Rethinking the SMASH Approach to Pronoun Interpretation

ANDREW KEHLER

1. Introduction

The last three decades of research in psycholinguistics and computational linguistics have produced an extensive body of work on pronoun interpretation. Despite the diversity of viewpoints, research methodologies, and ultimate conclusions provided by this literature, a majority of these studies has assumed (either explicitly or implicitly) that a particular type of process underlies pronoun interpretation. I will call this process (for lack of a better term) the SMASH paradigm—for Search, Match, and Select using Heuristics—which is characterized as follows:

1. Search: Collect possible referents within some suitable contextual window (usually the current utterance and 1–3 utterances prior).
2. Match: Filter out those referents that fail “hard” morphosyntactic constraints such as number, gender, and person agreement, and intrasentential syntactic binding constraints.
3. Select using Heuristics: Select a referent from those that remain by applying a set of heuristically based (“soft”) preferences. These are usually based on surface-level morphosyntactic factors, such as grammatical role ranking and grammatical role parallelism, among others.

The particular way this abstract procedure is instantiated varies in its details across different theories and algorithms, of course. The Centering
algorithm of Brennan et al. (1987), studied psycholinguistically by various authors (Hudson-D’Zouros 1989; Gordon, Gross, and Gilliom 1993; Brennan 1995; inter alia), performs Selection with respect to a ranking of centering transitions and a Centering rule (“Rule 1”) is added as a filter in the Match phase along with the more standard constraints mentioned earlier. The Select phase of Leppin and Less’s (1994) computational algorithm uses a combination of weighted preference factors to determine the salience of potential referents, from which the referent with the highest value is chosen. Hobbs’s (1978) well-known syntactic search mechanism, some predictions of which were psycholinguistically tested by Matthews and Chodorow (1988), uses an ordered Search phase that renders the Selection phase trivial: The referent chosen is the first one encountered during the search that satisfies the Match tests. Many other examples exist, not only of computer algorithms like these, but also with respect to the manner in which hypotheses are formulated and tested in the greater psycholinguistics literature.

In this chapter I argue that the SMASH way of framing the problem must be abandoned by any theory that seeks to explain pronoun interpretation within the human language processing mechanism. In so doing, I will offer a set of adequacy criteria that any analysis should at least be compatible with, if not ultimately explain. I will provide an outline of the form which I believe an adequate solution will ultimately take, although several questions will be left open for future work.

The remainder of the chapter is organized as follows. In the next section, I present a set of facts that are problematic for the Search and Match aspects of the SMASH paradigm. In section 3, I present facts that are problematic for the manner in which superficial morphosyntactically based ‘preferences’ are typically used in the Select phase. In light of these problematic facts, I will argue in section 4, that any adequate account of pronoun interpretation will have to be embedded in a larger model of discourse processing that accounts for the interaction between information structure and inferential processing mechanisms that underlie the establishment of coherence in discourse.

2. Problems for the Search and Match Approach

The third step of the SMASH approach uses heuristic preferences to choose a referent from the potential referents identified in Step 1 that survive a hard-constraint check in Step 2. If only one possible referent remains after the first two steps, that will be the referent trivially: in this case Step 3 has no work to do.

At first blush, this seems perfectly reasonable. In most cases an inability for a pronoun to refer to a locally introduced entity is due to the fact that a more salient compatible entity exists in the discourse context. For instance, Bush will (in most contexts) be preferred to Blair as the referent of He in passage (1).

(1) At the summit in Lisbon last Tuesday, George W. Bush met with Prime Minister Blair to discuss Middle East policy. He had to call the meeting short, however, due to a crisis that arose back home.

This accords with the predictions of most theories: even though Bush and Blair are both semantically plausible referents, the subject grammatical role is treated as a position that affords its occupant more salience than the object of a with-PP. This fact does not mean that a reference to Blair is inherently unpronominizable, however, as can be seen by considering passage (2).

(2) At the summit in Lisbon last Tuesday, National Security Advisor Condoleezza Rice met with Prime Minister Blair to discuss Middle East policy. He had to call the meeting short, however, due to a crisis that arose back home.

In this case, the person mentioned in subject position does not agree with the pronoun in gender. Thus, the reference to Blair is unambiguous—he is the only potential referent left when Step 3 is reached—and thus the reference will be correctly resolved by a SMASH procedure.

However, cases exist in which a pronoun cannot be felicitously used to refer to an entity even when it is the only one that satisfies the Match constraints. This unexpected fact casts serious doubt on the tenability of the Search and Match steps of the SMASH paradigm. The following subsections discuss three different types of example.

2.1 Not All Entities Are Salient Enough to Pronominalize

The first case I consider is example (3), which is a variant of an example from Gundel et al. (1993):

(3) Two Sears employees delivered some new appliances to my neighbors with the Doberman pinscher.

a. # It’s the same dog that bit Susan last summer.

b. That’s the same dog that bit Susan last summer.

The entity denoted by the Doberman pinscher does not license a subsequent pronoun reference; a demonstrative must be used per (3b). This failure to license a pronoun occurs despite the fact that the Doberman pinscher is (i) the most recently mentioned entity, (ii) occurs only one sentence back, and (iii) is in fact the only entity that satisfies the number restriction on the pronoun. As such, a SMASH procedure predicts such reference to be unambiguous and unproblematic.

Clearly this failure to pronominalize is not due to anything intrinsic to Doberman pinschers; placement in subject position allows for both pronouns
and demonstratives, as witnessed in (4), again adapted from an example used by Gundel et al.:

(4) My neighbor's Doberman pinscher bit a girl on a bike.
  a. It's the same dog that bit Susan last summer.
  b. That's the same dog that bit Susan last summer.

The grammatical position of the mention of the Doberman pinscher therefore appears to be of relevance: indeed, a natural argument to make here is that the syntactic position of the mention in (3)—a sentence-final noun phrase that is embedded in a prepositional phrase modifier to another noun phrase—is accorded too low a degree of salience to permit subsequent pronounization. If that were the case, then the Smash procedure could be consistent with these facts as long as it incorporated a lower bound on referent accessibility.

There is more to the story than that, however; compare passage (3) to passage (5):

(5) Lenox delivered new expensive china to my neighbors with the wild child. He'll have it all broken within a week!

The pronoun He in this case is felicitous, despite the fact that its referent is mentioned from the same grammatical position as the Doberman pinscher in (3). So, as Gundel et al. point out using similar examples (which I have modified so as to make each pronoun morphosyntactically incompatible with all but one entity-level referent), factors other than syntax must come into play.

It is tempting to speculate that the crucial difference between these examples lies in the speaker's purpose in uttering the prepositional phrases in each case. In (3), the most readily inferred speaker purpose behind the utterance of with the Doberman pinscher is that it restricts the reference of the NP to a unique set of neighbors. Once these neighbors are identified, the Doberman pinscher appears to have no further contribution to the overall proposition. While the prepositional phrase with the wild child may serve the same purpose in the first sentence of (5), the image that this sentence conjures up suggests that this particular choice for restricting the reference to the neighbors is not random—it is instead interpretable as a conversationally relevant description in the sense of Kronfeld (1990). That is, this choice may have been intended to create an expectation that the ensuing discourse will address the inadvisability of a couple with a wild child buying expensive china, which gives the wild child a role in the discourse that goes beyond restricting the reference to a unique set of neighbors.

If this line of reasoning is correct, it suggests that a discourse process as high-level as reasoning about the intentions that underlie a speaker's choice of linguistic expression is necessarily intertwined with the seemingly lower level process of pronoun interpretation.

2.2 Apparent Interference between 'Soft' Preferences and 'Hard' Constraints

In the last section, we saw an example of a reference that succeeded when the referent was mentioned in subject position (ex. 3), but did not when it was mentioned in an embedded, and therefore less salient, position (ex. 4). The success of the reference in (3) is unsurprising, since subject position is generally assumed to be associated with the highest degree of salience among grammatical roles.

However, there are examples in which even a subject referent cannot be felicitously pronounized in the subsequent utterance. (To my knowledge, examples of this sort were first noticed by Oehrle (1987).) Consider example (6):

(6) ?? Margaret Thatcher admires Ronald Reagan and George W. Bush absolutely worships her.

All of my informants have agreed that this example is infelicitous, assuming the pronoun is deaccented. These informants generally report a feeling that the pronoun should be assigned to Reagan, as if the speaker is confused about his gender. This is particularly striking for reasons similar to those cited with respect to example (3); not only is Thatcher evoked from the subject position of the previous sentence but (i) this pronoun assignment results in a completely plausible interpretation, and (ii) Thatcher is the only potential referent that meets the gender restriction of the pronoun. Again, a Smash algorithm would happily identify Thatcher as the referent regardless of the preferences it employs, since she is the only potential referent that survives the morphosyntactic constraint check in the Match phase.

As one might expect, the same effect occurs for mismatches in number, as in (7).

(7) ?? Republicans admire Ronald Reagan and George W. Bush absolutely worships them.

Again, the pronominal reference is infelicitous even though there is only one suitable referent available.

It may appear on the surface that these are examples in which a 'soft' preference for grammatical role parallelism actually trumps 'hard' constraints like gender and number agreement. This would in turn suggest that the Match step of the Smash procedure cannot come strictly before the Select step; they would instead need to be integrated in some manner. However, I will argue against this conclusion in section 4.5, where I claim that there are information structural factors governing accent placement at work in such examples that are completely orthogonal to whether a referring expression is realized as a pronoun. I will then argue that this fact denies the existence of any grammatical role parallelism preference at work in pronoun interpretation.
2.3 Conjunction

The final set of examples I consider in this section involves reference to an entity mentioned within a conjoined NP. First consider example (8):

(8) ?? Bush and Blair gave a press conference, and a reporter asked him a rude question.

The pronominal reference is infelicitous for reasons that should be apparent. Bush and Blair are evoked from equivalent grammatical positions, and are thus presumably accorded the same degree of salience in the discourse state. As such, they are indistinguishable as possible referents of the pronoun. Most SMASH theories would predict this fact, unless (i) they incorporate a preference for first mention, which if construed linearly (as opposed to hierarchically) would favor Bush, or (ii) recency, which if construed linearly would favor Blair.

A related fact that is less expected on the SMASH model is that a change in gender in one of the conjuncts does little to improve the felicity of the pronominal reference. Consider (9):

(9) ?? Rice and Blair gave a press conference, and a reporter asked him a rude question.

At a minimum, accent is required on him to make the reference felicitous. Repeating Blair’s name is more felicitous yet:

(10) Rice and Blair gave a press conference, and a reporter asked Blair a rude question.

Yet again we have a case in which one would expect reference with an unaccented pronoun to be unambiguous since there is only one male mentioned in the discourse, and as such a SMASH procedure would happily identify Blair as the referent. Note that the situation regarding (8) and (9) is in marked contrast to the difference between passages (1) and (2). Changing the subject referent from Bush in (1) to Rice in (2) made pronominal reference to Blair perfectly felicitous, whereas the same change from (8) to (9) did not.

Further complicating the situation are examples such as (11) and (12): 

(11) Bush gave a circumspect answer to a question on Iraq from Wolf Blitzer. He then brought up the hunt for Bin Laden.

(12) Bush and Blair each gave a circumspect answer to a question on Iraq from Wolf Blitzer. He then brought up the hunt for Bin Laden.

Most of my informants identify Bush as the referent of he in (11) and Blitzer as the referent of he in (12). We have already seen that the result for (12) is expected, since entities evoked from subject position are typically accorded greater salience than those evoked from within a PP modifier to another NP. The results for (12) are more interesting. Since all three possible referents are mentioned from embedded positions, it would be reasonable to expect that Bush and Blair would be more salient than Blitzer, since the former are embedded within the subject position. The “tie” in salience between Bush and Blair does not appear to cause infelicity as it did in example (8), but instead allows the pronoun to be assigned to a lower ranked entity, Blitzer. As before, the situation does not change appreciably if we change the gender of one of the conjuncts in subject position:

(13) Rice and Blair both gave a circumspect answer to a question on Iraq from Wolf Blitzer. He then brought up the hunt for Bin Laden.

These data might be taken as evidence that there is more to determining the referent of a pronoun than degree of salience, that is, that there is a topic-goal requirement in play (Gundel, Hedberg, and Zacharski 1993, inter alia). In each of the cases, the entities introduced in the conjoined subject NP presumably must either serve as topic together or not at all—it cannot be that one entity is the topic and the other is not. However, while topics are preferentially realized in subject position (in English), they need not be. Thus, a hearer trying to infer a potential topic from the first sentence of (13) as the referent of he will identify Blitzer, even though a more salient referent that agrees with the pronoun is (arguably) available. Analogous reasoning applies to the inability to identify a suitable topic as the referent of him in (9).

To summarize this section, the SMASH procedure predicts that if only one referent in the search window survives the application of ‘hard’ morphosyntactic constraints, then that entity is the referent trivially. We have seen here, however, three scenarios in which that is not the case. I am sure that others exist.

It should be acknowledged that one might object to the line of argumentation I have taken, citing the fact that since pronoun interpreters (human or machine) will generally not be confronted with passages like these (under the assumption that actual discourses are for the most part felicitous), the SMASH algorithm is doing the most reasonable thing by still selecting a referent. There are at least two reasons that I would not be swayed by such an objection. First, because these data reveal that the SMASH paradigm is a flawed way to conceive of pronoun interpretation, I would expect an analysis that is immune to these problems to have advantages in accounting for other data as well. Second, the problem with SMASH becomes more tangible when we look at it from a generation perspective. If left unenriched with additional information structural constraints, such algorithms will predict that pronominalization can occur during discourse production in all of the preceding problematic examples (again, on the basis of there being only one potential referent in the search window that satisfies Match constraints), thereby resulting in the generation of infelicitous discourse.
3. Problems for Selection Using Superficial Preferences

In the last section, I argued against the search-and-match aspects of the SMASH procedure—Steps 1 and 2—by appealing to three different types of example. In each case, infelicitous pronoun reference is predicted to be unambiguously and unproblematically by the SMASH approach.

In this section I consider Step 3 of the SMASH procedure, in particular its use of superficial morphosyntactic preferences to select pronoun referents. A variety of such preferences have been posited in the literature, based on orderings on grammatical roles, grammatical role parallelism, orderings on thematic roles, verb semantics, and the referential form of the antecedent, for example. Here I will focus on two of these: the preference for entities evoked from the subject position of the previous clause (henceforth the SUBJECT ASSIGNMENT preference), and the preference for entities evoked from the parallel grammatical role of the previous clause (henceforth the GRAMMATICAL ROLE PARALLELISM preference). I will argue that these preferences should not be blindly applied to all examples as SMASH procedures invariably do. Instead, we will see that each preference is in force only in a particular class of cases. I will argue that these preferences are in fact epiphenomena of deeper interpretation processes that only operate in particular contextual circumstances.

3.1 The Status of the Subject Preference

Many approaches to pronoun interpretation encode a preference for a pronoun to refer to the subject of the previous sentence. This preference is usually based on one of two related (but not equivalent) claims. The first is that subject position accords greater degree of salience to its occupant than do other grammatical roles; pronouns are thus presumed to be sensitive to the degree of salience of their possible referents. The second is that the subject position is the canonical place from which to mention a discourse topic (Chafer 1976; Gundel, Hedberg, and Zacharski 1993; Lambrecht 1994: inter alia); pronouns in this view are considered to be indicators of topic continuation.

We have already seen examples that support this preference: example (1) is repeated as (14):

(14) At the summit in Lisbon last Tuesday, George W. Bush met with Prime Minister Blair to discuss Middle East policy. He had to call the meeting short, however, due to a crisis that arose back home.

There are two possible referents of He. Bush and Blair, which are equally plausible semantically. The fact that Bush was placed in subject position appears to create the preference for him to be the referent. As such, if we were to switch the mentions of Blair and Bush, Blair becomes the preferred referent:

(15) At the summit in Lisbon last Tuesday, Prime Minister Blair met with George W. Bush to discuss Middle East policy. He had to call the meeting short, however, due to a crisis that arose back home.

Likewise, mentioning both Bush and Blair in a conjoined subject renders the reference infelicitous, as we saw for similar examples in section 2.3:

(16) At the summit in Lisbon last Tuesday, George W. Bush and Prime Minister Blair met to discuss Middle East policy. # He had to call the meeting short, however, due to a crisis that arose back home.

Since the first sentence in each of these three passages describes the same situation, the choice of syntactic form—which in turn determines which referents are mentioned from which grammatical roles—appears to be the crucial factor in determining the reference assignments.

There are a variety of situations in which the subject assignment preference is mysteriously neutralized, however. One such class includes constructions that indicate a transfer of possession or change of mental state. Stevenson et al. (1994), for instance, argue that occupants of different thematic roles in such constructions have different degrees of centrality within a hearer's mental representation of the end state of an event, based on examples such as (17a–b) and (18a–b).

(17) a. John seized the comic from Bill. He ... [began reading]
    b. John passed the comic to Bill. He ... [began reading]

(18) a. Ken impressed Geoff. He ... [knows a lot about cars]
    b. Ken admired Geoff. He ... [knows a lot about cars]

In a set of sentence completion experiments, Stevenson et al. found that hearers are more likely to continue passage (17a) in way that requires he to refer to John (as suggested by the bracketed text that I have added), whereas their completions of (17b) more often required he to refer to Bill. This result for (17b) belies the fact Bill is embedded within a sentence-final prepositional phrase, a position normally accorded less salience than the subject position. The property that these examples share is that the preferred referent occupies the GOAT thematic role of its respective predication, whereas the dispreferred entity occupies the SOURCE role. The analogous pattern was found for passages like (18a) and (18b), in which the referent assignments favored the occupant of the STIMULUS role (Ken and Geoff, respectively) over the occupant of the EXPERIENCER role (Geoff and Ken, respectively).

Another set of counterexamples manifest a causal relation between the clauses, as in (19) and (20).

(19) Bush blamed CIA director Tenet for the mistake. He had not properly vetted the speech.

(20) Colin pushed Don. He tumbled to the ground.
In each case the preferred referent is the occupant of the object position, even though the subject position entity is morphosyntactically compatible with the pronoun. In these cases, it would appear that semantic and world knowledge considerations are the determining factors.

The final class of counterexamples that I will discuss includes cases in which a pronoun favors the occupant of a parallel grammatical role, which is the topic of the next section.

3.3 The Status of the Grammatical Role Parallelism Preference

A variety of authors have argued for a basic preference in pronoun interpretation based on grammatical role parallelism (Kameyama 1986; Smyth 1994; Chambers and Smyth 1998; inter alia). This heuristic states that a pronoun will preferentially be associated with an antecedent in a parallel grammatical role. Various examples appear to support the existence of such a strategy.

First, the grammatical role parallelism preference is supported by examples in which a subject pronoun corefers with the subject of the previous sentence, as in examples (14) and (15). However, because this preference makes the same prediction as the subject assignment preference, examples with pronouns in other grammatical positions need to be considered to differentiate between the two. A canonical example of this sort is given in (21):

(21) Margaret Thatcher admires Hillary Clinton, and George W. Bush absolutely worships her.

There is an extremely strong preference to assign the (unaccented) pronoun her to Clinton. This example thus counterexemplifies the subject assignment preference, since Thatcher—which is not only consistent with the pronoun but preferred with respect to semantic plausibility—occupies that position. Intuitively, there appears to be a parallelism effect at play here that supersedes the subject preference.

We have already seen a number of examples, however, that call this preference into question. These include Stevenson et al.'s examples (17b) and (18a), and similarly the causally related examples in (19g) and (20). In which a subject pronoun is preferentially assigned a nonsubject referent. Likewise, examples in which a nonsubject pronoun prefers a subject referent are not hard to find. For instance, Kameyama (1996) reports that a majority of native informants prefer to identify the object pronoun in (22a) and (22b) with the subject of the preceding clause:

(22) a. John kicked Bill. Mary told him to go home. [= John ]

b. Bill was kicked by John. Mary told him to go home. [= Bill ]

(23) a. John kicked Bill. Mary punched him. [= Bill ]

In contrast, her study revealed a preference for grammatical role parallelism in (22c), creating an inconsistency among these data.

Thus, both the subject assignment and the grammatical role parallelism preferences are associated with examples that appear to definitively support their respective existences while counterexemplifying the other. In the next section I discuss how this confusing state of affairs can be resolved.

3.3 Do Preferences Compete?

So how are we to reconcile this contradictory evidence? Some authors have proposed that some type of competition among these preferences is at play. For instance, in ultimate arguments for the primacy of a grammatical role parallelism preference, Smyth (1994) acknowledges a role for the subject assignment preference:

I conclude from these observations that pronoun interpretation in conjoined sentences involves an obligatory search for a morphologically compatible antecedent which meets the binding theory (Chomsky, 1981) criteria for coreference and which, in addition, has the same grammatical role as the pronoun. If a match is found, the the parallel interpretation is obligatory, unless the pronoun is stressed, in which case it is selectively blocked. If no match is found, resolution is less certain, but will most often result in SA [= subject assignment], although if the pronoun or the first clause verb is stressed, alternative strategies govern the selection of an antecedent.

On this view, SA is a default strategy for sentences in which the degree of nonparallelism exceeds some limit: FF [= parallel function] is a specific outcome of the more general principle that the probability of parallel resolution depends on the number of features shared by the pronoun and the candidate antecedents. Retaining SA in the model allows us to account for an otherwise mysterious asymmetry between subject and nonsubject pronouns.

Similarly, Stevenson et al. (1995) carried out a set of question-answering experiments to determine whether the subject assignment and parallel grammatical role preferences jointly contribute to the interpretation of pronouns. They found that subjects more often resolved the pronoun to the subject position entity when both strategies indicated that preference, as compared to when only the subject assignment strategy applied. Furthermore, they found that nonsubject assignments were actually preferred when the two strategies disagreed. They conclude from this that the subject assignment and parallel function heuristics operate jointly, and that this in turn “implies a model of discourse processing in which a number of constraints compete in the interpretation of noun phrases.”

As I argue in Kehler (2002), however, these studies miss a crucial factor that distinguish the data, in particular the coherence relation that is manifest between the clauses. (My discussion here will out of necessity have to be quite brief; see Kehler (2002) for further details.) In that work, I categorize
coherence relations into three classes: Resemblance relations, Contiguity relations, and Cause-Effect relations. Examples of relations in these three categories, respectively, follow: the relation definitions are taken or adapted from those of Hobbs (1990):

Parallel: Infer $p(a_1, a_2, \ldots )$ from the assertion of $S_1$ and $p(b_1, b_2, \ldots )$ from the assertion of $S_2$, where $a$ and $b$ are similar for all $i$.

Occasion: Infer a change of state for a system of entities from each of $S_1$ and $S_2$, where the final state in $S_1$ provides the initial state for $S_2$.

Explanation: Infer $P$ from the assertion of $S_1$ and $Q$ from the assertion of $S_2$, where normally $Q \rightarrow P$.

Examples in which these different relations are operative show a characteristic pattern with respect to what surface-level interpretation preferences they provide support for. For instance, cases that support the subject assignment preference, for example (22a–b), are instances of the Occasion relation. (Certain cases that are problematic for this preference, such as example (17b), are instances of Occasion also. I return to this point in section 4.3.) Occasion is the relation typically operative in a narrated sequence of events: it allows one to express a situation centered around a system of entities by using intermediate states of affairs as points of connection between partial descriptions of that situation. The hearer's job, therefore, is to infer any information necessary to identify the initial state of the eventuality described by an utterance with the final state of the one that came before it. On the other hand, the examples that support a grammatical role parallelism preference, such as (21), are typically instances of the Parallel relation. This relation and others in the Resemblance category (Contrast, Exemplification, etc.) have quite a different character than Occasion: coherence with these relations requires that commonalities and contrasts among corresponding sets of parallel relations and entities be recognized using inference processes based on comparison, analogy, and generalization. Finally, the examples that were problematic for both approaches in (19) and (20) are instances of Cause-Effect relations, specifically Explanation (defined earlier) and Result (which is essentially the same relation except with the clause order reversed), respectively. These relations are established in yet a third way, as they require the identification of a causal chain that connects the propositions denoted by the uterances.

This pattern suggests that the characteristics of the different inference processes that underlie the establishment of these different types of coherence relation are likely to be in part responsible for the otherwise puzzling distribution of the data. Perhaps the most definitive evidence for this claim can be seen from examples that have more than one possible coherence interpretation. Consider example (23):


This passage has both a Parallel and a Result reading, and crucially, the assignment for the unaccented pronoun him is parasitic on this choice of interpretation. Under the Parallel interpretation, in which coherence is licensed by the similarity between deifying and punishing, the pronoun must be interpreted as referring to Cheney. This result accords with the grammatical role parallelism preference. On the other hand, under the Result reading of (23) in which coherence is licensed by the causal knowledge that a person who defies someone might get punished for it, him will be interpreted to refer to Powell. This result accords with the subject assignment preference. It is therefore hard to see how a preference-based system could predict this ambiguity (at least without predicting that all cases in which these two preferences compete are equally ambiguous, which we have already seen is not the case), since obviously the morphosyntactic properties of the passage remain constant for the two interpretations.

In the next section, I discuss analyses that integrate coherence and pronoun interpretation with respect to these and other data. I then revisit the status of the subject assignment and grammatical role parallelism preferences, and conclude that (i) the status of the subject preference is considerably more complicated than it might first appear, and (ii) that there is in fact no basic preference for parallelism at the level of grammatical roles in pronoun interpretation.

4. The View from Above

In the previous two sections, I argued on empirical grounds that the SMASH model cannot be correct. The data pose problems for both the search-and-match component of the model (section 2), and the reliance on superficial preferences (and combinations thereof) to rank candidate referents (section 3).

Beyond these empirical deficiencies, the SMASH approach also presents a larger conceptual problem. The amount of processing that has to occur at the time the pronoun is encountered in many SMASH algorithms (e.g., consider Smyth's (1994) and Stevenson et al.'s (1995) preference combining systems from the last section, or the Centering algorithm described in section 4.2, for instance) would seem to be at odds with a very basic fact: that the appropriate use of a pronoun generally has the effect of facilitating discourse interpretation, and not hindering it. After all, in choosing to use a pronoun, a speaker elects to use a potentially ambiguous expression that, under the SMASH paradigm, may require a computationally intensive effort for the hearer to resolve, rather than a less ambiguous or even unambiguous one that would presumably not (such as a proper name). Yet, as shown in the experiments of Gordon et al. (1993), for instance, discourses may be read more slowly if a proper name is used to refer to a focused entity instead of a pronoun. This suggests a rather obvious, but often ignored, facilitation paradox for a theory of pronoun interpretation. That is, if pronoun interpretation was really as hard as some SMASH approaches suggest, why would a speaker do the disservice to her hearer of using one?
The holy grail of a theory of pronoun interpretation, then, is to provide an answer to the facilitation paradox and at the same time explain the complex patterns of behavior one finds in the data. Picking up where we left off in section 3.3, I will now explore analyses that address the interaction of pronoun interpretation and coherence establishment with an eye toward these desiderata. In section 4.1 I discuss Hobbs's well-known treatment, in which pronoun interpretation is viewed purely as a side-effect of coherence establishment. While this account offers a potential explanation for the facilitation paradox, several facts regarding pronominal behavior are left unexplained. I then consider Centering theory in section 4.2. Although I believe Centering is correct in modeling the relationship between attentional state and pronoun interpretation explicitly (unlike Hobbs's account), it does not capture the type of interaction effects between information structure and coherence that the data suggest need to be modeled. In section 4.3 I begin to develop a theory that postulates such an integration. My claim is that by properly accounting for the division of labor between information structure, coherence establishment, and pronoun interpretation, we will find that much of the complexity that some Smash procedures consider to be part of pronoun interpretation should actually be accounted for elsewhere in the discourse processing mechanism, will ultimately argue that not only does a consideration of coherence lead to a (perhaps partial) resolution of the facilitation paradox, but it also explains a number of otherwise contradictory facts that have to this point resisted a satisfactory analysis.

4.1 Coherence and Inference

To my knowledge, Hobbs (1979) was the first to develop a theory of pronoun interpretation specifically based on the establishment of coherence relations. In fact, in his analysis pronoun interpretation is not an independent process at all, but instead results as a by-product of more general reasoning about the most likely interpretation of an utterance. Pronouns are modeled as free variables in logical representations that become bound during these inference processes; potential referents of pronouns are therefore those which result in valid proofs of coherence.

Let us illustrate with passages (24a) and (24b), adapted from an example from Winograd (1972).

(24) The city council denied the demonstrators a permit because ...
      a. … they feared violence.
      b. … they advocated violence.

In Hobbs's account, the correct assignment for the pronoun in each case falls out as a side-effect of the process of establishing the Explanation relation (here signaled by because), the definition of which is repeated as follows:

Explaination: Infer P from the assertion of S, and Q from the assertion of S', where formally Q → P.

Oversimplifying considerably, I will code the world knowledge necessary to establish Explanation for (24) within a single axiom, given in (25). (See Hobbs et al. (1993, p. 111) for a more detailed analysis of a similar example.)

(25) 
      fear(X, Y, V) ∧ advocate(Y, V) ∧ enable_to_cause(Z, Y, V) ⇒ deny(X, Y, Z)

This axiom says that if some X fears some Y, some Y advocates that same Y, and some Z would enable Y to bring about V, then X may deny Y of Z. To make this more concrete, the instantiation of this rule that is relevant for example (24) would say that if the city council fears violence, the demonstrators advocate violence, and a permit would enable the demonstrators to bring about violence, then this might cause the city council to deny the demonstrators a permit.

The first sentence in (24) can be represented with the predication given in (26).

(26) deny(city_council, demonstrators, permit)

This representation matches the consequent of axiom (25), triggering a process of abductive inference that can be used to establish Explanation. At this point, X will become bound to city_council. Y to demonstrators, and Z to permit.

Each of the follow-ons (24a–b) provides information that can be used to help "prove" the predications in the antecedent of the axiom, thereby establishing a connection between the clauses. Clause (24a) can be represented with predicate (27), in which the unbound variable T represents the pronoun they.

(27) fear(T, violence)

When this predicate is used to match the antecedent of axiom (25), the variables T and X are necessarily unified. Since X is already bound to city_council, the variable T representing they also receives this binding, and the pronoun is therefore resolved.

Likewise, clause (24b) can be represented as predicate (28).

(28) advocate(T, violence)

This predicate also matches a predicate within the antecedent of axiom (25), but in this case, the variables T and Y are unified. Since Y is already bound to demonstrators, the representation of they also receives this binding.

Thus, the correct referent for the pronoun is identified as a by-product of establishing Explanation in each case. The crucial information determining the choice of referent is semantic in nature, based on the establishment of the relationship between the predication containing the pronoun and the predication containing the potential referents. The fact that coreference came "for free" captures the effortlessness with which people appear...
to be able to interpret pronouns, offering a potential explanation for how the choice to use of pronoun can actually facilitate, rather than hinder, the process of discourse comprehension.

As perhaps the most elegant proposal out there for pronoun interpretation, it would certainly be nice if this was all there was to it. There are a variety of facts that are left unaccounted for, however. First, and most obviously, are sentence pairs in which pronominal assignments switch based on non-semantic factors, such as in passages (22a–b), repeated here as (29a–b):

(29)  
a. John kicked Bill. Mary told him to go home. [= John]  
b. Bill was kicked by John. Mary told him to go home. [= Bill]

Since Hobbs’s inference processes are carried out on representations that are purely semantic, syntactic distinctions such as voice are lost and thus are of no use in predicting this variation. Indeed, examples (24a–b) and (29a–b) create an interesting contrast: Whereas (29a–b) suggest that grammatical role is the primary determinant of pronominal interpretation (these examples keep semantics roughly constant), (24a–b) seem to suggest that semantics is the primary determinant, since these examples keep the relevant syntactic relationships constant, differing only with respect to the verb that occurs after the pronoun.

In fact, however, even variants of (24a–b) can be used as evidence that there is more to pronoun interpretation than coherence-driven reasoning. Consider the same passage except where the initial clause is passivized, as in (30):

(30)  
The demonstrators were denied a permit by the city council because...

a. ... they feared violence.

b. ... they advocated violence.

A strong majority of my informants assign the pronoun they to the subject (the demonstrators) in both follow-ons, even though the semantics of sentence (30a) might be expected to cause the pronoun to be identified with the city council in the same way that was just depicted for (24a). (Caramazza and Gupta 1979). In fact, performed experiments that showed this distinction using other stimuli. The alternation displayed by (24a–b) is simply absent in (30a–b). This is a problem for a purely coherence-based approach, since the logical relationships expressed by the first sentences of (24) and (30) are the same.

I take this as evidence that information structure is still important for pronoun interpretation, and that any adequate model will have to integrate it with coherence in a suitable way. Surely coherence-external factors were never fully disposable to begin with, as there are a variety of phenomena involving pronominal reference for which a coherence-based theory offers little help, such as intracausal anaphora (John likes his mother) and exophora (Look at him!).

Likewise, a model of information structure and its effect on attention will surely prove to be crucial for explaining certain garden-path effects that suggest that pronouns are interpreted incrementally. This phenomenon has in part motivated a framework that has been argued to integrate such factors with discourse coherence, namely Centering theory. In the next section I examine Centering and show that it does not provide answers to the kind of problems identified here.

4.2 Centering as a Theory of Pronoun Interpretation

The Centering theory of Gross et al. (1995, henceforth GJW) is largely motivated by two related facts about language that are not explained by purely content-based models of reference and coherence, such as that of Hobbs (1979). The first of these is that the coherence of a discourse does not depend only on semantic content but also on the type of referring expressions used. GJW illustrate this point with passage (31), which is meant to be interpreted as part of a longer segment that is currently centered on John.

(31)  
a. He has been acting quite odd. [He = John]  
b. He called up Mike yesterday.

c. John wanted to meet him quite urgently.

The third sentence in this passage is odd, despite the fact that the pronoun him in (31c) unambiguously refers to Mike. The oddness of this sentence stems from the choice of referring expressions used. In particular, the fact that the entity that is more central to the discourse (John) is not referred to with a pronoun whereas the less central element (Mike) is. As such, the same passage is perfectly felicitous if the final reference to John is pronounialized instead.

The second motivating fact is the existence of garden-path effects in pronoun interpretation, in which a pronoun appears to be interpreted before adequate semantic information has become available. GJW discuss passage (32):

(32)  
a. Terry really goofs sometimes.

b. Yesterday was a beautiful day and he was excited about trying out his new sailboat.

c. He wanted Tony to join him on a sailing expedition.

d. He called him at 6 a.m.

e. He was sick and furious at being woken up so early.

The passage is perfectly acceptable until sentence (32e), which causes the hearer to be misled. Whereas semantic plausibility considerations indicate that the intended referent for He is Tony, hearers tend to initially assign Terry as its referent, creating a garden-path effect. Such examples provide evidence
that more is involved in pronoun interpretation than simply reasoning about semantic plausibility. In fact, they suggest that hearers assign referents to pronouns at least in part based on other factors, before interpreting the remainder of the sentence.

In what follows, I will focus on Brennan et al.'s (1987, henceforth BFP) centering-based algorithm for pronoun interpretation. Each utterance in a discourse has exactly one backward-looking center (denoted C_i) and a partially ordered set of forward-looking centers (C_1, ..., C_n). The notation C_i(U_n) is used to refer to the C_i of utterance n, and C_i(U_m) to refer to the C_i list of utterance n. As a shorthand, the highest ranked forward-looking center C_i of utterance n is called the preferred center, or C_P(U_n). Roughly speaking, C_P(U_n) contains all entities that are referred to in utterance n: amongst this list is C_i(U_n). C_i(U_m) is the most highly ranked element in C_i(U_n) that is realized in U_m. The entities are ranked on the C_i list with respect to a grammatical role obliqueness hierarchy: subject, object, indirect object, other subcategorized functions, and adjuncts. They define four transitions between utterances depending on the relationship between C_i(U_n), C_i(U_m), and C_i(U_m+1), shown in Table 5.1.

Now we can give the two rules of centering:

Rule 1: If any element of C_i(U_n) is realized by a pronoun in U_m, then the C_i(U_m+1) must be realized by a pronoun also.

Rule 2: Transition states are ordered. Continue is preferred to Retain is preferred to Smooth-Shift is preferred to Rough-Shift.

The BFP algorithm is a SmAsh procedure, in which (i) Rule 1 is used during the Match phase, along with the other morphosyntactic constraints mentioned in the introduction, and (ii) The Select phase chooses the pronoun assignment(s) that result in the most preferred transition per Rule 2. This strategy correctly predicts that He and him in sentence (32a) refer to Terry and Tony, respectively, since this assignment results in a Continue relation, whereas the Tony/Terry assignment results in a less-preferred Retain relation. The algorithm also accounts for the oddness of sentence (32e), since assigning he to Tony results in a Smooth-Shift, whereas assigning he to Terry results in a Continue.

There are a few misconceptions about centering floating around the literature that I believe are worth clearing up. For instance, several previous authors (Hudson-D’Zmura 1989; Chambers and Smyth 1998) have claimed that the centering algorithm incorporates a subject assignment preference of the sort discussed in section 3.1. It does not; it instead incorporates what one might call a topic continuation preference, in which the C_i is taken to represent the sentence topic. Indeed, it is exactly this property that distinguishes a centering-based account from a purely salience-based strategy. Consider, for instance, example (33a-b), with possible follow-ons (33c) and (33d):

\[(33)\]
\[
\begin{align*}
\text{a.} & \quad \text{Terry is always willing to go sailing.} \\
\text{b.} & \quad \text{Tony dropped by his house yesterday.} \\
\text{c.} & \quad \text{He knocked but no one answered.} \\
\text{d.} & \quad \text{He wasn’t home.}
\end{align*}
\]

Whereas a subject preference predicts that He in (33c) refers to Tony, centering predicts that it will refer to Terry, since that assignment leads to a Continue transition (Tony gives rise to a Smooth-Shift). Whereas the subject preference gets it right in this case, it doesn’t always turn out that way: the preferred referent for He in follow-on (33d) is Terry, as predicted by centering. The reality is that discourses that end in a Retain, like (33a-b), often create an ambiguity for a following subject pronoun, since it is unclear whether or not the placement of a new entity in subject position was motivated by an intention to shift topic.

A second misconception is that centering provides an incremental mechanism for pronoun interpretation, modeling "a speaker’s immediate tendency to interpret a pronoun" (Brennan 1995). This is also not the case. Detailed discussion of this issue would take us too far afield; the reader is referred to Kehler (1997) for examples and further discussion. Briefly, the problem is that the preferred assignment for a pronoun in the BFP algorithm cannot necessarily be determined until the entire sentence has been processed. The reason stems from two properties of the algorithm: that determining the transition type between a pair of utterances U_n and U_m requires the identification of C_i(U_m), and a noun phrase (pronominal or not) can occur at any point in the utterance that will alter the assignment of C_i(U_m). This fact compromises the algorithm’s ability to model the effects that are a result of incremental pronoun interpretation, unlike systems that use preference-driven strategies more directly (e.g., the subject assignment strategy).

The main problem with centering for my purposes here, however, is that the notion of coherence it captures is primarily entity-based—for instance, of the type needed to account for the behavior of example (37)—and not of the type based on semantic relations and inference that is necessary to address the reference patterns we have seen thus far. Indeed, the constructs centering uses to assign referents are quite restricted: the C_i of the current and previous sentence and the C_i of the current sentence, neither of which are determined by semantic considerations. As such, centering offers no handle with which to explain why pronoun interpretations are different in several minimal pairs for which...
these constructs are invariant, including (17a–b), (18a–b), (22a) and (22c), and (24a–b). I therefore conclude that while attentional state and topic continuation are likely to be important influences in pronoun interpretation, others ultimately need to be accounted for as well. This leads us to the next section.

4.3 Integrating Information Structure and Inference

The data we have discussed so far offer contradictory evidence about the mechanisms that underlie pronoun interpretation. In section 3.3, I argued that previous approaches have failed to control for the type of coherence relation that is operative in the examples they consider. In section 4.1, however, I also claimed that it is not enough to only consider coherence. Instead, I argue (as I have elsewhere (Kehler 2002)) that an adequate analysis must capture the interaction between coherence establishment and information structural constraints on pronouns and their referents. A complete account of this interaction is no doubt a complex matter, and this short section will leave much open for future work. My goal here is instead to outline what I hope to convince the reader is a more promising direction for research than the Smash paradigm.

As I indicated in section 3.3, different types of coherence are associated with different types of inference processes used to establish them. One might therefore expect their interactions with information structure to vary as well. Since herein will lie the answer to why different pronoun interpretation preferences appear to be in force when different coherence relations are operative. I will consider this question in the context of some of the pronoun interpretation data we have discussed thus far, particularly with respect to the three coherence relations introduced in section 3.3.

Let us begin with passage (21), repeated as (34), which is an instance of the Parallel relation:

(34) Margaret Thatcher admires Hillary Clinton, and George W. Bush absolutely worships her.

Recall that hearers universally assign Clinton as the referent of her in (22) if it is unaccented, even though world knowledge would strongly suggest Thatcher. Such examples initially appear to provide support for a grammatical role parallelism preference as advocated by Smyth (1994) and Chambers and Smyth (1998). However, as we saw in section 3.2, there are a variety of other examples that refute the existence of such a general preference, particularly ones characterized by coherence relations other than Parallel.

It turns out that the strong bias toward grammatical role parallelism evident in examples like (34) is a result of the interaction between information structural constraints imposed by Parallel relations and rules of accent placement in English, and thus cannot be profitably attributed to a pronoun-specific interpretation preference. As I indicated in section 3.3, the inference process associated with Parallel and other relations in its class first identifies pairs of parallel entities and predicates as arguments to the relation, and then attempts to identify points of similarity and contrast among the members of each pair. The commonalities in such constructions create a common topic in the sense of B. Lakoff (1971a), which in turn serves as the background against which focal elements are introduced. For instance, the common topic for (34) could be paraphrased roughly as how politicians feel about one another.

As I describe in Kehler (2005), the mapping among parallel elements creates a situation in which any element in the second clause that is not coreferential with its parallel element in the first becomes part of the focus on the sentence, even if it denotes given information. General rules governing focus and accent placement in English then require that such constituents contain an accent, which in this case means placing accent on the pronoun. So if the referent of her is intended to be Thatcher, it must be accented as in (35a):

(35) a. Margaret Thatcher admires Hillary Clinton, and George W. Bush absolutely worships HER.

b. Margaret Thatcher admires Hillary Clinton, and George W. Bush absolutely worships THATCHER/#Thatcher.

The fact that this accent requirement is not the result of a pronoun-specific strategy is easily seen by replacing the pronoun in (35a) with the proper name Thatcher as in (35b), for which accent is still required. Deaccenting Thatcher here is felicitous because Thatcher and Clinton are not coreferent. Thus, the fact that a pronoun is used in (35a) and likewise in (34) is simply irrelevant: Constraints on pronominalization of the non-parallel referent Thatcher are met in (34), but independent constraints on deaccentuation are not.

We can now explain the effect we saw in section 2.2, in which a 'soft' preference for grammatical role parallelism appeared to interfere with a 'hard' gender agreement constraint in example (6), repeated as (36):

(36) ?? Margaret Thatcher admires Ronald Reagan, and George W. Bush absolutely worships her.

It is now clear that there is nothing 'soft' about this parallelism effect. For the pronoun to remain unaccented, it has to corefer with its parallel element, regardless of agreement considerations. The clash between gender and information structural constraints is irreconcilable in (36), hence its infelicity.

These same constraints on deaccentuation do not apply for relations outside of the Resemblance class, which is why felicitous examples that violate the parallel grammatical role preference are readily found. Recall from section 3.3 the class of examples from Stevenson et al. (1994), repeated here as (37a–b), which participate in the Occasion relation:

(37) a. John seized the comic from Bill. He...

b. John passed the comic to Bill. He...
Recall that Stevenson et al. found that hearers are more likely to interpret he to refer to John in passage (17a) and to Bill in (17b), despite the fact that in (17b) Bill is mentioned from within a sentence-final prepositional phrase. This lends support for a preference for occupants of the Goal thematic role over occupants of the Source role, since the Goal is presumably more central to the final state of the eventuality. That is, such constructions create the expectation that the recipient of the goods will be focused on next (cf. Arnold (2003)).

Recall that the inference process used to recognize Occasion attempts to connect the initial state of the eventuality described by an utterance with the final state of the one that came before it. As such, at the time that a pronoun is encountered, the referent most attended to should be the one that is most prominent with respect to the hearer's conceptualization of the end state of the previous eventuality. While this will often be the subject of the preceding sentence, this is not always so, as we see from example (37b). We therefore predict that specific evidence for a Goal preference will only be found when an Occasion relation is operative; no such preference is expected for examples of other coherence relations, such as the Parallel relation in (38a):

(38)  

a. John passed the comic to Bill. He threw the book to Fred.  
(Parallel, He = John)

b. John passed the comic to Bill. He didn’t want it anymore.  
(Explanation, He = John)

As expected, the unaccented subject pronoun in (38a) can felicitously refer only to John. As such, there can be no general preference for Goals over Sources, as we only find a preference for Goals where we would expect it: in Occasion relations, which crucially involve connecting the content of an utterance to the end state of the previous eventuality.

The same point is demonstrated by example (38b), an instance of the Cause-Effect relation Explanation. The inference process associated with such relations attempts to identify a causal chain between the semantics of the clauses being related. In this case, the manner in which the referent is determined depends on how the contributions to the discourse instantiate pre-existing causal knowledge; in (38a), this knowledge favors John as the referent. We saw how different semantic predications can lead to different pronoun assignments in otherwise identical passages in our discussion of examples (24a–b) in section 4.1. However, while there is no explicit model of information structure or attention in Hobbs’s analysis, we have seen two reasons to believe that such a model is necessary.

The first reason is the existence of garden-path effects, such as we saw in example (32). While that example involved an Occasion relation, garden-paths can occur when Cause-Effect relations are operative also. Consider the adapted version of (24b) given in (39).

(39)  

The city council denied the demonstrators a permit because they decided that the best way to draw attention to issues is to advocate violence.

As in (24b), the pronoun they in (39) is intended to refer to the demonstrators, but in this case the crucial information that leads to the eventual coherence relation realized—and with it, the intended pronominal referent—comes too late after the pronoun to shift attention away from the city council in time. Therefore, since there is an initial information structural bias toward the subject the city council, the hearer interprets it as the intended referent and a garden-path results.

The second reason is the effect of passivization that we saw in examples (30a–b), repeated here as (40a–b):

(40)  

The demonstrators were denied a permit by the city council because...

a. ... they feared violence.

b. ... they advocated violence.

Recall that a purely coherence-driven strategy does not explain why hearers interpret they to refer to the demonstrators in both follow-ons, unlike the alternation seen in the active voice versions in (24a–b). There are crucial differences between the active and passive forms that are no doubt responsible. First, they are not on equal footing attentionally. In (24a–b), the non-subject potential referent is in object position, which is a relatively salient position. In (40a–b), on the other hand, the non-subject potential referent is embedded in a sentence-final adjunct by-clause, and thus much less salient. Second, they are not equivalent with respect to topichood. Whereas the subject and object positions of active clauses can both serve as topics (with perhaps a moderate preference for the subject), the subject is the presumed topic in a passive clause. Under the assumption that pronouns refer primarily to (potential) topics, the difference between (24a–b) and (40a–b) is consistent.

To summarize to this point, we have seen that a variety of surface-level preferences that have been posited in the literature are actually epiphenomena of interactions between information structure and coherence establishment processes. When the predictions of preferences such as grammatical role parallelism and those of the coherence analysis diverge, the predictions of the coherence analysis win. Note that we can now explain the previously mentioned contrasts between Kameyama’s examples (22a–c), repeated here in (41a–c).

(41)  

a. John kicked Bill. Mary told him to go home. [= John ]

b. Bill was kicked by John. Mary told him to go home. [= Bill ]

c. John kicked Bill. Mary punched him. [= Bill ]
As we might now expect, in the cases of Occasion in (41a) and (41b) we see the appearance of a subject assignment preference, whereas in the case of Parallel in (41c) we see the appearance of a grammatical role parallelism preference.

Whereas I have argued that the grammatical role parallelism preference is epiphenomenal, it is more appropriate to characterize the subject assignment preference as being derivative. The reason stems from its status as the default position from which to mark topic, coupled with the influence of topichood on pronoun interpretation. Consider again the difference between examples (41a) and (41b). These are both instances of Occasion; the only difference is which entity is more topical going into the second clause, that is, the entity in subject position. Importantly, however, occupants of other syntactic positions in active voice clauses can serve as potential topics to which coherence establishment can shift attention, whereas we have seen that this is not so readily done for passive clauses. Thus, it may not be appropriate to reduce the role of topichood to a single subject assignment preference, as the effect of being in subject position is almost certainly dependent on the particular syntactic construction that the subject participates in.

While I have argued that surface-level morphosyntactic cues are not the ultimate determinants of pronominal assignments, it must be pointed out that they do appear to have an indirect role, in that they may affect what coherence relations are recognized (in ways, I might add, that are frankly not well-understood). For instance, strong syntactic parallelism may bias an interpretation toward the Parallel relation more than a less parallel structure would. The crucial point is that the connection between these superficial cues and pronoun interpretation is not direct, but mediated by the recognition of coherence.

This observation, in fact, underlies my primary criticism of more recent study of pronoun interpretation in the context of coherence establishment by Stevenson et al. (2000). Stevenson et al. compare two hypotheses, one based on ‘semantic focusing’ and the other termed ‘relational’. In semantic focusing, focus is placed on an entity central to the final state of an event as posited in Stevenson et al. (1994), but an intervening connective can alter it. The ‘relational’ hypothesis is basically the Hobbs analysis; the view I have presented falls into this category also. At issue is whether connectives directly affect focus, or whether they affect coherence establishment, which in turn affects focus.

Based on their results (which space concerns preclude me from discussing in detail), Stevenson et al. conclude that the semantic focusing account is superior. However, there are a variety of facts discussed herein with which their account is inconsistent. First, they assume that there is a one-to-one mapping between coherence relations and predicted referents, which as we have already seen is not the case. Passages (24a–b), for instance, are both instances of the Explanation relation, but the preferred pronominal assignments are different as predicted by default world-knowledge relationships. Second, Parallel relations are conspicuously absent from their study, and it is hard to see how they could be accounted for with semantic and connective focusing. For instance, their analysis predicts that the pronoun assignments for passages (37b) and (38a–b) will be the same: semantic focusing starts with Bill, and there is no connective to change it. Third, and relatedly, since there cannot be a shift from the semantic focus without a connective, their analysis makes the wrong predictions for the garden-path example (32b).

In their analysis, Patients receive semantic focus over Agents, and thus reference to Tony is predicted to be preferred (and therefore felicitous). Fourth, the analysis does not account for the differences between active and passive voice in examples like (24a–b) and (40a–b), since the thematic roles and connectives are constant in these passages. Finally, their analysis cannot predict the ambiguity of passages like (23), repeated here as (42), since the factors that determine semantic focusing are obviously invariant between the two readings:


In the analysis presented in this section, the ambiguity is predicted by the existence of two distinct coherence construals for this passage. All these problems notwithstanding, Stevenson et al.‘s model rightly distinguishes the dynamic effects that linguistic structure has on discourse state from the process of pronoun interpretation itself, and thus it is not a SMASH procedure.

5. Conclusions

To conclude, an adequate model of pronoun interpretation must explain a variety of facts, including (i) why in some cases a pronoun cannot be felicitously used even when only one referent is available, (ii) why we find evidence for so-called interpretation ‘preferences’, and why different preferences appear to prevail in different contextual circumstances, and (iii) how pronouns could have the effect of facilitating pronoun interpretation within a broader theory of discourse processing and comprehension. The commonly assumed SMASH paradigm fails at all three, and thus should be abandoned as a framework for theorizing about pronoun interpretation. I believe that much of the confusion in the literature is a direct result of casting pronoun interpretation in terms of this untenable paradigm. Furthermore, no matter what paradigm is assumed, future psycholinguistic studies should take great care in controlling for the operative coherence relations in their stimuli, since different relations will provide support for different preferences.

Much of the complexity that some SMASH procedures consider to be part of pronoun interpretation undoubtedly originates from sources independent of it. In this chapter I have focused on one: the interaction between information structure and coherence establishment. Because coherence establishment processes fundamentally differ with respect to the type of coherence they establish, we can explain why we see evidence for various (mostly epiphenomenal)
preferences only in certain contextual circumstances. Furthermore, because the complexity in the data results from these processes and not pronoun interpretation itself, we can explain how pronouns can actually facilitate comprehension by expressing topic continuity, as opposed to more elaborated referential forms that would require additional processing.

Several of the phenomena I have discussed as problematic for previous approaches raise questions about the respective roles of activation, subjecthood, and topichood in pronoun interpretation, including the examples discussed in sections 2.1 and 2.3, and the active-passive incongruences discussed in section 4.3. A full investigation of these and other remaining questions will require a more precise and elaborated model of discourse processing than I have provided here, and thus are subjects for future work.

NOTES

1. Thus, I will not argue that the SMASH paradigm should be abandoned in computer algorithms developed with engineering goals in mind (for instance, to achieve high accuracy on a pronoun interpretation task). Limits of current technology (to parse accurately, to model world knowledge, etc.) and the imperative to handle frequently occurring patterns reliably (which often means ignoring rarer phenomena that might nonetheless be the most illustrative for human language processing) are two of several reasons why such systems should not necessarily strive to model human language processing.

2. For rhetorical convenience, I will use the term ‘information structure’ in a particularly general way for the remainder of the paper, to include not only standard pragmatic notions such as topic, but also linguistic factors that bear on the cognitive notion of attention in language processing. The proper division of labor among these concepts is not always clear, but it is not my goal here to go into such questions in any detail.

3. All of the analyses of pronouns that I discuss are restricted to unaccented cases, unless otherwise indicated. In the case of example (6), accent must be placed on the predicate absolutely worships instead. The reason for this will become clear when we consider this example again in Section 4.3.

4. This is the case for the Lappin and Leass algorithm and Hobbs algorithm, for instance. Centering does not specify its ranking in enough detail to determine a prediction for this example.

5. Indeed, a reference to Rice in (13) would make it a so-called contrastive topic (Büring 1999, inter alia), which would in turn require it to be accepted.

6. A more recent paper by Stevenson and colleagues (Stevenson et al. 2000) does address coherence; I discuss their paper in section 4.3.

7. All of the centering constructs described here are from GJW, except that BFP split out GJW’s Shift transition into two subcases and modify Rule 2 accordingly. The algorithm for interpreting pronouns using these constructs is BFP’s.

8. To be clear, this does not mean that we will never see grammatical role parallelism in examples involving coherence relations like Occasion, Explanation, and Result. We clearly do, for instance, in any case in which a subject pronoun is identified with a subject antecedent. The point is that there is always an explanation for the particular choice of reference assignments other than the existence of a grammatical role parallelism preference.

REFERENCES


Good-Enough Representation in Plural and Singular Pronominal Reference

Modulating the Conjunction Cost

SUNGRYONG KOH, ANTHONY J. SANFORD, CHARLES CLIFTON JR., EUGENE J. DAWYDIAK

This chapter concerns the nature of representations set up when two individuals can be referred to by means of plural pronouns. The conditions under which plural pronominal reference is possible, or preferred from a processing perspective, has been the subject of recent psycholinguistic research (Albrecht and Clifton, 1998; Koh and Clifton, 2002; Eschenbach, Habel, Herweg, and Retikampfer, 1989; Kaup, Keiter, and Habel, 2002; Moxey, Sanford, Sturt, and Morrow, 2004; Sanford and Moxey 1995).

As part of the attempt to understand plural reference, the concept of complex reference object has been used as an explanatory construct: plural reference is deemed possible if a plural reference object can be formed (see Eschenbach et al., 1989, Kaup et al., 2002, and Koh and Clifton, 2002 for descriptions). The issue then becomes translated into what permits a complex reference object to be formed. In this chapter, we are concerned with singular and plural anaphoric reference to one or more individuals evoked by a previous sentence, as in (1):

(1) Stan and Pam asked the usherette for assistance.

Such a sentence, containing a noun phrase (NP) with two conjoined individual names (denoting atomic individuals), can be followed by a pronominal reference either to the pair of people (They appreciated the help) or to one individual (e.g., He appreciated the help), and still be intelligible. The same applies
Reference

*Interdiscipliary Perspectives*

Edited by
Jeanette K. Gundel and Nancy Hedberg

OXFORD UNIVERSITY PRESS
2008