
26. Kennedy, C., Boguraev, B.: Anaphora for everyone: Pronominal anaphora resolution without a parser. Proceedings of the 16th International Conference on Computational Linguistics (COLING '96) (1996) 113-118
27. Kintsch, W., van Dijk, T.: Toward a model of text comprehension and production. Psychological Review **85** (1978) 363-394
28. Lappin, S., Leass, H.: An algorithm for pronominal anaphora resolution. Computational Linguistics **20** (1994) 535-561
29. Mathews, A., Chodorow, M.: Pronoun resolution in two-clause sentences: Effects of ambiguity, antecedent location, and depth of embedding. Journal of Memory and Language **27** (1988) 245-260
30. Mitkov, R.: Anaphora Resolution. Longman, London (2002)
31. Prat-Sala, M., Branigan, H.: Discourse constraints on syntactic processing in language production: A cross-linguistic study in English and Spanish. Journal of Memory and Language **42** (1999) 168-182
32. Speas, M.: Phrase Structure in Natural Language. Kluwer Academic Publishers, Dordrecht (1990)
33. Stevenson, R., Crawley, R., Kleinman, D.: Thematic roles, focus and the representation of events. Language and Cognitive Processes **9** (1994) 519-548
34. Stevenson, R., Knott, A., Oberlander, J., McDonald, S.: Interpreting pronouns and connectives: Interactions among focusing, thematic roles and coherence relations. Language and Cognitive Processes **15**(3) (2000) 225-262
35. Ward, G.: The Semantics and Pragmatics of Preposing. PhD thesis, University of Pennsylvania (1985)