Discourse Coherence and Attention: A Theory of Pronouns

Abstract

Traditionally, pronouns are treated as ambiguous between bound and demonstrative uses. Bound uses are non-referential and function as bound variables, and demonstrative uses are referential, and pick out objects as determined by their linguistic meaning and surrounding non-linguistic cues – e.g., an accompanying demonstration or an appropriate and adequately transparent speaker’s intention. In this paper, we challenge tradition and argue that both demonstrative and bound pronouns are dependent on, and co-vary with, antecedent expressions. Moreover, the semantic value of a pronoun is never determined, even partly, by extra-linguistic cues; it is fixed, invariably and unambiguously, by features of its context of use governed entirely by linguistic rules. We exploit the mechanisms of Centering and Coherence theories to develop a precise and general meta-semantics for pronouns, according to which the semantic value of a pronoun is determined by what is at the center of attention in a coherent discourse. Since the notions of attention and coherence are, we argue, governed by linguistic rules, we can give a uniform analysis of pronoun resolution that covers bound, demonstrative, and even discourse bound (“E-type”) readings. Just as the semantic value of the first-person pronoun ‘I’ is conventionally set by a particular feature of its context of use—namely, the speaker—so too, we argue, the semantic value, e.g., of ‘he’ is conventionally set by a particular feature of its context of use. In this paper, we elucidate what this feature is and how it works.

1 Introduction

The third-person pronoun ‘he’ doesn’t seem ambiguous in the way, for example, that the word ‘bank’ is.¹ Yet its tokens can take different semantic values in different contexts of use, and they seem to do so in entirely different ways. Uttered pointing at Bill, ‘he’ refers to Bill; and pointing at Sam, it refers to Sam. Yet in ‘Every man shaved himself’ and ‘A man walked in and he sat down,’ its occurrences do not seem to refer at all: the interpretations in these cases are dependent on, and co-vary with, the possible instances in the ranges of the quantifiers ‘Every man’ and ‘A man’. No doubt, linguistic meaning plays some role in determining the semantic values of the third-person male pronoun in these diverse contexts, but what are these semantic values exactly and exactly what role does meaning have in delivering them?

¹ Some think that what we usually regard as an ambiguous word is really two words. We reject this view (see Hawthorne & Lepore, 2011). But the point in the text can be restated: there doesn’t seem to be a plurality of ‘he’s in the way that there is a plurality of ‘bank’s.
The standard answer begins by positing an ambiguity: bound uses of 'he' are non-referential and have the semantic function of bound variables. Demonstrative uses, by contrast, are referential and take as a semantic value their referent, an object picked out jointly by linguistic meaning and a further cue—an accompanying demonstration, an appropriate and adequately transparent speaker's intention, or both. Such supplementary mechanisms are standardly understood to be extra-linguistic or pragmatic cues. In particular, they are not linguistically encoded.

In this paper, we challenge the standard view by arguing that the semantic value of a demonstrative pronoun is analogous to that of a bound variable, and so, it too is dependent on, and can co-vary with, an antecedent expression. Moreover, this semantic value is not determined, even partly, by extra-linguistic cues; rather, it is fixed, invariably and unambiguously, by features of its context of use, which themselves are entirely governed by linguistic rules. Thus, we reject the ambiguity analysis as well as its appeal to extra-linguistic determinants of semantic values. In effect, our treatment of 'he' is much like the standard treatment of the first-person pronoun 'I'; just as the latter's semantic value is set by a particular feature of its context of use—namely, the speaker—so too, the semantic value of 'he' is set by a particular feature of its context of use. Our challenge is to elucidate what this feature is, and how it works.

Before we proceed, we want to register that our view has far-reaching philosophical consequences beyond the semantics and meta-semantics of pronouns. Philosophers often appeal to context-sensitivity in analyzing philosophically interesting terms, such as 'know', 'good', 'might' and 'all', where their semantic values are alleged to be determined by contextual parameters such as the epistemic standards, standards of precision, value commitments, implicit restrictions on quantifier domains, etc. In building theories about how the values of these parameters are fixed in context, philosophers have been invariably guided by the standard view of the meta-semantics of pronouns: the values of these parameters are determined jointly by linguistic meaning and non-linguistic cues - a speaker's intention, or an epistemic cue that renders the intention manifest, or both. Often, in fact, both the linguistic analyses of these terms, and the philosophical arguments exploiting these analyses, depend on this model of context dependence. If we are right, however, this view is

2 Proponents of the view are many. Kaplan (1989b), King (2014, a.; 2014, b.), Neale (2004), Reimer (1992) all advocate some version of a view of this general kind, though there are important differences between them that needn't concern us here. We believe these views are misguided, but here we focus primarily, though not exclusively, on our positive view.

3 That the meta-semantics of context-sensitive expressions generally should be modeled on the received view of the meta-semantics of demonstrative pronouns is often either implicitly assumed, or explicitly endorsed. (See. e.g. Stanley & Szabó Gendler, 2000, King & Stanley, 2005, King, 2014 b., Cohen, 1999, Dowell, 2011.)
misguided even for pronouns: the role of context in fixing the semantic values of pronouns is fully constrained by grammar. This suggests we should be wary of meta-semantic theories about other philosophically interesting expressions built on the standard model of the meta-semantics of demonstrative pronouns. In fact, our view recommends a project of its extension to these other cases of context sensitivity; if this is right, then philosophers will have to rethink the role of context-sensitivity in their arguments.

In what follows, we proceed to develop our approach to the semantics and meta-semantics of pronouns. We call it the *Attention-Coherence Approach*, for it contains two crucial components—a ranking of candidate interpretations according to relative prominence (the attention component), and implicit mechanisms affecting this ranking (the coherence component). The role of both components in our theory will be made precise in what follows. We begin in the next section to develop the attention component; the theory will not receive its full shape until after we present the coherence component in later sections.

2 Attention

On a rough first pass, our theory is that pronouns, as a matter of linguistic meaning, (their *character*, if you like), always pick out whatever is ‘in the center of attention’ in a discourse. Our theory borrows resources from what is known in linguistics as *Centering* theory (Sidner, 1983; Grosz, Joshi, & Weinstein, 1995; Bittner, 2014), according to which, candidate resolutions of a pronoun in a discourse are ranked according to their relative prominence: those higher in the ranking are preferred as interpretations of pronouns over those lower in the ranking. We call this ranking the *attentional state* of an ongoing discourse. As a discourse progresses, utterances bring new candidates into focus and adjust the prominence of old ones, thereby changing the attentional state. Pronouns, as a matter of linguistic meaning, always pick out the top-ranked candidate in the current attentional state; and further, the attentional state itself is maintained through linguistic rules.

To motivate the dynamics of an attentional state, we begin by considering two ways in which a speaker might use (1), and its associated contrasting interpretations:

1 A man walked in. He sat down.

---

4 It is important to note that, while we borrow resources from Centering theory, we depart from it in spirit. For example, Centering theory is traditionally understood as *pragmatic* theory, in the sense that the preference for higher ranked entities is meant to be input to a holistic process of interpretation that resolves ambiguity by finding the overall most plausible interpretation. We, instead, hold that the ranking is maintained and updated by linguistic rules, *and moreover* has a genuinely linguistic effect on interpretation insofar as the preference for the higher ranked entities is built into the meaning of the pronoun.
One reading is that (1) is part of the flow of a continuing narrative. In this case, ‘He’ is interpreted as co-varying with ‘A man’. More precisely, on this construal, the discourse in (1) is true just in case some man both walked in and sat down. The second reading is that the speaker is describing past events in relation to the current situation. For example, the speaker might utter ‘He,’ while pointing at a small child, Bill, who is present nearby. In this case, ‘He’ is interpreted as referring to Bill. The discourse is true just in case a man came in and Bill sat down.

Thus, speakers can use (1) in different ways, and hearers can point to reasons why certain interpretations are recovered in particular circumstances, and others are not. Pronoun resolution is guided by an implicit organization that knits the information in the speaker’s discourse together. On its anaphoric reading, the discourse begins with a description involving the indefinite ‘A man’ and then proceeds directly to develop a narrative: on this organization, ‘He’ is interpreted as dependent on ‘A man’. On its deictic reading, the speaker marks a transition from talk of past events to the present situation with a demonstrative gesture: on this organization, ‘He’ is interpreted as dependent on the individual demonstrated, and not on ‘A man’.

To formalize these ideas into a framework that allows us to provide logical forms for discourses containing pronouns, we first need a model for bound variables that captures the relevant interpretive dependencies. We provide such a model by cleaving closely to the familiar Tarskian machinery for variable binding. In a Tarskian framework, recall, the truth of a formula is specified relative to an assignment of values to variables. A variable \(x_i\) is interpreted by retrieving the \(i^{th}\) element of a current assignment. Bound variables have dependent interpretations because quantifiers vary the assignments that are in force within their scope. An obvious shortcoming is that interpretation on a Tarskian approach is limited to the syntactic scope of a quantifier. But, in discourse, quantifiers can introduce dependencies that persist across subsequent sentences, as in the anaphoric reading of (1). We therefore need a framework for extending the familiar notions of variable and binding to the kind of dependent interpretations we find in (1). To this end, we exploit an approach known as dynamic semantics, which is designed precisely for this purpose.\(^5\)

\(^5\) See Dekker (2011) for an introduction and overview to dynamic semantics. We employ dynamic semantics as a way of formalizing linguistic dependencies relevant for the interpretation of demonstrative pronouns, because it is the simplest and most elegant framework for handling such dependencies (including the problematic donkey and cross-sentential anaphora). However, bear in mind that our philosophical point – namely, that pronominal interpretation is fixed through linguistic rules – is independent of the particular way we formalize it, and potentially, could be implemented in other semantic frameworks (e.g., see Elbourne, 2005.).
In dynamic semantics, just as in a Tarskian framework, variable binding creates dependent interpretations by varying assignment functions. However, unlike the standard Tarskian approach dynamic semantics assigns truth conditions to formulas relative to a pair of assignment functions, not to just a single one. The first assignment function is old hat: it accounts for interpretive dependencies potentially already available in the discourse for interpreting a formula; the second one, however, accounts for interpretive dependencies the formula itself makes available for interpreting subsequent discourse. So, think of a formula, in part, as an instruction to update the available interpretive dependencies from an input assignment to an output one, in effect, as an extra bit of bookkeeping within a discourse – tracking potential interpretive dependencies from a prior discourse, encoded in the input assignment function, as well as tracking interpretive dependences that a formula introduces for interpreting subsequent discourse, encoded in the output assignment function. Just as in Tarskian semantics, we still interpret a bound variable $x_i$ by retrieving the $i^{th}$ element of the input assignment, but now we can define the existential quantifier in a way that allows for alternative output assignments, so that the interpretation of subsequent variables can continue to take a bound interpretation throughout the subsequent discourse. Overall, a formula is true on a given input assignment if there is an output assignment that makes it true. For example, due to the existential quantifier in (1), the (formula corresponding to the translation of the) first sentence in (1) is true on a given input assignment function just in case there is an output assignment function according to which it’s true that some unspecified man – a witness for the existential – walked in. The (formula corresponding to the translation of the) second sentence, then, 

\[ \text{[C]} \]

\[ \text{[C]}[(g, h)] \]

A detailed exposition of our formal system is in the Formal Appendix. In brief, we interpret formulas via updates that relate an input assignment $g$ to an output assignment $h$. The simplest update contributes information fixed by a condition $C$, written as $[C]$ and interpreted as a partial identity relation between assignments. If $g$ is an input assignment, and $h$ an output one, $[C](g, h)$ iff $g=h$ and $C$ holds on the interpretation of variables given by $g$. In standard fashion, a sequence of updates $H$ and $K$ is represented by a single update $H; K$ that performs the update $H$ followed by $K$. An update is true with respect to an input assignment iff the input assignment is related to some output one by the update relation. In addition to the simplest updates, we need updates that affect the prominence ranking of the candidate referents. We introduce these presently. Another systematic formal development of an account that combines dynamic semantics and Centering theory as a framework for handling anaphoric relations is Bittner (2014). Our approach is in the spirit of Bittner’s, but crucially different. First, we introduce updates that correspond to implicit shifts in attention. Furthermore, we have a very different take on the conventionality of different attention-shifting mechanisms, pointing to their linguistic nature.

Note that any output assignment function that makes it true that there’s a man who walked in will do. This is the sense in which the witness for the existential is an “unspecified” man – any one will do.
takes this output assignment as a new input one, and the pronoun ‘he’ can then be interpreted as a bound variable that takes as its value whatever the witness for ‘a man’ was in this new input assignment function, rendering the anaphoric reading of (1). Overall, the discourse is true on the initial input assignment just in case some man walked in and sat down.

But how does the interpretation of ‘he’ in (1) get bound by the existential quantifier ‘A man’? In Tarskian and dynamic semantics alike, variables behave like temporary names. A quantifier like \( \exists x \) re-defines what \( x \) names within its scope; we can then interpret occurrences of \( x \) as local names for potential witnesses of the quantifier. So, in essence, if ‘A man’ is translated with an ‘\( \exists x \)’, then ‘he’ gets translated with an ‘\( x \)’, thus looking for the \( i \)th element of the input assignment; this element, given the prior discourse, will be required to be a man who walked in.

We adopt the main ideas of dynamic semantics, but describe how discourse dependences are created differently. Instead of organizing interpretive dependencies by name, we opt for prominence. To achieve this, we treat assignments as stacks, in the sense of theoretical computer science. Each assignment specifies a sequence of possible individuals ordered by prominence. The most prominent individual is in initial position – the top of the stack – and ones in subsequent positions are deeper in the stack, thereby, receding in prominence. Quantifiers introduce new possibilities for dependent interpretations in subsequent discourse by pushing values onto the stack: inserting them at a specified position, for example, at the top, thereby, decreasing the prominence of other candidates in the discourse. In this way, they vary output assignment functions. We thus treat a variable not like a temporary name, but rather like a marker of prominence; for example, we can define a variable that picks out the top of the current stack, and thus, co-varies in its interpretation with whichever quantifier most recently pushed a new value there.\(^8\) We use the symbol ‘@’, then, as such a variable that is interpreted relative to an assignment \( g \) as specifying the top-ranked element of \( g \). (The mnemonic is that ‘@’ is at the center of attention.)

The meaning of ‘@’ is our first approximation to the meaning of an English pronoun. To illustrate, let the formula \([\text{man(@)}]\) be a condition that requires that whatever is at the center of attention be a man. And let ‘( \( \alpha \) )’ indicate a dynamic existential quantifier, which changes the input assignment function by introducing a new unspecified top-ranked individual at the center of attention, while all other

\(^8\) Nothing about using stacks to model prominence requires a commitment to the discourse scope of dynamic semantics. This tool can be used with traditional syntactic binding as well. For example, de Bruijn indices offer a stack-based method for writing and interpreting terms of the simply-typed \( \lambda \)-calculus (de Bruijn, 1972). However, we think stack-based variables and discourse scope of dynamic semantics are particularly powerful together for formalizing the logical forms of natural language discourses in perspicuous ways. Thus, we shall combine the two tools.
candidates are demoted one position in the ordering. This update is our first approximation to the meaning of the English indefinite article. Then, (1), on its ‘anaphoric’ interpretation, is represented as (2):

\[
\langle \alpha \rangle; \text{[man(@)]}; \text{[walk.in(@)]}; \text{[sit.down(@)]}
\]

(2) begins with an existential quantifier, corresponding to the indefinite article, that introduces its witness as the top-ranked resolution for subsequent variables; the witness is then constrained to be a man, to have walked in, and then—once picked up by the anaphoric pronoun—to have sat down. Thus, the formula is true just in case some man walked in and sat down.10

By contrast, consider a deictic reading of (1), uttered, suppose, while pointing at a small child, Bill, as in (3):

\[
\text{(3)} \quad \text{A man walked in. He [pointing to a small child, Bill] sat down.}
\]

We know its indefinite NP puts its witness at the top of the ranking, while the pronoun, as a matter of linguistic meaning, resolves to the top-ranked candidate. Yet, ‘He’ in (3) doesn’t resolve to the witness for ‘A man’ because of a further shift in attention—one that is triggered by a pointing gesture. This shift updates the attentional state so that the entity indicated by the pointing gesture now becomes top-ranked.

More precisely, we introduce a family of updates, written ‘\( \langle \pi c \rangle \)’, where ‘\( \pi \)’ corresponds to the act of pointing and ‘\( c \)’ to some individual being pointed at.11 This update stores \( c \) as the top-ranked entity, and (as always) pushes all others down one position deeper in the ordering. It thereby represents the effect of the pointing that accompanies the use of the pronoun. (3), then, is represented formally as (4):

\[
\langle \alpha \rangle; \text{[man(@)]}; \text{[walk.in(@)]}; \langle \pi b \rangle; \text{[sit.down(@)]}
\]

9 This is, thus, a non-deterministic update, in the jargon of dynamic semantics. More precisely, \( \langle \alpha \rangle \) relates an input assignment function \( g \) to an output one \( h \) iff \( h \) potentially differs with \( g \) in the top position, and for every subsequent position \( i \), \( g_i = h_{i+1} \).

10 Recall that the formula is true for an input assignment \( g \) iff it is true for some output assignment. In fact, though, the formula in (2) has no free occurrences of variables, and so, its truth doesn’t depend on an initial input assignment at all. See the Formal Appendix for the relevant formal definitions.

11 More generally, ‘\( \langle \alpha \rangle \)’ corresponds to the contribution of indefinites, and ‘\( \langle \pi c \rangle \)’ to the contribution of definites, where ‘\( n \)’ is an individual denoting expression. We do not attempt to offer a full account of definites other than pronouns here.
In (4), we formalize ‘He sat down’ with ‘[sit.down(@)]’, just as in (2). But in its context in (4), the condition is true just in case Bill sat down. After the ‘(α)’ that introduces a man, we find an intervening expression ‘(πb)’, corresponding to the act of pointing, that updates the attentional state. As a result, whomever the speaker has pointed at—Bill—is at the top of the assignment when ‘[sit.down(@)]’ is interpreted.

Note that, as formalized in (4), changes of attention are reflected in logical form. Traditionally, a shift in attention provoked by a pointing gesture accompanying the use of a so-called demonstrative pronoun has been treated as a result of a pragmatic process in which the hearer comes to recognize the speaker’s referential intention by exploiting the epistemic cue provided by the speaker’s pointing gesture. The idea is that the only way to make sense of the speaker having chosen to accompany her use of the pronoun with a pointing gesture is that she intends to refer to some object roughly in the indicated direction. Crucially, the pointing gesture, although supplementing the incomplete linguistic meaning of a demonstrative, does not render the demonstrated object prominent as a matter of a convention, but merely serves as an epistemic cue to constrain the interpretation of a speaker’s utterance. It succeeds to help manifest the speaker’s intention just because it makes sense to point at an intended referent.

We adamantly reject this view. The change of attention contributed by the pointing gesture in (3) should be reflected in logical form, because, as we shall show presently, the sorts of gestures that secure pronoun resolution in cases like (3) do so as a matter of grammar; as such, these gestures are integral to the linguistic utterance itself much like other linguistic features are. The rules of language dictate that a demonstrative act of pointing introduces a new candidate referent at the top of the stack. This particular utterance of (3) includes the action of pointing at Bill, which contributes to meaning in a way governed by rules of language.¹²

We submit several reasons for treating deictic gestures as governed by grammar. First, as argued by Kendon (2004), non-verbal means of indicating an entity are governed by linguistic rules sensitive to form, meaning, and the relationship with ongoing speech. For example, English speakers seem to count deixis as well-formed only when the pointing action is synchronized in an appropriate way with the prosody of the accompanying utterance. As Kendon illustrates, English speakers often repair utterances when their performance fails to align speech and gesture in

¹² For an example of a grammar integrating gesture and speech in this way, see Alahverdzhieva et al (2011).
time, as one would expect if the requirement for synchronicity were dictated by an underlying convention.\(^{13}\)

In addition, Kendon observes that although English speakers can use a range of hand shapes when they indicate an object, their particular choice affects the semantic contribution they make. For example, a gesture with the index finger and thumb extended, and the other fingers curled closed, uses the direction of the index finger to single out an object as an individual distinct from its alternatives. By contrast, a gesture with a flat hand open toward the audience, its four fingers extended in a tight line, uses the direction of the fingers to exhibit an object as a representative of a broader class. Moreover, there are indefinitely many other possible gestures that are not typically used as deictic gestures for English speakers though they could have been. For example, English speakers use the thumb in the “thumbs up” hand shape, with the thumb extended from a tight fist, to demonstrate movement in the direction that the thumb points, but normally not to demonstrate an object located in that direction, though there’s nothing in the gesture itself that could have precluded it from playing this other function. The apparently arbitrary fact that some spatially directed actions are taken to indicate objects—while apparently analogous ones are not—shows convention at work in English speakers’ demonstrations. Had the convention been different, the “thumbs up” gesture could have served as a deictic one; this possibility is certainly not ruled out on grounds of rationality. Numerous languages, unlike English, allow speakers to indicate objects by deictic gestures of the lips, while not exploiting the extended index finger as a pointing gesture. Cuna, a language spoken in Panama, is one of many examples of such languages (Wilkins, 2003).\(^{14}\) Such conventionality and variability is a hallmark of linguistic meaning.\(^{15}\)

\(^{13}\) This is similar to the way prosodic focus is grammatically constrained to appear at a particular position with a particular contour, in order to contribute a particular semantic contribution.

\(^{14}\) Sarah Murray (pc) pointed out to us that a similar phenomenon is found in Cheyenne.

\(^{15}\) We understand convention in the sense of Lewis (1969), that is, as a solution to a reoccurring coordination problem. This notion of a convention requires arbitrariness in the choice of the solution to the coordination problem; there needs to be at least one alternative solution that is not ruled out on the basis of rationality. Thus, the arbitrariness manifested in these variations across linguistic communities showcases the conventionality of these gestures. Perhaps, one might still wonder whether we should understand these conventions as linguistic; we think the answer is clearly positive, since the deictic gesture has a direct impact on the interpretation of linguistic utterances. In this regard, deictic gestures are no less linguistic than, say, prosody.
Further, in typical cases, these acts that are signaling shift in attention are *indispensable* for the appropriate interpretation. In (3), where a candidate antecedent for the pronoun is linguistically available, explicit signaling of a shift of attention is necessary to establish a deictic reading, even to a particularly salient individual in the situation of utterance; otherwise, grammar seems to commit the speaker to an anaphoric reading of the pronoun. Other epistemic cues that might indicate real-world salience are not enough in the absence of such a signal. Even when Bill is jumping up and down in the faces of all the interlocutors—and thus, attracting the psychological attention of the interlocutors in an extremely compelling way—unless there is an overt signal that establishes Bill as the referent of 'he', the audience will still prefer the default anaphoric reading. Indeed, some means of explicit signaling of a shift of attention is required even when a linguistic antecedent is unavailable, such as when a pronoun is used deictically in an utterance that initiates a conversation. If deictic gestures were merely epistemic cues on a par with other epistemic cues that suggest real-world salience, this would be quite mysterious.

But don't we still need some pragmatic way of determining the *demonstratum* of a demonstrative gesture? We think not. We represent the pointing gesture in (3) as an *act of pointing at Bill*, and not as an act of pointing *simplitcer*, with the intended *demonstratum to be determined by context*. In other words, we are denying that pointing is semantically interpreted as having some uniform character that, given a context, determines a referent. An act of pointing at Bill, say, is not semantically interpreted as *pointing at x*, where *x* is contextually determined to be Bill. Rather, we think of pointing gestures as ambiguous between multiple possible forms, e.g., *an act of pointing at Bill*, or *pointing at Bill's shirt*, or *pointing at Bill's button*, etc. Pointings are thus, on our account, ambiguous. The interpretive work in (3) comes in settling whether the *form* of the gesture is one of pointing at Bill or of pointing at something else. This interpretive work of disambiguation is part of pre-semantics, in the sense of Kaplan (1989) – namely, the interpretive work needed to settle the *linguistic form* of an utterance. This should not be confused with the interpretive

---

16 Note that this demonstrates that “real-world” salience does not suffice for pronoun resolution. Our notion of prominence, in contrast to the intuitive, “real-world” notion of salience, is linguistically governed; what’s prominent, for us, is determined by rules of language.

17 As we are arguing, grammar specifies a diverse set of resources for raising entities to prominence (deictic gestures and indefinite NPs among them). Our formalism is expressive enough to make such resources available even for utterances that do not (merely) follow up an ongoing discourse. The precise prediction of our view is therefore not that deixis without demonstration is impossible at the beginning of a discourse, but that it succeeds only for utterances that recognizably accomplish acts that independently require construing the referent as the center of attention in the current state of the ongoing discourse.
work involved in the semantic interpretation of an already given linguistic form. Though, according to us, general epistemic cues play no role in semantic interpretation, they can play a role in allowing observers to recognize which form of a pointing gesture is in play. But this is the same sort of role they’d play in helping to disambiguate any other ambiguity (e.g., in “John was close to the bank when he fell into the river.”). And even here, just as with other ambiguities, conventions governing demonstrative actions constrain the possible disambiguations. As we have seen, a flat hand shape with the fingers towards the audience allows for a certain range of interpretations, but not others; similarly for an extended index finger.

To sum up: there are reasons to treat the effect of an attention shift triggered by an indefinite NP or an act of demonstration as governed by rules of language and as represented in logical form. And the recognition of these systematic effects on the prominence ranking of candidate referents in a context allows us to assign a single meaning to each pronoun that fully determines its resolution in a given context—roughly, pronouns pick out the most prominent candidate, the one at the center of attention. Refinements, however, are clearly required; we turn to them directly.

3 Refinements

Obviously, the idea that ‘he’, as a matter of meaning, takes on the most prominent candidate interpretation, however appealing as an intuition this may be, cannot be the whole story. ‘He’ in (5) cannot have an interpretation that depends on, and co-varies with, ‘A girl’. ‘A girl’ may introduce a prominent candidate interpretation for subsequent dependent elements, but not for ‘He’. ‘He’ must be resolved to a male.

5 A girl came in. He sat down.

Likewise, ‘She’ in (6) cannot be interpreted as co-varying with the witness for the quantifier ‘Some girls’. An interpretation that depends on ‘Some girls’ involves a plurality, whereas ‘she’ must have a singular interpretation.

6 Some girls came in. She sat down.

Finally, ‘we’ and ‘they’ in (7) cannot be interpreted as co-referential since ‘they’ requires a third-person interpretation—one disjoint from speaker and addressee. But ‘we’ is a first-person pronoun: its interpretation includes the speaker.

7 We came in. They sat down.

What these data reflect, of course, is that pronouns come with specific person, gender and number requirements that must be satisfied in the process of pronoun
resolution. Furthermore, these requirements are clearly linguistic: it matters whether the speaker utters ‘he’, ‘she’ or ‘they’ precisely because they differ in meaning.

Further, in (8), ‘him’ cannot semantically co-vary with Paul, even though Paul would normally be an eligible referent for ‘him’ (satisfying the person, gender and number requirement).

8 Paul met him.

The usual explanation for this failure is that reference is further constrained by syntactic principles—in this case, Condition B of Chomsky’s (1981) Binding Theory, which requires non-reflexive pronouns to be free in their governing category. Roughly, this means that the non-reflexive pronoun must not be bound by an expression in the same clause, that is, that it cannot share a clause with its own antecedent. So, in (8), where ‘him’ and ‘Paul’ are clausemates, ‘him’ cannot refer to Paul on pain of violating Principle B.

---

18 Of course, it can sometimes happen that a speaker uses a third-person pronoun ‘he’ to refer to the speaker. Witness a Perry-type scenario (Perry, 1979), where the speaker seeing, but not recognizing, himself in the mirror says, “He is the messy shopper.” For another example, consider a piece of discourse whereby the speaker is distancing himself from his earlier self, as in the following: “That day a boy was born. He was named John. That boy was me.” However, such cases are arguably marked in the sense that the exploitation of an extant rule makes for the surprising effect. These cases can be accommodated if one and the same object can be represented by distinct discourse referents (see Cumming, 2014). Note that we do not discuss here the appropriate treatment of the first-person singular pronoun ‘I’ and the second-person pronoun ‘you’. One way to do so would be to reserve a particular position on the list of salient referents for the speaker and the addressee respectively. One might then argue that each utterance appropriately updates the stored value for the speaker and the addressee, which is in line with the standard idea that the context includes the parameters for the speaker and the addressee. Then, ‘I’ requires its referent to be the speaker, and, roughly, ‘you’ requires its referent to be the addressee. This also fits the idea that ‘I’ and ‘he’ are represented by distinct discourse referents.

19 This is true regardless of whether a speaker can manage to ‘speaker-refer’ (Kripke, 1977) to a woman with the pronoun ‘he’ (or a definite ‘that man’).

20 More precisely, the antecedent of a non-reflexive anaphoric pronoun must not be local, or c-command the pronoun—a non-reflexive pronoun must not be bound in its local domain. By contrast, a reflexive pronoun must have a local antecedent (roughly, it must be bound by an antecedent in the same clause), according to the
With this in mind, we naturally refine our analysis by saying that, in general, a pronoun denotes the highest-ranked discourse referent whose value respects any specified grammatical features and any applicable syntactic principles. To interpret a pronoun, we consider the candidates in order of prominence within the current attentional state of the discourse, until we find one that satisfies the operative linguistic constraints.\(^{(9)}\) (9) integrates the constraints of semantics and syntax for the pronoun 'he'—it captures the character of 'he', if you like.

When interpreted with respect to an assignment function \(g\) modeling the available dependent interpretations on an occasion of use, 'he' denotes \(g\)'s highest-ranked entity that is singular, masculine, and disjoint from the speaker and addressee of the utterance, and that yields an interpretation where the occurrence of the pronoun is free in its governing category.\(^{(22)}\)

To incorporate (9) into our formal language, let 'he' be a predicate representing the constraints associated with the English third person singular male pronoun and '@he' be an individual expression that denotes the highest-ranked entity from the current assignment function that satisfies the property denoted by 'he'. (2) and (3), then, become respectively (10) and (11).

\[\langle \alpha \rangle; [\text{man}(\@)]; [\text{walk.in}(\@)]; [\text{sit.down}(\@he)]\]

Condition A; so, e.g. in 'Pete told Sam to explain himself', 'himself' can refer only to Sam, not to Pete. For more details, and precise definitions, see Chomsky, (1981).

\(^{21}\) The proposal to resolve the pronoun to the top-ranked candidate referent thus involves search, rather than a simultaneous imposition of potentially competing constraints of attention ranking and linguistic constraints associated with the pronoun. In other words, we are not saying the pronoun refers to the top-ranked discourse referent unless it violates a constraint; rather, it refers to the highest ranked discourse referent that satisfies the constraints. The idea of searching for the referent of an anaphor is a familiar ingredient of models of attention in the centering tradition (Grosz & Sidner, 1986; Brennan, Friedman & Pollard, 1987).

\(^{22}\) Roberts (2002) proposes similar rules associated with demonstrative pronouns, although with crucial differences. Instead of a linguistically governed notion of attentional prominence ranking, she uses a liberal notion of contextual, "real-world" salience. As seen above, this notion of "real-world" contextual salience does not suffice for establishing reference. Moreover, she treats semantic constraints on pronoun resolution as presuppositions of a pronoun, whereas we understand them as a part of their character. Finally, she does not explicitly build the constraints on syntactic binding directly into the constraints on resolution associated with the meaning of the pronoun.
This accommodates considerations raised above. However, even with these qualifications, we still fall short of a complete account. For example, (12) and (13) point to some lingering deficiencies.

12 A man met Sam. He greeted him.
13 John was disappointed with Tim.
   i. He fired him.
   ii. He did sloppy work.

In (12), English speakers would normally resolve the occurrence of 'He' to the man introduced in the first sentence, and, 'him' to Sam.\(^{23}\) Although the syntactic constraint on binding explains why English speakers do not treat the pronouns in (12) as co-referential (it would violate Principle B), it won't explain why the first pronoun is resolved to the man introduced in the first sentence, and not to Sam. The explanation can't be that only 'A man' and not 'Sam' updates the top of the stack to allow for a new dependent interpretation; a name can clearly also affect attention focus.\(^{24}\) So, why is the former, and not the latter, the highest-ranked candidate for 'He'? Similarly, given what we've said so far, (13) is puzzling. In (13i), the pronoun 'He' is resolved to John; in (13ii), to Tim. Clearly, both interpretations can't be top-ranked at the same time. But what accounts for this difference? What is it that makes John prominent in (13i) and Tim in (13ii)? We take up these challenges in turn.

For (12), an explanation is readily available: in English, the ranking of attention normally mirrors the grammatical role in which noun phrases are realized; in particular, the noun phrase in subject position takes precedence over the one in object position (Kameyama, 1996; Kehler, 2002; Bittner, 2014). This is why, with an utterance of (12), the first occurrence of the pronoun 'He' gets resolved to the candidate referent introduced by the previous subject, 'A man', rather than the one introduced by the previous object, 'Sam'. Crucially, this preference for referents introduced by the noun phrases in the subject position seems to be a grammatical feature of English, as it is not universally shared across languages.\(^{25}\) But what about

---

\(^{23}\) We are, of course, assuming that (13) is not uttered with accompanying demonstrative gestures. That the NP introduced in a subject position is preferred as an antecedent for the subsequent anaphor over the one introduced in the object position is well documented by a number of corpus and psycholinguistic studies. (See Kameyama 1999, 1986; Walker, Masayo, & Cote, 1994; Bittner, 2014.)

\(^{24}\) Clearly, names, and other definites can raise discourse referents to prominence (as witnessed by sentences like 'Bill came in. He sat down').

\(^{25}\) For example, languages with a more flexible syntax can exploit word order to indicate prominence. And some languages grammaticize prominence with
the occurrence of 'him'? The dependent interpretation linked to 'A man' remains the most prominent one when we turn to interpreting 'him'; the ranking remains constant throughout the second sentence of (12). But we already know why 'him' cannot be resolved this way. Principle B prevents 'him' from taking the same interpretation as the subject of its own clause. Thus, 'him' must resolve to the highest-ranked candidate lower in the ranking that does satisfy the relevant constraints, and that is the one for 'Sam'.

The most streamlined way of capturing this aspect of English grammar formally is to link grammatical roles to specified positions on the list of candidate interpretations: for example, the subject corresponds to position 0, the direct object to position 1, the indirect object to position 2. We then use expressions of the form '〈αn〉' and '〈πnc〉' to encode updates that push referents to position n in the stack (arbitrarily limiting ourselves to n = 0, 1 or 2 and assuming that the value of n follows compositionally from the grammatical status of the expression being represented). On this strategy, the first sentences of (1) and (12) are formalized as (14) and (15) respectively.

14 〈α0〉; [man(x0)]; [walk.in(x0)];
15 〈α0〉; [man(x0)]; 〈π1s〉; [met(x0, x1)]; [greeted(x0, x1)]

Given these representations, in formalizing successive sentences across discourse, we should also encode the fact that the pronoun is assigned a particular grammatical role in its own sentence, and represent this by an explicit further update to the attentional state of the discourse. (To represent the effect of Principle B in resolving the pronoun, we introduce '＠hex0' to represent the most prominent discourse referent that satisfies the pronoun other than the discourse referent x0—meaning the top-ranked male other than the subject.) We get (16) and (17):

16 〈α0〉; [man(x0)]; [walk.in(x0)]; 〈π0@he〉; [sit.down(x0)]
17 〈α0〉; [man(x0)]; 〈π1s〉; [met(x0, x1)];
     〈π0@he〉; 〈π1@hex0〉; [greeted(x0, x1)]

morphemes like topic markers that crosscut word order and grammatical role. (See Kameyama 1999, 1986; Walker, Masayo, & Cote, 1994; Bittner, 2014.)

Since the pronoun itself is in the subject position, the update—'〈π0@he〉'—that in this formalism corresponds to the pronoun, promotes the candidate referent associated with 'he' as the top-ranked referent. Note that it does not introduce a new referent; it rather re-stores (i.e., pushes a copy of) an old one to the designated position on the list of prominent candidates. If the entity was already at the top-ranked position, as is the case in both (16) and (17), the update associated with this pronoun will not change the possible interpretations of subsequent ones.
This approach achieves the most uniform possible syntax-semantics interface. In particular, the formal representation has the elegant property of consisting entirely of updates that can be read directly off the lexical items that comprise the utterance, along with their grammatical roles.

What about (13), our second challenging example? While (13i) could be explained in an exactly parallel fashion to (12), by appealing to subject preference for why the occurrence of 'he' takes John, and Principle B for why the occurrence of 'him' takes Tim, (13ii) would seem to run directly counter to this explanation, for in (13ii), even though there is a prominent, accessible antecedent in subject position ('John') satisfying the character of the pronoun, the preferred resolution of the occurrence of 'he' is to the referent introduced by the antecedent in the object position ('Tim'). How can we explain this? Here is where the coherence part of the Attention-Coherence Approach enters.

4 Coherence

Coherence theory starts from an obvious but often ignored observation about discourse: a discourse is more than a sequence of grammatical sentences. Successive contributions to discourse must be linked together by a recognizable flow of interpretive relationships; the speaker must signal how she articulates her complex ideas, stories and arguments step by step, according to shared standards. We see the requirement of coherence in both discourses (18) and (19) (Hobbs, 1979).

18 John took the train from Paris to Istanbul. He has family there.
19 John took the train from Paris to Istanbul. He likes spinach.

Discourse (18) doesn’t just reveal two interesting facts about John: it suggests that the reason he went to Istanbul was to visit his family there. A coherence relation of Explanation links the second sentence to the first (Cf. Asher & Lascarides, 2003; Kehler, 2002). Coherence theory recommends we represent this interpretive connection explicitly in order to capture the correct interpretation of (18). Conversely, the requirement that discourse must be coherent—that its implicit organization must be readily retrievable in virtue of the interlocutors’ general, linguistic and background knowledge—is strikingly evident in the interpretive effort that (19) elicits. Given the apparently unrelated facts about John in (19), we feel unsatisfied: we search for a connection. Is Istanbul known for its spinach? Is the train? Is eating spinach a symptom of some general timidity that also encompasses fear of flying? Clearly, interlocutors must use the common ground to disambiguate between discourses harboring different coherence relations, just as they must use the common ground to resolve other ambiguities. But just as clearly, a failure to
acknowledge any of these ingredients of interpretation constitutes a failure to understand the discourse.\textsuperscript{27}

Kehler (2002) advances the view that coherence relations cluster into (at least) three qualitatively different sorts, reflecting alternative strategies for organizing discourse.\textsuperscript{28} These are illustrated in (20)-(22).

20 Max spilt a bucket of water. He tripped on his shoelace.
21 Max spilt a bucket of water. He spilt it all over the rug.
22 Max spilt a bucket of water. John dropped a jar of cookies.\textsuperscript{29}

Discourse (20) illustrates the same kind of an explanatory discourse we’ve already considered in (18). For Kehler (2002), Explanation is an instance of a broader class of cause-effect (or event-result) relations that speakers can utilize to organize discourse. Discourse (21), meanwhile, gives an extended description of unfolding events—thus, a Narrative connection, which for Kehler epitomizes a broader class of Contiguity relations. And, finally, (22) exemplifies what Kehler calls Resemblance relations, organizing a discourse to draw comparisons and contrasts. In (22), there’s a Parallel between Max’s and John’s respective accidents. The different coherence relations are alike in signaling relationships among propositions in discourse. However, as we shall see, these coherence relations also shape how other material in discourse is interpreted. This is particularly important for context-dependent elements. The best way to capture these interpretive effects formally, we will argue, is to represent coherence relations explicitly in logical form.

\textsuperscript{27} It is worth noting that the idea of representing coherence relations explicitly doesn’t necessarily mean that an explanatory relation is a part of the truth-conditional or asserted content of the discourse in (18), although we choose to remain neutral one way or the other. In what follows, we will argue, however, that the connections signaled by coherence relations, as Explanation in (18), are not conveyed as a matter of merely implicatured, pragmatic content. This is not at odds with this content being non-asserted, or not at issue, in the sense of Tonhauser at al. (2013), since conventional encoding of non-asserted content is not unusual (e.g., semantic presuppositions, conventional implicatures, expressive content, etc.). For simplicity, our formalism includes only a single dimension of semantic content, in the sense of Potts (2005). That is, it does not attempt to model not-at-issue, or non-asserted content as a separate contribution to utterance meaning, but we could easily modify our framework to do this. However, this is a task for a different paper.

\textsuperscript{28} Kehler’s typology is useful in giving us an intuitive picture of the kinds of interpretive connections we find in discourse. Other typologies might be better suited for other purposes. (See e.g. Mann & Thompson, 1988; Knott, 1996.)

\textsuperscript{29} Examples (20) and (21) are adapted by Kehler (2002) from Asher and Lascarides (2003).
Coherence theorists view the two problems of identifying coherence relations and resolving semantic ambiguities as mutually constraining.\textsuperscript{30} In (13ii), for example, only resolving ‘he’ to Tim allows for a plausible explanation of John’s disappointment. In (18), only resolving ‘there’ to Istanbul rather than to Paris allows for a plausible explanation of John’s trip. In (20)–(22), meanwhile, we infer a temporal relation between the spill described initially and the tripping, spilling or dropping described next that matches the inferred coherence relation. Reference and coherence relations fit together in such cases.


23 Phil tickled Stanley, and Liz poked him.

Speakers tend to interpret (23) in either of two ways.\textsuperscript{31} One way assumes Liz’s action was prompted by Phil’s. Liz is perhaps reacting with disapproval to what Phil has done. This cause-effect interpretation comes with the understanding that ‘him’ must refer to Phil. The second reading assumes Liz’s action was similar to Phil’s in certain respects. This parallel interpretation comes with the understanding that ‘him’ must refer to Stanley. Crucially, the choice of a coherence relation (Result or Parallel) and the pronoun resolution (to Phil or to Stanley) go in tandem.\textsuperscript{32}

\textsuperscript{30} Importantly, though Coherence theorists acknowledge the mutual co-dependence of these problems, they take the preferences for certain pronoun resolutions, suggested by coherence relations, to be an input to a holistic, pragmatic process of interpretation that resolves the ambiguity by finding the overall most plausible interpretation. Unlike us, they do not hold there is a conventionalized effect of coherence on the attentional state, dictated by grammar; nor do they posit a formal representation that makes the two ambiguities—in coherence, and in pronoun resolution—align with each other as a matter of grammar.

\textsuperscript{31} Again, we assume no demonstrative gestures are accompanying (23). When (23) is embedded within a larger discourse, we might be able to get other resolutions of the occurrence of the pronoun, but this is because such embeddings could license a different coherence relation. Similarly, the effect of focus on the pronoun could license a different resolution, but this too is to be expected on our account, since focus can affect prominence just as demonstrative gestures do.

\textsuperscript{32} As noted, filling in the context might change the interpretive dependences that affect the pronoun resolution either by changing the coherence relations organizing the discourse, or by otherwise giving rise to other linguistic constraints on the interpretation. For example, one might imagine a scenario in which it is a part of the background knowledge that Liz always copies Phil, and tries to mimic whatever he
Pragmatic theories of reference resolution—even on standard coherence approaches—take this as evidence of an inferential relationship between a speaker’s intention in organizing the discourse and her referential intentions. According to us, however, these studies confirm an even tighter connection between coherence relations and pronoun resolution than Coherence theorists have been inclined to posit.

The contrast between these interpretations leads us to conclude it’s a mistake to treat (23) as harboring two separate ambiguities that audiences must resolve in turn—one about discourse coherence, another about pronoun resolution. These examples suggest that, once the choice of the coherence relation is established, there is not much further choice in the resolution of a pronoun; a certain interpretation is automatically set up as the most prominent one. Guided by this observation, we hypothesize that pronoun resolution itself is settled by the coherence relations that organize the discourse. Thus, coherence relations are another kind of mechanism that affects changes in the attentional state of a discourse. More precisely, we follow Coherence theory in formally representing coherence relations in the logical form of a discourse, but go further in proposing that this requires not only representing inferential connections but also representing the shifts in attention that are associated with coherence relations. To put it simply, coherence relations have a two-fold function—they establish an interpretive relation, and moreover, (as we shall argue) as a matter of their linguistic contribution, they can promote certain entities to prominence. Thus, we depart from standard Coherence theorists inasmuch as we maintain that coherence relations come with grammatically encoded shifts in attention, much like NPs or pointing gestures do. It is only after these shifts in attention are acknowledged as a part of the linguistic contribution of coherence relations that we can represent the intuitively correct interpretations of

does. So, then, one could say “Phil tickled Stanley, and as a result, she poked him”, where ‘him’ is then interpreted as Phil. Two things are important to note here. First, one should not assume that coherence relations are mutually exclusive. In many cases, more than one relation is needed to capture the structure of the discourse (see Asher and Lascarides, 2003). This might give raise to a more complex pattern of promoting referents to prominence. But, second, and perhaps more importantly, the mere presence of an explicit descriptive signal such as “as a result” does not automatically mean that a particular relation – Result, in the given case – is organizing the discourse (see Webber at al., 2003). Coherence relations are a matter of linguistic dependences and not just any way of describing how events are related in the real world will give rise to a particular coherence relation. Note that, similarly, not just any way of describing cause-effect relations will count as Explanation. What “as a result” in the given example is doing, instead of giving rise to the Result relation, is targeting the background presupposition about Liz’s known pattern of behavior. And there is reason already to think that presuppositions put further constraints on possible antecedents (see Hobbs, 1993). (Thanks to Alex Lascarides for pressing this point.)
pronouns, while giving pronouns the uniform and unambiguous meanings we have already proposed. We first explain how this account captures the data from (13i) and (13ii), and then, present several reasons in defense of the view that the attention-shifts associated with coherence relations are grammatically encoded, and are not merely pragmatically implied.

According to us, the difference between the resolutions in (13i) and (13ii) is fixed by the coherence relation that figures in their respective representations; these coherence relations come with different attention shifting updates, which affects, semantically, the resolution of subsequent pronouns. In (13i), the coherence relation is Narration: the content of the second sentence follows up on the content of the first, providing an extended description of unfolding events. In (13ii), the coherence relation is Explanation: the content of the second sentence is taken to explain the content of the first. According to Coherence theory, one or the other of these relations surfaces in the formal representations of (13i) and (13ii). Our proposal for pronoun resolution is that, as a matter of language, these distinct coherence relations affect the attentional state of the discourse, promoting one or the other of the candidate referents to the top-ranked position. When a formal representation features Narration, an attention-shifting operation is featured that raises John (the subject) to prominence; and when it features Explanation, an attention-shifting operation is featured that raises Tim (the object) to prominence. Accordingly, we represent (13i) as (24), and (13ii) as (25):

\[
\begin{align*}
24 & \langle \pi 0 j \rangle ; \langle \pi 1 t \rangle ; \left[\text{was.disappointed.with}(x_0, x_1)\right] ; \\
& \left[\text{Narration}(x_0)\right] ; \langle \pi 0 x_0 \rangle ; \\
& \langle \pi 0 @ \text{he} \rangle ; \langle \pi 1 @ \text{he}^{x_0} \rangle ; \left[\text{fired}(x_0, x_1)\right]
\end{align*}
\]

\[
\begin{align*}
25 & \langle \pi 0 j \rangle ; \langle \pi 1 t \rangle ; \left[\text{was.disappointed.with}(x_0, x_1)\right] ; \\
& \left[\text{Explanation}(x_0, x_1)\right] ; \langle \pi 0 x_1 \rangle ; \\
& \langle \pi 0 @ \text{he} \rangle ; \langle \alpha 1 \rangle ; \left[\text{work}(x_1)\right] ; \left[\text{sloppy}(x_1)\right] ; \left[\text{did}(x_0, x_1)\right]
\end{align*}
\]

(24) introduces John into the subject position, and Tim into the direct object position, requiring that John was disappointed by Tim. The second sentence continues a narrative about John; we represent the contribution of Narration as: ‘[Narration(x_0)]’; ⟨π0x0⟩, an update that requires that the discourse continues the narrative about x_0, and correspondingly promotes the prominence of x_0 in the

---

33 As we shall see in the next section, not all explanatory coherence relations raise the direct object to prominence, but that’s because not all of them are alike. Coherence relations are linguistic entities that encode linguistic interdependences in discourse, not mere common-sense dependences between events in the world. Thus, which entity is raised to prominence by which coherence relation is a matter of empirical linguistic inquiry, not a matter of an a priori judgment, or reasoning about the nature of common-sense relations between the events in the world.
attentional state of the discourse. The occurrence of the pronoun, 'He', now picks out the currently most prominent candidate appropriate for it, which is John, while the occurrence of 'him' picks out the currently most prominent candidate other than John, which is Tim. Further, it is required that the subject (John) fired the object (Tim).

The first line in (25) is the same as in (24). The crucial difference comes next. This discourse is modeled as harboring the coherence relation of Explanation, which we formalize using the update ‘[Explanation(x0, x1) ]; (π0x1 )’. This update requires that the two bits of discourse stand in an explanatory relation, and promotes the entity in the object position to prominence. The pronoun ‘he’ continues to pick out the top-ranked applicable interpretation, but thanks to the update associated with coherence, this is now Tim. The formula then proceeds with an additional dynamic existential quantifier, and further conditions that ensure that its witness is sloppy work done by the subject, Tim.

Crucially, for us, differences in pronoun resolution follow from specifications of the coherence relations that organize the discourse and update the attentional state, and not from an open-ended process of pronoun resolution. It might seem surprising that we are describing the attention-shifts associated with coherence as grammatically encoded. By contrast, you might be inclined to hold that the attention shifts contributed by coherence relations only reflects the speaker’s intentions and the hearer’s common-sense inference—that is, that they arise through pragmatic processes rather than through linguistic rules. This view is in fact standard among Coherence theorists. So construed, an attention shift that guarantees the correct interpretation of the subsequent pronoun is a result not of the linguistic effect of a coherence relation, as we are urging, but rather of pragmatic reasoning that occurs once a hearer has established that a particular coherence relation is structuring the discourse. For example, a hearer reasons that in (13i), since the particular relation of Narration holds between the two sentences, the speaker must be intending to promote a certain referent, in this case, John, to the center of attention. Crucially, on this view, intention recognition affects the adequate re-ranking of the list of prominent referents.

Note that this representation of Narration suffices to capture the effects on attention that concern us here. For other purposes, we might want to refine the representation of Narration to account for the spatio-temporal and causal links we find in narrative discourse. For example, Hobbs formalizes these inferential connections in terms of a relationship between the eventualities described in successive sentences. Asher and Lascarides (2003), meanwhile, model it as a relationship between the dynamic propositions expressed by successive sentences. Such extensions can be developed straightforwardly in our system by extending the formalism to a suitably enriched ontology. Inevitably, however, it incurs significant technical complexity to do so. Since it’s simply the effect of coherence on attention that concerns us here, we stick with the simpler formalism.
We can see why some might feel some sympathy for this alternative picture. It certainly makes sense for attention to shift in the ways that (24) and (25) suggest. Indeed, it would be perverse to reverse these preferences, so that when we came to consider a narrative about John who’s disappointed in Tim, the referent of ‘he’ was Tim, or when we came to explain what about Tim made John disappointed in him, suddenly the referent for ‘he’ was John. To organize discourse in such a confusing manner would make it much harder to communicate our ideas concisely. Taking the metaphor of attention seriously, such patterns would aim the spotlight of discourse directly away from the matters of greatest interest.

However, we invite you to ask not where coherence relations shift attention, but rather in what circumstances they do so. It is this question that gives us our reasons for maintaining that attention-shifting operations are grammatically encoded. We make our case by observing first that speakers and hearers take a rather restricted set of cues into account in instances of pronoun resolution. They privilege specifically linguistic cues, over the broader constraints of background knowledge and rational inference that they might potentially consider. For a perfect example of this regularity, consider (26), from Kehler (2002):

26 *Margaret Thatcher admires Ronald Reagan, and George W. Bush absolutely worships her.

Something is not right with (26); Kehler reports it is generally judged to be infelicitous by his subjects. What explains this observation? By virtue of following ‘admires’ with ‘absolutely worships’—a stronger term in an obvious scalar relationship—the speaker provides clear evidence that (26) is organized around a contrast between Margaret Thatcher’s and George W. Bush’s comparable attitudes. Coherence theory predicts that this parallel should make Reagan (the object of the first clause) the preferred referent for the occurrence of the pronoun in the object position in the second clause. And, indeed, reading (26), it seems as if the speaker has erred, inadvertently referring to Reagan as ‘her’. Of course, since Thatcher has been evoked in the previous sentence, in the subject position, and is a well-known object of Bush’s admiration, you’d expect it would be rather easy to refer to Thatcher with ‘her’ were the effect of coherence on prominence merely a by-product of the general pragmatic, common-sense reasoning that interpreters use to recognize a plausible interpretation; yet, this is not what we find.

This point is analogous to what we observed earlier about the conventionality of demonstration. Thatcher may attract our visual attention with what she’s doing in the current situation, but if the speaker is pointing elsewhere while saying ‘her’—or continuing an ongoing discourse about someone else—then we don’t take ‘her’ to

---

35 We note that Kehler, contrary to us, is not committed to this preference being a result of a grammaticized contribution of the coherence relation in question. He uses this example for different purposes, but it perfectly illustrates our point.
refer to Thatcher. Just so here, we suggest, the *Parallel* coherence relation encoded in (26) accomplishes a kind of inferred demonstration, indicating Reagan in a way that's difficult for any common-sense inference to override.36

Further support for our thesis that these characteristic cues for reference resolution are conventionalized comes once again from the variation we find across language communities. Many languages have explicit operations for shifting attention, such as grammatical topic marking or a distinction between topic and non-topic pronouns. It happens that some of these languages are much more constrained than English in the sorts of shifts they permit to take place implicitly. Our rough tentative inspection of Serbian, a pro(-noun) drop language, that allows (and sometimes requires) pronouns to be “dropped”, i.e. remain unpronounced in certain grammatical positions, seems to suggest that it is a language of that kind.37 For example, consider two possible translations of (13i):

27 Džon je bio razočaran Timom.
John-NOM is-PRS-3ms be-PPA-3ms disappointed-ADJ-3ms Tim-INS.
i. Otpustio ga je.
Fired-PPA-3ms him is-PRS-3ms.
ii. On ga je otpustio.
He him is-PRS-3ms fired-PPA-3ms.

When (13i) is translated into Serbian with the third-person male singular pronoun in (13i) "dropped" (as in (27i)), the natural interpretation is that it was John who fired Tim. Yet, with the pronounced third-person singular pronoun (27ii), the natural interpretation is that it was Tim who fired John. It seems that the overt pronoun signals a change in prominence that the covert counterpart does not require. English, lacking this explicit mean of signaling a shift in attention, is more flexible with implicit shifts in attention; thus, we witness an ambiguity in (13i)-(13ii).

36 Of course, putting a focal stress on 'her' makes the sentence felicitous, with the referent of ‘her’ Thatcher (as Kehler himself reports). But note also, that this focal stress comes with its own interpretive requirements. Stress seems to change what we take the point of the discourse to be: with stress, it explains how Bush follows Thatcher’s opinions, not how conservative politicians feel about Reagan. There seems to be no reading where the relation in question is *Parallel* and the pronoun is resolved to Thatcher. But this should be easily achievable if the attention-shifting effect of *Parallel* were merely a pragmatic, default effect.
The Greenlandic language Kalaallisut, as described by Bittner (2007), is another clear case of this. To illustrate, Bittner contrasts (28i) in English with (28ii) in Kalaallisut.

28  

Traveling to Denmark
i. When I came to Denmark, I bought my ticket six months in advance.
ii. Danmarkimut tikikkama, Danmarki-mut tikit-ga-ma, Denmark-sg.DAT come-FCT-1s
When I came to Denmark,... 
  qaammatit arvinillit siuqqullugu billitsisivunga.
  qaammat-t arvinili-t siuqqu-t lu-gu billtisi-si-pu-nga.
  month-pl six-pl v.ahead-ELA-3s ticket-get-IND.1V-1s
  I got a ticket (for some other event) six month ahead (of that event).

Following Moens and Steedman (1988) and Webber (Webber (1988), Webber et al (2003)), we take sentences with subordinate clauses, such as ‘when’ clauses, to be mini-discourses, with coherent interpretive connections between the clauses that, in many ways, mirror the interpretive connections found between successive sentences in a discourse, or between conjuncts in a conjunction. Crucially, for (27i), English speakers find a natural interpretation where buying the ticket early is a description that elaborates how the speaker came to Denmark. For its translation (27ii), however, the analogous interpretation is unavailable to Kalaallisut speakers. The Kalaallisut sentence requires an interpretation where the main clause describes what happens after the speaker came to Denmark – which is compatible with the narrative interpretation, but not elaboration. Why do we witness this discrepancy? What happens is that Kalaallisut grammatically marks the resolution of temporal anaphora, so that an event verb forces the temporal progression, making the time “right after the event described” prominent for temporal anaphora. Since such a temporal progression, dictated by an eventive verb, is consistent with the narrative interpretation, we find this interpretation available. However, though this language has grammatical means of signaling shifts in attention for temporal anaphora, it, unlike English, does not seem to associate coherence relations (in particular, Elaboration) with attention shifting updates. That is, in English, Elaboration contributes an attention shifting update that makes the time prior to coming to Denmark prominent for temporal anaphora. In Kalaallisut, Elaboration lacks this attention shifting update, and therefore, absent other grammatically marked ways of rendering the time before the trip to Denmark prominent, the only way to resolve temporal anaphora in (27ii) is for the trip to Denmark to proceed buying the ticket. This interpretation is only consistent with Narrative, not with Elaboration, and thus, there’s no Elaboration interpretation of (28ii).

Note that the discrepancy between English and Kalaallisut would be puzzling if the shifts induced by Elaboration in English were merely a result of pragmatic
reasoning. That is, if in English Elaboration makes certain interpretation prominent just by means of pragmatic reasoning, then we would expect to find the same range of interpretations in Kalaallisut. Buying tickets for trips is the same the world over; if the resolution of context-dependent meanings in (28i) and (28ii) were just common sense, we would expect the same interpretation in both cases. However, that is not what we find. Only if languages interpret analogous expressions differently as signals of transitions in discourse can we accommodate the difference we actually find. Our idea then is that each coherence relation in logical form carries with it a rule-governed contribution to attention in discourse. Then, given the state of attention, the resolution of pronouns follows.

Our view, of course, is consistent with general reasoning being crucial to interpretation. When the grammar delivers multiple candidate readings, hearers need to pick the one that makes the most sense on a given occasion. For example, (23) is ambiguous (at least) between the discourse containing the relation Result and one featuring the relation Parallel. Some general reasoning might be invoked in disambiguating between these two interpretations, much as some general reasoning might be involved in figuring out whether a speaker means the financial institution, or the river bank, with a use of 'bank', or which quantifier scope is intended with a use of 'Every boy kissed a girl'. In order to assign an interpretation, a hearer must first settle these disambiguations. In our view, this may involve assessing the plausibility of the inferential links conveyed by coherence relations. It may even involve evaluating whether the reading engenders a plausible resolution of pronouns. Such reasoning constitutes an important principle of disambiguation for us, but it cannot contribute its own content to logical form. It serves to privilege a particular logical form of the discourse, among the available ones the grammar delivers. But, once a certain coherence relation is established, its linguistic effect requires a re-ordering of the prominence ranking of candidate referents; that is, once a coherence relation is established, the resolution of a pronoun is determined as a matter of grammar, and not general reasoning.

In sum, there is strong evidence in favor of treating the attention-shifting effects of coherence relations as contributions governed by linguistic rules rather than as a byproduct of pragmatic reasoning. These attention-shifting updates change the attentional state of the discourse, thus setting the parameters of the context that determine the resolution of an occurrence of a pronoun. A pronoun, in turn, as a function of these parameters, automatically selects a referent, according to its linguistic meaning. Thus, not only is the resolution of a pronoun determined fully by its linguistic meaning as a function of context, but moreover, the relevant features of context that fix this reference are themselves orchestrated by the rules of language. So, pronominal resolution is fully linguistically determined—through and through. The linguistic mechanisms of attention-shifting updates and the rules governing pronominal usage together determine pronominal resolution on any occasion of use.
5 Conclusion

We have argued for a joint attention-coherence account for pronoun resolution. This account assigns a single linguistic meaning to each pronoun, but is nevertheless powerful enough to account for a wide range of data. The pronoun resolution responds to one overarching principle: pronouns are variables with dependent interpretations, but they are interpreted with respect to a ranking of prominence rather than a rigid method of names or indices. Moreover, each resolution is restricted by various additional constraints specified by the linguistic meaning of the pronoun, including person, gender and number features, and independently motivated syntactic constraints. They trigger a search for a matching interpretation. Think of the character of a pronoun as incorporating these constraints. The character constitutes the linguistic meaning of the pronoun and on our account it is transparently visible in logical form. More importantly, it determines the resolution of 'he' automatically, as function of context.

To make this work, the context must be appropriately set up. This is achieved by what we have called the attentional state of the discourse. This state is the result of a series of attention-shifting updates, which intuitively re-rank candidate resolutions for pronouns but are formalized as pushing new entities onto a stack of values for variables (using the Tarskian apparatus of assignment functions to model dependent interpretations). These updates are contributed by the diverse mechanisms that speakers have learned for structuring discourse and shaping the interpretation of pronouns, including evoking discourse entities in specific grammatical roles, demonstrating entities with non-verbal (yet linguistic) actions, and signaling the direction of discourse through various interpretive connections between clauses. Although the mechanisms contributing attention-shifting updates are heterogeneous, we have argued that this contribution is governed by linguistic rules, rather than by pragmatic reasoning, and as such, should be formally represented in the logical form for discourse.

Of course, our account is provisional in several respects. We haven't talked much about the modal profile of context-sensitive utterances. Dynamic semantics has the expressive resources to distinguish the dynamics of reference across discourse from the propositions that utterances express and the attitudes they can attribute to agents. Developing a formalism that combines these ingredients remains a task for future work. However, based on Bittner's (2014) successful foray into modal dynamic semantics with centering, we are optimistic that such formalisms can be

38 Our formal semantics does have resources to capture the basic facts about the modal profile of utterances containing pronouns. See the Formal Appendix for details.

39 See Brasoveanu (2010) and Cumming (2008) for detailed discussion of these issues.
straightforwardly adapted to take into account the context-dependent character of pronouns and the attention-shifting operations of discourse that we have proposed.

We also have remained silent about a great many of the ambiguities associated with discourses containing pronouns—both within English and across languages. Cases like Kehler’s (21) and (25)—among others—show that fleshing out our view in full detail will require us to postulate a variety of coherence-triggered attention shifts within complex sentences. Hobbs (1990), among others, has long argued that the effects of discourse coherence are visible not just when sentences are combined together, but in the many inferential connections that must be established to knit the interpretation of each clause together. So far, however, these relationships remain largely unexplored in formal semantics, and call for future work.

Relatedly, we haven’t committed to, or provided, any particular (exhaustive) list of coherence relations. The question of which coherence relations there are is an empirical one that requires careful further research to regiment the particular interpretations of different discourses within a consistent general framework. Our work suggests that patterns of pronoun interpretation can provide a powerful new source of constraints on this project. Conversely, our theory must confront the findings of this empirical research. Is the range of pronoun interpretations compatible with a plausible taxonomy of attention-shifting coherence relations? To see what’s at stake, consider the following, illustrative example (Winograd, 1972):

29

   The city council denied the demonstrators a permit.
   i. They feared violence.
   ii. They advocated violence.

Both (29i) and (29ii) exhibit an explanatory relation, yet the two occurrences of the pronoun ‘they’ are resolved differently. Our suggestion is that, though in both (29i) and (29ii) the content of the second sentence (or the event it describes) is taken to explain the content of the first one (or the event it describes), these explanatory relations are qualitatively different. In (29ii), the council’s decision about the demonstrators can be explained on the basis of the former’s beliefs about the latter: the relevant explanation being that it is because the demonstrators are potentially violent, or at least believed to be so by the city council, that the council has denied them a permit. Meanwhile, in (29i), the council’s decision about the demonstrators can be explained based on (other aspects of) the council’s attitudes: it is because the council feared violence that they decided to deny the demonstrators a permit. These qualitative differences suggest two distinct coherence relations are at play. So it’s no surprise that they might come with different attention-shifting operations—when the explanation goes via subject, the subject is promoted to prominence; when it goes via object, the pattern is reversed. To flesh out this suggestion, we need to substantiate systematic differences between these kinds of explanations. We don’t

---

40 For example, we do not offer a treatment of dependent clauses.
want to simply re-label the difference in pronoun resolution as a difference in coherence. At the same time, however, we are not concerned about the prospect of acknowledging a more nuanced array of relations than previous work sometimes appeals to. Coherence relations are a matter of the linguistic interdependences that we find in a discourse, and not a matter of simply characterizing real-world connections between events. But, we have already seen that explanatory relations fall under a broader class of cause-effect (or event-result) relations that comprise many different relations that epitomize this broader class in different ways (for example, we distinguish Result from Explanation). Thus, as already noted, only further empirical research can guide us closer to having a firm grasp on the taxonomy. We submit our proposal here to the scrutiny of this future research.

We ourselves remain most excited about the philosophical ramifications of the kinds of tools we have developed here. As we noted at the outset, philosophers often use linguistic examples to argue for context dependence—in cases like knowledge attributions, the implicit domain of quantifiers, standards for vague scalar terms, and many others—and then employ this context dependence to push broader philosophical views. The kind of context dependence they have in mind, implicitly, is generally a traditional one: they see context-dependent elements as freely selecting one interpretation from an open-ended array of candidates by unspecified, broadly pragmatic and open-ended mechanisms. Often, in fact, both the linguistic analyses and the philosophical arguments depend on this model of context dependence.

But it turns out—if we are right—that context is not as powerful as philosophers have traditionally presumed. In fact, as we have seen in interpreting even a simple and straightforward case like (1), much of our discourse ultimately lacks any dependency on non-linguistic context. And many of our practices in using language presuppose that our discourse has constrained and self-contained interpretations: when we judge truth and falsity, when we report one another, and when we talk about language and meaning itself. This observation suggests a project of extending our account to other cases of apparent interpretive variability—in domain restriction, the resolution of implicit arguments, and in lexical semantics. If our approach succeeds in capturing the apparent interpretive variability of such cases with uniform meanings and constrained variation, then philosophers will have to give up many of their customary appeals to context sensitivity.

6 Bibliography


Appendix: Formal definitions
for a fragment of the Attention-Coherence Approach

We have the following basic types in our language:

- individuals (type $e$)
- possible worlds (type $s$)
- truth values (type $t$)
- sequences of individuals and worlds (type $c$)

We reserve the symbol $a$ for the actual world.

We use the following notation to describe operations on sequences:

- $i_m$
  If $m$ is an integer, then $i_m$ is the $m$th element of $i$.

- $i_{m,n}$
  If $m$ and $n$ are integers, then $i_{m,n}$ is a sequence containing the subsequence of elements of $i$ in order from element number $m$ up through the element that precedes $n$ (if any).

- $i_{m...}$
  If $m$ is an integer, then $i_{m...}$ is the sequence containing the complete subsequence of elements of $i$ in order beginning from element number $m$.

- $i + j$
  If $i$ is a sequence and $j$ is a sequence, then $i + j$ is the sequence containing the elements of $i$ in order followed by the elements of $j$ in order.
  Note then that $i = i_{0,k} + i_{k...}$

- $u.i$
  If $u$ is an individual and $i$ is a sequence, then $u.i$ is the sequence that begins with $u$ and continues with the elements of $i$ in order.

- $w(i)$
  If $i$ is a sequence, then $w(i)$ is the first element of $i$ that is of world type.

We define frames and models in the usual way:
• **A Frame** is a tuple $F = (D_w, D_e, D_t)$ such that $D_t$ is a domain of truth values ($D_t = \{0, 1\}$), $D_w$ is a domain of possible worlds, $D_e$ is a domain of individuals, $D_t \cap D_w \cap D_e = \emptyset$, with $R$, a (transitive and reflexive) accessibility relation defined over $D_w$.

• **A Model** is a pair $M = (F, I)$, where $F$ is a frame and $I$ an interpretation function, which assigns to each individual constant an element of $D_e$ and each predicate constant a set of pairs of $D_w$, and an n-tuple of $D_e$.

The language of formulas and its semantics is as follows:

• Individual expressions

  – if $t$ is an individual constant, then $t$ is an individual expression
    (represents the name of an individual)
  
  – the variable $x_m$ is an individual expression
    (represents a discourse reference contributed by argument structure)
  
  – if $p$ is a unary predicate, then $@p$ is an individual expression
    (represents a syntactically unconstrained anaphor)
  
  – if $p$ is a unary predicate and $o$ is an individual expression, then $@p^o$ is an individual expression
    (represents a syntactically constrained anaphor)

The interpretation of individual expressions at a sequence $i$ and world $w$:

• $[[t]]^{i,w} = I(t)$ for interpretation function $I$.
  (Access constants from model.)

• $[[x_m]]^{i,w} = i_m$.
  (Look up values of variables. We need variables to manage argument structure; otherwise, it will be very cumbersome to deal with the syntax–semantics interface for transitive and ditransitive verbs; we need to potentially distinguish the order in which arguments are introduced, how salient they are after the utterance, and what role they play in the described event. Having variables clears this all up. Basically, $x_0$ will correspond to the subject, $x_1$ to the direct object, $x_2$ to the indirect object, and so forth.)

• $[[@p]]^{i,w} = i_0$ if $(w, i_0) \in I(p)$.
  $[[@p]]^{i,w} = [[@p]]^{i_1...w}$ otherwise.
  (Find most prominent referent that agrees with anaphor.)
\[ \llbracket \alpha k \rrbracket (w, i, j) \] if and only if \( j = i_0 + o.i_1 \ldots \) for some individual \( o \) of type \( e \).

\[ \llbracket (\pi k t) \rrbracket (w, i, j) \] if and only if \( o = \llbracket t \rrbracket i, w \) and \( j = i_0 + o.i_1 \ldots \) for some individual \( o \) of type \( e \).

\[ \llbracket [\varphi] \rrbracket (w, i, j) \] if and only if \( j = i \) and \( \llbracket [\varphi] \rrbracket i, w \) is true.

\[ \llbracket H; K \rrbracket (w, i, j) \] if and only if there is some sequence \( h \) such that \( \llbracket H \rrbracket (w, i, h) \) and \( \llbracket K \rrbracket (w, h, j) \).
• $$\Box K(w, i, j)$$ if and only if $$j = i$$ and for all worlds $$v$$ accessible from $$q$$, there is some $$k$$ such that $$[K](v, i, k)$$.

We represent the initial sequence $$J = (a, x, y)$$ where $$a$$ is the actual world, $$x$$ is the speaker of the utterance, $$y$$ is the addressee. In a more general language, $$J$$ could be extended by whatever parameters are appropriate for the interpretation of relevant indexical elements.

Finally, you have some general definitions:

• $$H$$ is true in a model $$M$$, when uttered at $$a$$, by $$x$$ to $$y$$ if and only if there is some sequence $$i$$ such that $$[H](a, J, i)$$.

• $$H$$ is valid if and only if it’s true in all models.

• $$H$$ entails $$K$$ (version 1: $$K$$ is a summary of $$H$$)
  For any model $$M$$, $$a$$ and individuals $$x$$ and $$y$$ in $$a$$, if $$H$$ is true when uttered at $$a$$ by $$x$$ to $$y$$ in $$M$$, then $$K$$ is true when uttered at $$a$$ by $$x$$ to $$y$$ in $$M$$.

• $$H$$ entails $$K$$ (version 2: $$K$$ doesn’t add information to $$H$$)
  For any model $$M$$, any world $$a$$ in $$M$$, and any assignment $$i$$ such that $$[H](a, J, i)$$, there is an assignment $$j$$ such that $$[K](a, J, j)$$.
  Version 2 is the one that’s usually given in treatments of dynamic semantics designed to accommodate anaphora, because it allows anaphoric links not only between the premises but from the premises to the conclusion.

Worked out examples:

• A man met Sam. He greeted him.

  formula  | gloss  | output
  --|--|--
  $$\langle \alpha \rangle ; \text{[man}(x_0)\text{]}$$ | “A man (is the subject)” | $$(m, \ldots)$$ where $$m$$ is a man
  $$\langle \pi \rangle s$$ | “Sam (is the object)” | $$(m, s, \ldots)$$
  $$\text{[met}(x_0, x_1)\text{]}$$ | “(the subject) met (the object)” | $$(m, s, \ldots)$$ where $$m$$ met $$s$$
  $$\langle \pi \rangle \text{[he]}$$ | “He (is the subject)” | $$(m, m, s, \ldots)$$ since $$m$$ is a he
  $$\langle \pi \rangle \text{[he]}$$ | “him (is the object)” | $$(m, s, m, s, \ldots)$$ since $$m, s$$ are he but $$m = x_0$$
  $$\text{[greeted}(x_0, x_1)\text{]}$$ | “(the subject) greeted (the object)” | $$(m, s, m, s, \ldots)$$ where $$m$$ greeted $$s$$

Recall that you should be reading “0” as “the subject” and “1” as “the direct object,” which is their meaning in the formalism.
John was disappointed with Tim. He fired him.

<table>
<thead>
<tr>
<th>formula</th>
<th>gloss</th>
<th>output</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \langle \pi 0 j \rangle )</td>
<td>“John (is the subject)”</td>
<td>( (j, \ldots) )</td>
</tr>
<tr>
<td>( \langle \pi 1 t \rangle )</td>
<td>“Tim (is the object)”</td>
<td>( (j, t, \ldots) )</td>
</tr>
<tr>
<td>([\text{disapp.with}(x_0, x_1)])</td>
<td>“(the subject) was disappointed with (the object)”</td>
<td>( (j, t, \ldots) ) where ( j ) was disappointed with ( t )</td>
</tr>
<tr>
<td>([\text{Narration}(x_0)]; \langle \pi 0 x_0 \rangle )</td>
<td>“continue a narrative about the subject”</td>
<td>( (j, j, t, \ldots) ) where ( j ) was disappointed by ( t ) since ( t ) is a he</td>
</tr>
<tr>
<td>( \langle \pi 0 @ \text{he} \rangle )</td>
<td>“He (is the subject)”</td>
<td>( (j, j, j, t, \ldots) ) since ( j ) is a he</td>
</tr>
<tr>
<td>( \langle \pi 1 @ \text{he}^{\text{m}} \rangle )</td>
<td>“him (is the object)”</td>
<td>( (j, t, j, j, t, \ldots) ) since ( j, t ) are he, but ( j = x_0 )</td>
</tr>
<tr>
<td>([\text{fired}(x_0, x_1)])</td>
<td>“(the subject) fired (the object)”</td>
<td>( (j, j, j, t, \ldots) ) where ( j ) fired ( t )</td>
</tr>
</tbody>
</table>

John was disappointed with Tim. He did sloppy work.

<table>
<thead>
<tr>
<th>formula</th>
<th>gloss</th>
<th>output</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \langle \pi 0 j \rangle )</td>
<td>“John (is the subject)”</td>
<td>( (j, \ldots) )</td>
</tr>
<tr>
<td>( \langle \pi 1 t \rangle )</td>
<td>“Tim (is the object)”</td>
<td>( (j, t, \ldots) )</td>
</tr>
<tr>
<td>([\text{disapp.with}(x_0, x_1)])</td>
<td>“(the subject) was disappointed with (the object)”</td>
<td>( (j, t, \ldots) ) where ( j ) was disappointed with ( t )</td>
</tr>
<tr>
<td>([\text{Explanation}(x_0, x_1)]; \langle \pi 0 x_1 \rangle )</td>
<td>“explain appealing to the object”</td>
<td>( (t, j, j, t, \ldots) ) where ( j ) was disappointed by ( t )</td>
</tr>
<tr>
<td>( \langle \pi 0 @ \text{he} \rangle )</td>
<td>“He (is the subject)”</td>
<td>( (t, j, j, t, \ldots) ) since ( t ) is a he</td>
</tr>
<tr>
<td>( \langle \alpha 1 ]; [\text{work}(x_1)] \rangle )</td>
<td>“A work (is the object)”</td>
<td>( (t, w, j, t, \ldots) ) where ( w ) is some work</td>
</tr>
<tr>
<td>([\text{sloppy}(x_1)])</td>
<td>(the object) is sloppy</td>
<td>( (t, w, j, t, \ldots) ) where ( w ) is sloppy</td>
</tr>
<tr>
<td>([\text{did}(x_0, x_1)])</td>
<td>“(the subject) did (the object)”</td>
<td>( (t, w, j, t, \ldots) ) where ( j ) did ( t )</td>
</tr>
</tbody>
</table>

We can reduce the difference in the minimal pair to an attentional shift associated with two different coherence relations, Narration and Explanation, represented explicitly in logical form.