Discourse and Logical Form:

Pronouns, Attention and Coherence

1   Introduction

The meaning of the pronoun 'he' does not seem ambiguous in the way that, say, the noun 'bank' is. Yet what we express in uttering 'He is happy,' pointing at Bill, differs from what we express when uttering it, pointing at a different person, Sam. 'he' in the first case refers to Bill, in the second to Sam. And further with an utterance of 'A man always drives the car he owns,' or 'A man came in. He sat down,' without an accompanying pointing gesture, 'he' refers not at all; its interpretation, instead, co-varies with possible instances in the range of the quantifier 'A man'. Those uses in which the pronoun is interpreted referentially are usually called 'demonstrative,' and those in which it is interpreted non-referentially are usually called 'bound'. We immediately face two questions: what is the semantic value of a pronoun in a context, given that it can have both bound and demonstrative uses? And, since this value can vary with context, how does context fix it? The first question concerns the semantics of pronouns, and the second their meta-semantics.

The received view answers the first question by positing an ambiguity: bound uses of 'he' function as bound variables; demonstrative ones, by contrast, are referential. The received view's answer to the second question is that the semantic value of a bound use co-varies with possible instances in the range of a binding expression, and for demonstrative uses, linguistic meaning constrains, but does not fully determine, its semantic value in context. For instance, in uttering 'He is happy,' pointing at Bill, its linguistic meaning is supplemented by a pointing gesture in fixing the referent of 'he', and, in general, a surrounding contextual cue, either an accompanying demonstration, or an appropriate and adequately transparent speaker intention, or even both, is required to fix the referent of a demonstrative pronoun.[1] Such supplementation is understood to be extra-linguistic or pragmatic, and, most certainly, not linguistically encoded.

This paper challenges the received view on both counts: we reject the ambiguity thesis and any meta-semantics that invokes extra-linguistic determinants of semantic values in context. We argue that the semantic value of a demonstrative

---

[1] Proponents of the received view are many. Notable ones are Kaplan (1989., b), King (2014a, b), King (2014a, b), Neale (2004), Reimer (1992). There are important differences between these different versions, but they do not matter for our present purposes. We believe these views are misguided, but our focus is primarily here on our positive view. For our criticisms of such views, see [Redacted].

pronoun, like that of a bound variable, depends on, and can co-vary with, antecedent expressions; and further, the linguistic meaning of a pronoun, on both uses, fully determines its semantic value in a context. In particular, 'he' is like the first person pronoun 'I' since for both context automatically fixes a semantic value on any occasion of use. Just as 'I,' given its character, refers to the speaker of a context, so too, contra the received view, 'he', given its character, determines its semantic value as a function of a fixed feature of context. On first pass, its semantic value is the male at the "center of attention" in a context. (We will precisify "center of attention" below.) In contrast to the received view, what lies at the center of attention is *not* governed by extra-linguistic mechanisms, speaker intentions or demonstrations, but entirely by linguistic rules. More precisely, linguistic mechanisms govern the dynamics of prominence in a discourse, rendering certain semantic values preferred for pronoun resolution in a context. The meaning of a pronoun fully determines its referent in a context—it picks out whatever linguistic rules determine to be "at the center of attention" at that point in a discourse. In what follows, we spell out these rules and how they affect pronoun resolution.

As a clarification, note that our claim that linguistic conventions play a role in pronoun resolution is not altogether novel. Most theories of pronoun resolution maintain linguistic conventions play *some* role in fixing semantic value. Many believe conventions play a role in constraining what speakers can even be taken to intend—e.g., in a normal context, one cannot reasonably intend to refer to a woman using 'he'. We believe, however, such constraints undermine the main ingredient of these accounts, namely, that speaker intentions, or additional epistemic cues, play a crucial role in fixing the semantic value of a pronoun. As we have argued elsewhere,[2] the meaning of 'he' is determined in conformity with linguistic conventions *regardless* of the presence or absence of a speaker intention, or an epistemic cue.

Briefly, to see what we mean, suppose while Sue is making herself highly salient by jumping up and down, you have been asked, 'Who came to the party last night,' to which you respond, 'Mary came to the party. She had fun.' In uttering 'she had fun', you make no overt gesture towards Sue, not even a glance towards her, but nevertheless you intended to refer to Sue. It is our intuition that what you have said is *Mary* came to the party and had fun, despite your intention or Sue's salience. In short, contrary to the received view, an intention is neither necessary nor sufficient to fix a referent on an occasion of use. To say in such cases that linguistic conventions constrain what can reasonably be intended is to agree with us that these conventions fix a referent on an occasion of use.

We also note that various authors argue that linguistic rules do play a role in resolving bound uses of pronouns, and more generally, so-called anaphoric uses,

---

[2] We have provided extensive criticism of such accounts elsewhere. [Reference redacted].

where pronoun resolution is dependent on a prior linguistic expression.[3] Our view goes beyond, as well as diverges from, these authors in several respects.[4] For one, we argue for a unique linguistic meaning (a character) for *all* uses of pronouns—demonstrative, bound and anaphoric alike—and that, for *all* these uses, the interpretation of a pronoun is fully determined by linguistic rules. Secondly, we elucidate a wide range of linguistic mechanisms that affect the dynamics of prominence in a discourse, which have either gone unnoticed in the current debate, or otherwise are confused with pragmatic effects on interpretation. Given the mechanisms that maintain the dynamics of prominence in a discourse, any occurrence of pronoun picks out its semantic value automatically, according to its character—selecting whatever is most prominent, or "at the center of attention".

Finally, we want to register that our view has far-reaching philosophical consequences beyond the semantics and meta-semantics of pronouns. Philosophers regularly invoke 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 are invariably guided by the received view of the meta-semantics of pronouns: namely, that the values of these parameters are determined jointly by linguistic meaning and non-linguistic cues.[5] Often, in fact, both the linguistic

---

[3] Our distinction between bound and demonstrative uses of pronouns does not perfectly coincide with one between demonstrative and anaphoric uses of pronouns, where an anaphoric use is any sue the interpretation of which depends on an expression introduced earlier in the discourse. Namely, some anaphoric pronouns are clearly referential. To wit, consider: 'Bill came in. He sat down.' Since we will argue that all uses of pronouns have the same semantics, and are resolved by appeal to the same linguistic mechanism, all demonstrative, bound and anaphoric uses are, according to us, alike.

[4] Theorists who stress linguistic constraints on the interpretation of an anaphoric pronoun to various degrees include Chomsky's *Binding Theory* (*cf.* Chomsky (1981) as well as various forms of dynamic semantics (e.g., Kamp (1981) and Heim (1982)). Almost everyone agrees there are linguistic constraints on the resolution of anaphoric pronouns. Fiengo and May (1994, 2006), for instance, hold that any two co-indexed expression-tokens are tokens of the same expression type, and are co-referential as a matter of grammar, and point to a number of linguistic constraints on co-indexed, including, but not limited to, those of Binding Theory. Though we agree with these authors that anaphoric dependences are a matter of lingustic conventions, we both go beyond, and deverge from them, in several respects. First, we argue *all* pronoun resolution—whether demonstrative, bound or anaphoric—is determined by a lingusitc rules. Secondly, the theories described allow for a single pronoun to be representd by different expressions in the logical form (corrseponding to different indices). This suggests an ambiguity in the word 'he' like, e.g., the word 'bank'. We, by contrast, argue that 'he', has a unique, unambiguous linguistic meaning, and that all its occurences are represneted by a token of the same expression-type in logical form. Thirdly, and most importantly, we charcterize a wider specturm of hithero unappreciated linguistic mechanisms that govern pronoun resolution.

[5] 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

analyses of these terms, and the philosophical arguments exploiting these analyses, depend on this model of context dependence. If we are right, however, the standard view is misguided even for pronouns. This suggests we should be wary of meta-semantical theories about philosophically interesting expressions built on this model. In fact, our view recommends a project of its extension to other cases; if so, philosophers will have to rethink the role of context-sensitivity in their arguments.[6]

In what follows, we develop an approach to the semantics and meta-semantics of pronouns we call the *Attention-Coherence Approach*; it contains two components—a ranking of candidate interpretations according to relative prominence (attention), and implicit mechanisms affecting this ranking (coherence). The role of both components 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, (*character*, if you like), always pick out whatever is 'in the center of attention' in a discourse. To capture this, we borrow resources from *Centering* theory,[7] 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. Call this ranking the *attentional state* of an ongoing discourse.[8] As a discourse progresses, utterances bring new candidates into focus and adjust the prominence of old ones, thereby changing the attentional state. We shall argue that pronouns, as a matter of meaning, always pick out the top-ranked candidate in the current attentional state, and further, that the attentional state itself is maintained through linguistic rules.

To motivate the dynamics of an attentional state, we begin with (1):

1   A man walked in. He sat down.

One reading is that 'He' is interpreted as co-varying with 'A man'. On this construal, (1) is all true just in case some man walked in and sat down. On a second reading, the speaker, perhaps uttering 'He' pointing at a small child, Bill, is interpreted as referring to Bill. The discourse is true just in case a man came in and *Bill* sat down.

Speakers can use (1) in different ways and hearers can point to reasons why certain interpretations are recovered, and others not. Pronoun resolution is guided by an implicit organization that knits together information in discourse. On its anaphoric reading, the discourse begins with a description involving 'A man' and proceeds directly to develop a narrative: accordingly, 'He' is interpreted as dependent on 'A man'. On its deictic reading, however, 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'. We shall argue that the difference between these interpretations of (1) is a difference in the dynamics of attention, guided by linguistic mechanisms. It is because the description 'A man' makes its potential witnesses prominent, and the discourse proceeds to develop a narrative about whomever is a witness for 'A man', thus keeping it in the center of attention, that 'He' is understood as co-varying with potential witnesses in the domain of the existential 'A man'. Similarly, it is because the pointing gesture shifts the attention from potential witnesses for 'A man' to the child pointed at, Billy, that 'He' on the deictic reading is understood to refer to Billy. These changes in attention, governing the prominence of potential resolutions for a pronoun, we shall argue, are guided by linguistic mechanisms.

To make our proposal clear and precise, we shall endorse a particular way of formalizing our key ideas. But our philosophical point—that pronominal interpretation is fixed through linguistic rules—is independent of our formalization, and could be implemented in other frameworks. Accordingly, our preferred formalization is not essential to our argument.[9]

Since on one interpretation of (1), 'He' is understood as 'bound' by the indefinite 'A man', we adopt a model for bound variables that allows us to capture relevant interpretive dependencies across discourses. One such model cleaves closely to the familiar Tarskian machinery for variable binding, according to which the truth of a formula is specified relative to an assignment of values to variables. A variable, x_i, is interpreted by retrieving the ith element of a current assignment. Bound variables have dependent interpretations because quantifiers vary the assignments in force within their scope. One shortcoming is that interpretation so-construed is limited to the syntactic scope of a quantifier. But quantifiers can introduce dependencies that

---

[9] For alternative frameworks, see e.g., Elbourne (2005). See also Fiengo and May (1994, 2006).

persist across subsequent sentences, as in the anaphoric reading of (1). We need a framework for extending the familiar notions of a variable and binding to the kind of dependent interpretations in (1). A fix is provided by *dynamic semantics.*[10]

In dynamic semantics, just as in a Tarskian framework, variable binding creates dependent interpretations by varying assignment functions. However, unlike the Tarskian approach, dynamic semantics assigns truth conditions to formulas relative to a *pair* of assignment functions, not to a single one. The first assignment accounts for interpretive dependencies potentially already available in the discourse for interpreting a formula; the second, however, accounts for interpretive dependencies the formula itself makes available for interpreting subsequent discourse. A formula is interpreted as an instruction to update the available interpretive dependencies: it augments the potential interpretive dependencies from prior discourse, encoded in the input assignment, with the additional interpretive dependencies a formula introduces for interpreting subsequent discourse, so as to create the output assignment. Just as in Tarskian semantics, in dynamic semantics, we interpret a bound variable x_i by retrieving the ith 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 an output assignment makes it true.[11]

Due to the existential quantifier in (1), the formula modeling the first sentence in (1) is true on a given input assignment just in case there is an output assignment according to which some unspecified man—a witness for the existential—walked in.[12] The second sentence is modeled with a formula that takes this output assignment as a new input; 'he' can then be interpreted as a bound variable that takes as its value whatever the witness for 'a man' is in this new input assignment, resulting in the anaphoric reading of (1). The discourse is true on the initial input assignment just in case a man walked in and sat down.

---

[10] See Dekker (2011) for an accessible overview of dynamic semantics.

[11] In this framework, we interpret formulas via updates that relate an input assignment g to an output one 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 one H ; K that performs H followed by K. An update is true for an input assignment iff it 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. As noted earlier, we exploit this particular formal system just for the sake of concreteness and clarity; our philosophical ideas can be implemented in other frameworks. For detailed exposition of the formal system see the Appendix.

[12] Any output assignment 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.

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

By contrast, since we want to model prominence in a discourse, we will organize interpretive dependencies *not* with names, but by *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 by one position. In this way, they vary output assignment functions. We treat a variable not as a temporary name, but as a marker of prominence; 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. We use '@' as a variable, interpreted relative to an assignment *g,* as specifying the top-ranked element of *g.* (The mnemonic is '@' is *at* the center of attention.)

The meaning of '@' is a first approximation to the meaning of an English pronoun. To illustrate, let '[man(@)]' be a condition that requires that whatever is at the center of attention be a man. And let '$\langle \alpha \rangle$' indicate a dynamic existential quantifier, which changes the input assignment by introducing a new unspecified top–ranked individual at the center of attention, with all other candidates demoted one position in the ordering.[13] This update is a first approximation to the meaning of the English indefinite article. It affects the prominence ranking, changing what is at the center of attention. (1), on its 'anaphoric' interpretation, is represented as (2):

2 $\qquad$ $\langle \alpha \rangle$; [man(@)]; [walk.in(@)]; [sit.down(@)]

(2) begins with an existential quantifier, corresponding to the indefinite article, that introduces its witness as the new top-ranked resolution for subsequent variables;

---

[13] This is a non-deterministic update in the jargon of dynamic semantics. $\langle \alpha \rangle$ 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, $h_i = g_{i-1}$.

the witness is constrained to be a man, to have walked in, and—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.[14]

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

3   A man walked in. He [pointing to a child, Bill] sat down.

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

We introduce a family of updates, written '$\langle$ πc $\rangle$', where 'π' corresponds to the act of pointing and 'c' to some individual being pointed at.[15] This update stores c as the top-ranked entity, and (as always) pushes all others down a position in the ordering. It thereby represents the effect of the pointing that accompanies the use of 'he'. (3) is represented formally as (4):

4          $\langle$ α $\rangle$; [man(@)]; [walk.in(@)]; $\langle$ πb $\rangle$; [sit.down(@)]

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. The intervening expression '$\langle$ πb $\rangle$', corresponding to the act of pointing at Bill, updates the attentional state. As a result, whomever the speaker has pointed at is at the top of the assignment when '[sit.down(@)]' is interpreted. In other words, a pointing gesture updates the prominence ranking, making Bill the center of attention.

As in (4), the shift of attention provoked by a pointing gesture is reflected in logical form. However, traditionally, the effect of a pointing gesture accompanying the use of a demonstrative pronoun has been treated *as the 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 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 in the indicated direction. Crucially, the pointing gesture does *not* render the demonstrated object prominent

---

[14] The formula is true for an input assignment *g* iff it is true for some output assignment. Since (2) has no free occurrences of a variable, its truth does not depend on an initial input assignment at all.
[15] More generally, '$\langle$ α $\rangle$' corresponds to the contribution of indefinites, and '$\langle$ πc $\rangle$' to definites, where 'c' is a denoting expression. We do not offer here an account of definites other than pronouns.

as a matter of a convention. It serves as an epistemic cue that helps manifest the speaker's intention because it makes sense to point at an intended referent.

We reject this view. The change of attention contributed by the pointing gesture in (3) is and *should be* reflected in logical form, we maintain, because the sorts of gestures that secure pronoun resolution in these cases do so as a matter of grammar. Our claim is that rules of language dictate that a demonstrative act of pointing introduces a new candidate referent at the top of the stack.[16] 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 rules sensitive to form, meaning, and the relationship with ongoing speech. English speakers seem to count deixis as well-formed only when the pointing action is synchronized appropriately with the prosody of the accompanying utterance. They often repair utterances when their performance fails to align speech and gesture in time, as one would expect if the requirement for synchronicity were dictated by an underlying convention.[17]

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. 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 many possible gestures that are *not* typically used as deictic gestures by English speakers though they could have been. 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 nothing in the gesture itself precludes it from playing this function. The arbitrary fact that some spatially directed actions are taken to indicate objects—while analogous ones are not—reveals convention at work in English speakers' demonstrations. Numerous languages, unlike English, allow speakers to indicate objects by deictic gestures of their lips, and not an extended index finger. Cuna, a language from

---

[16] For a grammar integrating gesture and speech, see Alahverdzhieva and Lascarides (2011).

[17] This is similar to how prosodic focus is grammatically constrained to appear at a particular position with a particular contour to contribute a semantic contribution.

Panama, is one of many examples.[18] Such conventionality and cross-linguistic variability is a hallmark of linguistic meaning.[19]

Further, in typical cases, acts signaling shifts in attention are *indispensable* for the appropriate interpretation. In (3), 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. Even with Bill jumping up and down in front of the interlocutors, attracting their psychological attention, unless an overt signal establishes Bill as the referent of 'he', the audience will prefer the default anaphoric reading. Indeed, some means of explicit signaling of a shift of attention is required even when a linguistic antecedent is *un*available, such as when a pronoun is used deictically in an utterance that initiates a conversation. If deictic gestures were epistemic cues that suggest real-world salience, this would be quite mysterious.[20]

Moreover, the linguistic contribution of a pointing gesture clearly affects the inference patterns (1) licenses. The inference from (1), on its anaphoric interpretation, where 'He' is understood as co-varying with 'A man,' to 'Some man sat down' is valid; but, on its deictic reading, it is not. Given that, as we have argued, the contribution of a pointing gesture is underwritten by linguistic conventions, and that this contribution affects which inference patterns (1) can license, we submit that the contribution of the pointing gesture should be reflected in the logical form of (1). Indeed, (2) entails that some man came in and sat down, as a matter of logical form; but (4), on the other hand, does not.[21]

Finally, we deny that pragmatic mechanisms determine the *demonstratum* of a demonstrative gesture. We represent the pointing in (3) as a *pointing at Bill*, not as pointing *simpliciter*, *with an intended demonstratum to be determined by context*. In other words, we deny pointing is semantically interpreted as having a uniform character that, given a context, determines a referent. Pointing at Bill is not interpreted as *pointing at x*, where *x* is contextually determined to be Bill. Rather,

---

[18] See Wilkins (2003).

[19] We understand convention a la Lewis (1969), though we need not subscribe to all of his conditions to register our point.

[20] Grammar specifies a diverse set of resources for raising entities to prominence (e.g., deictic gestures and indefinite NPs). Our formalism is expressive enough to make such resources available even for utterances that are not following up an ongoing discourse. Our prediction 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. We provide a more detailed account of deixis without overt accompanying demonstrative gestures in [Redacted].

[21] An inference from a formula K to H is valid iff for any assignments g, and h, if K relates g to h, then there is an assignment i such that H relates h to i. Formal definitions are in the Appendix.

pointing gestures are ambiguous between multiple possible forms, e.g., *pointing at Bill*, *pointing at his shirt*, *pointing at his button*. Acts of pointing are, on our account, ambiguous. The interpretive work in (3) comes in settling whether the *form* of the gesture is pointing at Bill or pointing at something else. But *disambiguation* is pre-semantics, in Kaplan's sense:[22] it involves the interpretive work needed to settle *the linguistic form* of an utterance. This should not be confused with what is involved in the semantic interpretation of a given linguistic form. Though, according to us, epistemic cues play no role in semantic interpretation, they *can* allow observers to recognize *which form* of a pointing gesture is in play. This is the role they'd play in helping to disambiguate *any* ambiguity, e.g., in disambiguating a use of 'bank'. And, as with other ambiguities, conventions governing demonstrative actions constrain 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: we have argued there are reasons to treat the effect of an attention shift triggered by an indefinite NP or an act of demonstration as rule governed and as represented in logical form. The recognition of these systematic effects on the prominence ranking of candidate referents in a context allows us to assign one 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 obviously required; we turn to them directly.

3 Refinements

That 'he', as a matter of meaning, takes on the most prominent candidate interpretation cannot be the whole story. 'He' in (5) cannot be interpreted as co-varying with 'A girl'. It must be resolved to a *male*.

5   A girl came in. He sat down.

'She' in (6) cannot be interpreted as co-varying with the witness for 'Some girls'. 'she' must have a *singular* interpretation.

6   Some girls came in. She sat down.

---

[22] See Kaplan (1989b.).

'we' and 'they' in (7) cannot be interpreted as co-varying since 'they' requires a third-person interpretation *disjoint* from speaker and addressee, and 'we' requires a first-person pronoun interpretation that includes the speaker.[23]

7   We came in. They sat down.

Pronouns come with person, gender and number requirements that must be satisfied in pronoun resolution. These requirements are *linguistic*: it matters whether a speaker utters 'he', 'she' or 'they' because they differ in meaning.[24]

Further, in (8), 'him' cannot semantically co-vary with Paul, even though Paul would normally be an eligible referent for 'him'.

8   Paul met him.

The explanation is that reference is constrained by syntactic principles—in this case, Principle B of Chomsky's Binding Theory, which requires non-reflexive pronouns to be free in their governing category. Roughly, this means 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.[25] So, in (8), where 'him' and 'Paul' are clausemates, 'him' cannot refer to Paul on pain of violating Principle B.

With this in mind, we say a pronoun denotes the value stored at the highest-ranked position on the stack that respects grammatical features and applicable syntactic principles. To interpret a pronoun, we consider candidates in order of prominence within the current attentional state of the discourse, until we find one that satisfies

---

[23] Sometimes a speaker can use 'he' to refer to himself or 'they' to pick out a group including himself. Witness a Perry-type scenario, where the speaker seeing, but not recognizing, himself in a mirror says, "He is the messy shopper" (Perry, 1979). Such cases are arguably marked in the sense that the exploitation of an extant rule makes for the surprising effect. Others suggest that the disjointedness condition is pragmatically implicated, not grammatically dictated (see Fiengo and May (2006)). By contrast, we treat number requirement as a part of the character of a pronoun. We agree that sometimes the speaker can manage to 'speaker-refer', in the sense of Kripke (1977), to himself with 'he', or to a group including himself with 'they'. Also, sometimes an audience can understand that the speaker meant to pick out himself with 'he', or some group including himself with 'they', in the same sense as we can sometimes understand a speaker who mistakenly attempted to pick out a male with 'her'. This is so even though the character of a pronoun encodes number, gender and person features. However, we could modify our account, if one insisted on excluding these features from the meaning of a pronoun. Whichever property P specifies the character of 'he', we maintain its referent is the top-ranked entity satisfying that property.

[24] Again, this is true regardless of whether a speaker can manage to 'speaker-refer' in the sense of Kripke, to a woman with the pronoun 'he' (or a definite 'that man').

[25] More precisely, the antecedent of a non-reflexive anaphoric pronoun must not be local, or c-command the pronoun. For precise definitions, see Chomsky (1981).

the operative linguistic constraints.[26] (9) integrates the constraints of semantics and syntax for the pronoun 'he'.

9   When interpreted relative to an assignment 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.[27]

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

10   ⟨ α ⟩; [man(@)]; [walk.in(@)]; [sit.down(@he)]

11   ⟨ α ⟩; [man(@)]; [walk.in(@)]; ⟨ πb ⟩; [sit.down(@he)]

However, even with these qualifications, (12) and (13) point to 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 'He' to the man introduced in the first sentence and 'him' to Sam.[28] Although the syntactic constraint on binding explains why English speakers do not treat the pronouns in (12) as co-referential (violation of 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 also affect attention focus.[29] So, why is the former, and not the latter, the highest-ranked candidate for 'He'? Similarly, given what we've

---

[26] The proposal to resolve the pronoun to the top-ranked element on the stack involves *search* rather than a simultaneous imposition of potentially competing constraints*;* the pronoun picks out the value stored at the highest ranked position on the stack *that satisfies the constraints*.

[27] Roberts (2002) proposes similar rules for demonstrative pronouns, but with crucial differences. Instead of a linguistically governed notion of attentional prominence ranking, she uses a liberal notion of contextual, "real-world" salience. Moreover, she treats semantic constraints on pronoun resolution as *presuppositions* of a pronoun, but we understand them as part of character.

[28] Assume (13) is uttered without accompanying demonstrative gestures. That the NP in subject position is preferred as an antecedent for the subsequent anaphor over the one in object position is well-documented by a number of corpus and psycholinguistic studies. See Kameyama (1999), Walker, Masayo, and Cote (1994), and Bittner (2014).

[29] Witness: 'Bill came in. He sat down'.

said so far, (13) is puzzling. In (13i), 'He' is naturally resolved to John, but in (13ii), to Tim. Clearly, both referents can't be top-ranked simultaneously. What makes John prominent in (13i) and Tim in (13ii)? We take up these challenges in turn.

For (12), an explanation is available. In English, attention ranking mirrors the grammatical role in which noun phrases are realized; the noun phrase in subject position takes precedence over the one in object position.[30] This is why, with (12), 'He' is resolved to the candidate referent introduced by the previous subject, 'A man', rather than one introduced by the previous object 'Sam'. Crucially, this preference for referents introduced by noun phrases in subject position is a grammatical feature of English, and not universally shared across languages.[31] But what about 'him'? The dependent interpretation linked to 'A man' remains most prominent when we turn to interpreting 'him'; the ranking remains constant throughout the second sentence of (12). But because 'him' cannot be so resolved, on pain of violating Principle B, it must resolve to the highest-ranked candidate lower in the ranking that satisfies the relevant constraints, i.e., Sam.

A streamlined way of representing this aspect of English formally is to link grammatical roles to specific positions on the stack of candidate interpretations: the subject corresponds to position 0, the direct object to 1, the indirect object to 2. We use expressions of forms '$\langle \alpha n \rangle$' and '$\langle \pi nc \rangle$' to encode updates that push referents to position n in the stack (arbitrarily limiting ourselves to n = 0, 1 or 2 and assuming the value of n follows compositionally from the grammatical status of the expression being represented), pushing all other one position down. On this strategy, the first sentences of (1) and (12) are formalized as (14) and (15) respectively.

14          $\langle \alpha\ 0 \rangle$; [man($x_0$)]; [walk.in($x_0$)];

15          $\langle \alpha\ 0 \rangle$; [man($x_0$)]; $\langle \pi 1s \rangle$; [met(x_0, x_1)]; [greeted(x_0, x_1)]

Given these LFs, in formalizing successive sentences across discourse, we should encode the pronoun as assigned a 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 '@he^{x_0}', an expression that denotes the most prominent referent whose value

---

[30] This observation is well documented. See e.g. Kameyama (1996), Kehler (2002), and Bittner (2014).

[31] Languages with a flexible syntax can exploit word order to indicate prominence. Some grammaticalize prominence with morphemes like topic markers that crosscut word order and grammatical role. See Kameyama (1986), Walker, Masayo, and Cote (1994) and Bittner (2014).

satisfies the pronoun other than x_0—meaning the top-ranked male other than the subject.) We get (16) and (17):[32]

16          $\langle \alpha\, 0 \rangle$; [man($x_0$)]; [walk.in($x_0$)]; $\langle \pi0@he \rangle$; [sit.down(x_0)]

17          $\langle \alpha\, 0 \rangle$; [man($x_0$)]; $\langle \pi1s \rangle$; [met(x_0, x_1)];

          $\langle \pi\, 0@he \rangle$; $\langle \pi1@he^{x_0} \rangle$; [greeted($x_0$, x_1)]

This approach achieves the most uniform possible syntax-semantics interface. In particular, our formal representation consists entirely of updates readable directly off the lexical items that comprise the utterance, along with their grammatical roles.

What about the second challenge created by (13)? While (13i) is explained in parallel fashion to (12), by appealing to subject preference for 'he' taking John, and Principle B for 'him' taking Tim, (13ii) seems to run counter, for, though there is a prominent, accessible antecedent in subject position ('John') satisfying the character of the pronoun, the preferred resolution of 'he' is to the referent introduced by the antecedent in object position ('Tim'). This is where coherence enters.

4   Coherence

Coherence theory starts from an obvious observation: a discourse is more than a sequence of grammatical sentences. Successive contributions must be linked together by a recognizable flow of interpretive relationships. We see the requirement of coherence in (18) and (19), from 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.

(18) doesn't just reveal two facts about John: it suggests the reason he went to Istanbul was to visit his family. A coherence relation of *Explanation* links the second sentence to the first.[33] Coherence theory recommends we represent this connection explicitly to capture the correct interpretation of (18). Conversely, the requirement that a discourse must be coherent is strikingly evident in the interpretive effort (19) elicits. Given apparently unrelated facts about John in (19), we search for a connection. Is Istanbul known for its spinach? Is the train? Clearly, interlocutors

---

[32] Since the pronoun is in subject position, the update—'$\langle \pi0@he \rangle$'—that corresponds to the pronoun promotes the candidate referent associated with 'he' as the top-ranked referent. It does not introduce a new one; 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 is already at the top-ranked position, as in (16)-(17), the update associated with the pronoun will not change possible interpretations of subsequent ones.
[33] See Asher and Lascarides (2003), Hobbs (1979) and Kehler (2002).

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.[34]

Kehler advances the view that coherence relations cluster into (at least) three qualitatively different sorts, reflecting alternative strategies for organizing discourse,[35] as 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.[36]

(20) illustrates the kind of an explanatory discourse we considered in (18). For Kehler, *Explanation* is an instance of a broader class of cause-effect (or event-result) relations that speakers can utilize to organize discourse. (21), meanwhile, gives an extended description of unfolding events—thus, a *Narrative* connection, which for Kehler epitomizes a broader class of *Contiguity* relations. (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. Different coherence relations are alike in signaling relationships among propositions in discourse. However, as we shall see, these relations also shape how other material in a discourse is interpreted. This is particularly important for context-dependent elements. The best way to capture these interpretive effects formally, we argue, is to represent coherence relations explicitly in logical form.

Coherence theorists view *identifying coherence relations* and *resolving semantic ambiguities* as mutually constraining. In (13ii), only resolving 'he' to Tim allows for a plausible explanation of John's disappointment. In (18), only resolving 'there' to Istanbul allows for a plausible explanation of John's trip.[37] In (20)–(22), meanwhile,

---

[34] Representing coherence relations explicitly does not mean an explanatory relation is a part of the truth-conditional content of (18). Though we will argue that connections signaled by coherence relations, like Explanation in (18), are not conveyed as a conversational implicature, this is not at odds with this content being not at-issue, in the sense of Tonhauser, Beaver, Roberts, and Simons (2013); conventional encoding of non-asserted content is not unusual (semantic presuppositions, conventional implicatures, expressive content). For simplicity, our formalism includes only one dimension of semantic content, in the sense of Potts (2005), but we could easily modify our framework to capture conventionalized not at-issue content. However, this task is for another paper.
[35] Kehler's typology is useful in providing an intuitive picture of the kinds of interpretive connections we find in discourse. Other typologies might be better suited for other purposes. *Cf*. Mann and Thompson (1988), and Knott (1996).
[36] See Kehler (2002). (20) and (21) are adapted by Kehler (2002) from Asher and Lascarides (2003).
[37] Note that changing (19) to 'John took a train from Paris to Istanbul. He likes spinach there,' still only allows resolving 'there' to Istanbul for a plausible explanation of John's trip. And since

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.

A number of studies have experimentally confirmed the interdependence of resolving pronouns and establishing coherence.[38] An illustrative example is (23), from Smyth (1994):[39]

23   Phil tickled Stanley, and Liz poked him.

Speakers tend to interpret (23) in either of two ways. 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 Stanley) go in tandem.[40] 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, these studies confirm a

Explanation is naturally signaled in such examples, as in (18) and (19), the pronoun is understood to refer to Istanbul. This further supports our claim that the choice of relation governs the resolution of the pronoun. (Of course, so modified the discourse is no longer as incoherent as the one sequencing random information in (19), because it is clear how the second sentence provides an explanation of the first, as is required by Explanation. But that is beside our point.)

[38] See, for instance, Kaiser, E. (2011), "On the Relation between Coherence Relations and Anaphoric Demonstratives in German," In I. Reich, E. Horch, & D. Pauly (Eds.), *Proceedings of Sinn & Bedeutung* 15, pp. 337–351, Saarbrücken: Saarland University Press, and references therein.

[39] We assume no demonstrative gestures are accompanying (23).

[40] When (23) is embedded within a larger discourse, we might be able to get other resolutions, but this is because embeddings can license different coherence relations. Similarly, prosodic marking can affect pronoun resolution. This is expected, since prosodic marking can affect prominence just as demonstrative gestures can. Likewise, filling in a context might change the interpretive dependences that affect pronoun resolution by virtue of changing the coherence relations organizing a discourse, or by giving rise to other linguistic constraints on interpretation. One might imagine a scenario in which it is a part of the common ground that Liz always copies Phil, and tries to mimic whatever he does. One could then say, "Phil tickled Stanley, and as a result, she poked him", where 'him' is interpreted as Phil. Two things to note: first, we should not assume coherence relations are mutually exclusive. In many cases, more than one is needed to capture the structure of a discourse (see Asher and Lascarides, 2003). This might give rise to a more complex pattern of promoting referents to prominence. And, second, the presence of an explicit descriptive signal such as "as a result" need not automatically mean a particular relation—*Result*—is organizing the discourse (see Webber, Stone, Joshi, and Knott (2003)). Coherence relations are a matter of linguistic dependences and not just any way of describing how events are related gives rise to a particular coherence relation. Similarly, not just any way of describing cause-effect relations counts as Explanation. "as a result" might be targeting a background presupposition about Liz's known pattern of behavior. There is reason to think presuppositions place further constraints on possible antecedents (see Hobbs, 1990).

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 separate ambiguities audiences must resolve in turn—one about discourse coherence, another about pronoun resolution. The examples suggest that, once a coherence relation is established, a certain pronoun resolution is automatically set up. We hypothesize that pronoun resolution is *settled* by whichever coherence relations organize a discourse.

We argue coherence relations are another sort of mechanism that effects changes in the attentional state of a discourse. More precisely, we formally represent coherence relations in the logical form of a discourse, but propose further that this requires representing not only inferential connections, but also shifts in attention associated with coherence relations. Put simply, coherence establishes an interpretive relation, and moreover, (as we shall argue) as a matter of linguistic contribution, they promote certain entities to prominence. 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 are acknowledged as a part of a *linguistic* contribution that we can represent the intuitively correct interpretations of pronouns, while giving pronouns the uniform and unambiguous meanings we have already proposed. We first explain how this account captures 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, not pragmatically implied.

According to us, the difference between resolutions in (13i) and (13ii) is fixed by the coherence relation that figures in their respective representations; these 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 formal representations of (13i) and (13ii). Our proposal 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*, the attention-shifting operation raises John (the subject) to prominence; and when it features *Explanation*, the

attention-shifting operation raises Tim (the direct object) to prominence.[41] Accordingly, we represent (13i) as (24), and (13ii) as (25):

24          $\langle \pi 0j \rangle$; $\langle \pi 1t \rangle$; [was.disappointed.with(x_0,x_1)] ;

            [Narration(x_0)]] ; $\langle \pi 0x_0 \rangle$;

            $\langle \pi 0@he \rangle$; $\langle \pi 1@he^{x_0} \rangle$; [fired$(x_0, x_1)$]]

25          $\langle \pi 0j \rangle$; $\langle \pi 1t \rangle$; [was.disappointed.with(x_0,x_1)] ;

            [Explanation(x_0, x_1)]] ; $\langle \pi 0x_1 \rangle$;

            $\langle \pi 0@he \rangle$; $\langle \alpha 1 \rangle$; [work(x_1)]; [sloppy(x_1)] ;[did(x_0, x_1)]]

(24) introduces John into subject position, and Tim into 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)]] ; $\langle \pi 0x_0 \rangle$', an update that requires that the discourse continues the narrative about x_0, and correspondingly, renders x_0 prominent in the attentional state of the discourse.[42] 'He' picks out the most prominent candidate appropriate for it, which is John, while 'him' picks out the currently most prominent candidate other than John, which is Tim. Further, it is required 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 *Explanation*, which we formalize using the update '[Explanation(x_0, x_1)]] ; $\langle \pi 0x_1 \rangle$'. This update requires that the two bits of discourse stand in an explanatory relation, and promotes the entity in object position to prominence. 'he' continues to pick out the top-ranked interpretation, but due to the update associated with *Explanation*, this is Tim. The formula then proceeds with an additional dynamic existential quantifier, and further conditions that ensure its witness is sloppy work done by the subject Tim.

Crucially, differences in pronoun resolution *follow from* specifications of the coherence relations that organize the discourse and update the attentional state,

---

[41] As we shall see below, not all explanatory coherence relations raise the direct object to prominence. Coherence relations 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 inquiry, not *a priori* judgment, or reasoning about common-sense relations between events in the world.

[42] This representation of *Narration* suffices to capture the effects on attention that concern us. For other purposes, we might want to refine the representation of *Narration* to account for spatio-temporal and causal links we find in narrative discourse. Hobbs (1979, 1990) formalizes these inferential connections in terms of a relationship between eventualities described in successive sentences. Asher and Lascarides (2003) model it as a relationship between dynamic propositions expressed by successive sentences.

and *not* from an open-ended process of pronoun resolution. It might seem surprising to describe attention-shifts associated with coherence as grammatically encoded. By contrast, you might hold that attention shifts contributed by coherence relations only reflect speaker intentions and the hearer's common-sense inference. This view is, in fact, standard among Coherence theorists. So construed, an attention shift that guarantees the correct interpretation of the subsequent pronoun is not a linguistic effect of coherence, as we urge, but rather of pragmatic reasoning that occurs once a hearer has established that a particular coherence relation is structuring the discourse. A hearer reasons that in (13i), since *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. On this view, intention recognition affects the re-ranking of the list of prominent referents.

It certainly *makes sense* for attention to shift in the ways (24) and (25) suggest. It would be perverse to reverse these preferences, so that when we came to consider a narrative about John who is 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.

However, ask not *where* coherence relations shift attention, but rather *in what circumstances* they do so. This question gives us reason for maintaining that attention-shifting operations are grammatically encoded. To see why, first observe that speakers and hearers take a restricted set of cues into account in instances of pronoun resolution. They privilege linguistic ones, 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.

Kehler reports (26) is generally judged to be infelicitous by subjects. What explains this? By virtue of following 'admires' with 'absolutely worships'—a stronger term in an obvious scalar relationship—the speaker provides clear evidence (26) is organized around a contrast between Margaret Thatcher's and George W. Bush's comparable attitudes. Coherence theory predicts this parallel should make Reagan (the object of the first clause) the preferred referent for the pronoun in object position in the second clause.[43] And, indeed, reading (26), it seems as if the speaker

---

[43] Kehler is not committed to this preference being a result of a grammaticized contribution of coherence. He uses (26) for different purposes, but it perfectly illustrates our point.

has erred, inadvertently referring to Reagan as 'her'. Of course, since Thatcher has been evoked in the previous sentence, in 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 the conventionality of demonstration. Thatcher may attract our visual attention with what she's doing, but if the speaker is pointing elsewhere while saying 'her'—or continuing an ongoing discourse about someone else—we do not take 'her' to refer to Thatcher. Likewise, the *Parallel* relation encoded in (26) accomplishes a kind of inferred demonstration, indicating Reagan in a way that's difficult for a common-sense inference to override.[44]

Further support for our thesis comes from the variation we find across languages. Many have explicit operations for shifting attention, such as grammatical topic marking, or a distinction between topic and non-topic pronouns. Some are more constrained than English in the sorts of shifts they permit to take place implicitly. A rough inspection of Serbian, a language that allows (sometimes requires) pronouns to be "dropped", i.e. remain unpronounced in certain grammatical positions, suggests it is a language of this kind. So, 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 its third-person male singular pronoun "dropped" (as in (27i)), the interpretation is that John fired Tim. Yet when the pronoun is overt as in (27ii), the interpretation is Tim fired John. It seems that the overt pronoun signals a change in prominence the covert counterpart does not. English, lacking this explicit means of signaling a shift in attention, is more flexible with implicit shifts in attention; thus, we witness an ambiguity in (13i)-(13ii).

---

[44] Focal stress on 'her' makes (26) felicitous, with the referent of 'her' being Thatcher (as Kehler reports). But this stress carries interpretive requirements, changing the point of the discourse: it explains how Bush follows Thatcher's opinions, not how conservative politicians feel about Reagan. There is no reading where the relation is *Parallel* and the pronoun resolves to Thatcher. But this should be achievable if the attention-shifting effect of *Parallel* were a pragmatic, default one.

The Greenlandic language Kalaallisut provides another clear case. Bittner (2014), 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 arvinilli-t siuqqut-llu-gu billitsi-si-pu-nga.
             month-pl six-pl v.ahead-ELA-3s ticket-get-IND.IV-1s
             I got a ticket (for some other event) six months ahead (of that event).

Following Moens, and Steedman and Webber, we take sentences with subordinate clauses, such as 'when' clauses, to be mini-discourses, with coherent interpretive connections between the clauses that mirror interpretive connections found between successive sentences in a discourse, or between conjuncts.[45] For (28i), 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 (28ii), 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. We witness this discrepancy because Kalaallisut 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 for signaling shifts in attention for temporal anaphora, it, unlike English, seems not to associate coherence relations (viz., *Elaboration*) with attention shifting updates. 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 so, absent other grammatically marked ways of rendering the time before the trip to Denmark prominent, the only way to resolve temporal anaphora in (28ii) is for the trip to Denmark to precede

---

[45] See Moens and Steedman (1988), Webber (1988), Webber *et al.* (2003).

buying the ticket. This interpretation is consistent with *Narrative*, not *Elaboration*, and so, there's no *Elaboration* interpretation of (28ii).

The discrepancy between English and Kalaallisut would be puzzling were the shifts induced by *Elaboration* in English a result of pragmatic reasoning. If it were so, we would expect to find the same range of interpretations in Kalaallisut. However, we do not. Only if languages interpret analogous expressions differently as signals of transitions in discourse can we accommodate the differences we find. So, we infer that each coherence relation in logical form carries a rule-governed contribution to attention in discourse. Given the state of attention, pronoun resolution follows.

Our view is consistent with general reasoning being crucial to interpretation. When the grammar delivers multiple candidate readings, hearers need to choose one that makes the most sense on a given occasion. (23) is ambiguous between a discourse containing *Result* and one featuring *Parallel*. Some general reasoning might be invoked in disambiguating between these, much as it might be involved in figuring out whether a speaker means a financial institution or a river bank, with a use of 'bank', or which quantifier scope is intended with a use of 'Every boy kissed a girl'. To interpret, a hearer must first settle disambiguations. This may involve assessing the plausibility of inferential links conveyed by coherence relations. It may involve evaluating whether a reading engenders a plausible resolution of pronouns. Such reasoning constitutes an important principle of disambiguation, but it cannot contribute content to logical form. It serves to privilege a logical form of discourse, among available ones grammar delivers. But once a coherence relation is established, pronoun resolution is determined by grammar, not general reasoning.

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

5 Conclusion

We defended a joint attention-coherence account for pronoun resolution that assigns one linguistic meaning to each pronoun. Pronouns are variables with dependent interpretations, but are interpreted relative to a prominence ranking. Each resolution is restricted by additional constraints, 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; on our account, it determines the resolution of a pronoun automatically as function of context. To make this work, context must be appropriately set up. This is achieved by the attentional state of a discourse, the result of a series of attention-shifting updates, which intuitively re-rank candidate resolutions for pronouns; formally, they are pushing new entities onto a stack of values for variables, demoting others. These updates are contributed by the diverse mechanisms for structuring discourse and shaping pronominal interpretation, 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 these mechanisms are heterogeneous, their contribution is governed by rules, not pragmatic reasoning, and as such, should be formally represented in logical form.

Our account is provisional in several respects. We said little about the modal profile of context-sensitive utterances.[46] We remained silent about many ambiguities associated with discourses containing pronouns—both within English and across languages.[47] And we have not provided a (exhaustive) list of coherence relations. To see what's at stake, consider (29) from 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 occurrences of 'they' are resolved differently. Our suggestion is that, though in (29i) and (29ii) the content of the second sentence is taken to explain the content of the first, 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 because the demonstrators are potentially violent, or believed to be so by the city council, the council denied them a

---

[46] Still, our formal semantics has resources to capture basic facts about these modal profiles.
[47] For example, we have not offered a treatment of dependent clauses.

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 distinct coherence relations are at play. So, it's no surprise they 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. Only further empirical research can guide us closer to having a firm grasp on all of these issues, and we submit our proposal here to the scrutiny of this future research.

We remain excited about the philosophical ramifications of the tools developed here. Philosophers often use linguistic examples to argue for context dependence. The kind of context dependence they have in mind, implicitly, is where 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, both the linguistic analyses and ensuing philosophical arguments depend on this model of context dependence. But if we are right, context is not as powerful as philosophers have presumed. In interpreting even a straightforward case like (1), our discourse ultimately lacks a dependency on *non-linguistic* context. This suggests a project of extending our account to other cases of apparent contextual variability. If we can succeed in capturing theie apparent contextual variability with uniform meanings and constrained variation, philosophers will have to give up many customary appeals to context sensitivity.

6   Bibliography

Alahverdzhieva, K., & Lascarides, A. (2011). An HPSG Approach to Synchronous Deixis and Speech,. In S. Muller (Ed.), *Proceedings of the 18th International Conference on Head-Driven Phase Structure Grammar (HPSG)*, 6–24. Seattle.

Asher, N., & Lascarides, A. (2003). *Logics of Conversation.* Cambridge Press.

Bittner , M. (2012). Perspectival Discourse Referents for Indexicals. *Proceedings of SULA 7.* Amhrest: GLSA, UMass.

_____ (2007). Online Update: Temporal, Modal, and De Se Anaphora in Polysynthetic Discourse. In C. Barker, & P. Jacobson (Eds.), *Direct Compositionality,* 363–404. Oxford: Oxford Press.

Bittner, M. (2014). *Temporality: Universals and Variation.* Oxford: Wiley-Blackwell.

Chomsky, N. (1981/1993). *Lectures on Government and Binding: The Pisa Lectures.* Mouton de Gruyter.

Cohen, S. (1999). Contextualism, Skepticism, and The Structure of Reasons. *Philosophical Perspectives: Epistemology , 13*, 57–89.

Dekker, P. (2011). Dynamic Semantics. In K. von Heusinger, C. Maienborn, & P. Portner (Eds.), *Semantics: An International Handbook of Natural Language Meaning, Vol. 1.* (pp. 923–945). Berlin: Mouton De Gruyter .

_____ (1994). Predicate Logic With Anaphora. In L. Santelmann, & M. Harvey (Ed.), *Proceedings From SALT IX,* 79–95. Ithaca, NY: Cornell University.

Dowell, J. L. (2011). A Flexible Contextualist Account of Epistemic Modals. *Philosophers' Imprint , 11* (14), 1–25.

Elbourne, P. (2005). *Situations and Individuals.* MIT Press.

Fiengo, R., & May, R. (2006). *De Lingua Belief.* Cambridge, MA: MIT Press.

_____. (1994). *Indices and Identity .* Cambridge, MA: MIT Press.

Grosz, B., & Sidner, C. (1986). Attention, Intentions and The Structure of Discourse. *Computational Linguistics , 12*, 175–204.

Grosz, B., Joshi, A. K., & Weinstein, S. (1995). Centering: A Framework for Modelling the Local Coherence of Discourse. *Computational Linguistics , 21*, 203–225.

Heim, I. (1982). *The Semantics of Definite and Indefinite Noun Phrases.* PhD Thesis, University of Massachusettes Amherst.

Hobbs, J. (1979). Coherence and Coreference. *Cognitive Science , 3*, 67–90.

_____. (1990). *Literature and Cognition* (Vol. Lecture Notes 21). Stanford, California: Center for the Study of Language and Information.

Kameyama , M. (1996). Indefeasible Semantics and Defeasible Pragmatics. In M. Kanazava, C. Pinon, & H. de Swart (Eds.), *Quantifiers, Deduction and Context.* 111–138. CSLI Publications.

_____. (1986). A Property-Sharing Constraint in Centering. *Proceedings of 24th Meeting of Assoc. for Computational Linguistics*, 200–206. New York.

_____. (1999). Stressed and Unstressed Pronouns: Complementary Preferences. In P. Bosch, & R. van der Sandt (Ed.), *Focus: Linguistic, Cognitive and Computational Perspectives* (pp. 306–321). Cambridge University Press.

Kamp, H. (1981). A Theory of Truth and Semantic Representation. In J. Groenendijk, T. Janssen, & M. Stokhof (Eds.), *Formal Methods in the Study of Language* (pp. 277–322). Amsterdam: Mathematisch Centrum.

Kaplan, D. (1989b). Afterthoughts. In J. Almong, J. Perry, & H. Wettstein (Eds.), *Themes from Kaplan.* Oxford University Press.

Kaplan, D. (1989a). Demonstratives. In J. Almong, J. Perry, & H. Wettstein (Eds.), *Themes from Kaplan,* 481–565. Oxford University Press.

Kehler, A. (2002). *Coherence, Reference & the Theory of Grammar.* CSLI Publications.

Kehler, A., Kertz, L., Rohde, H., & Elman, J. (2008). Coherence and Coreference Revisited. *Journal of Semantics , 25*, 1–44.

Keiser, E. (2009). Effects of Anaphoric Dependencies and Semantic Representations on Pronoun Interpretation. In S. L. Devi, A. Branco, & R. Mitkov (Eds.), *Anaphora Processing and Applications* (pp. 121–130). Heidelberg: Springer.

Kendon, A. (2004). *Gesture: Visible Action as Utterance.* New York: Cambridge Press.

King, J. C. (2014 a.). Speaker Intentions in Context. *Nous , 48* (2), 219–237 .

_____. (2014 b.). The Metasemantics of Contextual Sensitivity. In A. Burgess, & B. Sherman (Eds.), *Metasemantics: new essays on the foundations of meaning* ,97–118. Oxford: Oxford University Press.

King, J. C., & Stanley, J. (2005). Semantics, Pragmatics, and The Role of Semantic Content. In G. Zoltán, *Semantics vs Pragmatics,* 111–164. Oxford Press.

Knott, A. (1996). *A Data Driven Methodology for Motivating a Set of Coherence Relations.* Department of Artificial Intelligence, University of Edinburgh.

Kripke, S. (1977). Speaker's Reference and Semantic Reference. In P. French, T. Uehlig, & H. Wettstein (Eds.), *Midwest Studies in Philosophy, Vo. II.* 255–276.

Lewis, D. (1969). *Convention.* Cambridge: Harvard University Press.

Mann, W. C., & Thompson, S. A. (1988). Rhetorical Structure Theory: Toward a functional theory of text organization. *Text , 8* (3), 243–281.

Moens , M., & Steedman, M. (1988). Temporal Ontology and Temporal Reference. *Computational Linguistics , 14* (2), 15–28.

Neale, S. (1990). *Descriptions.* MIT Press.

_____. (2004). This, That, and the Other. In M. Reimer, & A. Bezuidenhout (Eds.), *Descriptions and Beyond* (pp. 68–182). New York: Oxford University Press.

Perry, J. (1979). The Problem of the Essential Indexicaal. *Nous , 13* (1), 3–21.

Potts, C. (2005). *The Logic Of Conventional Implicatures.* Oxford University Press.

Reimer, M. (1992). Three Views of Demonstrative Reference. *Synthese , 93*, 373–402.

Roberts, C. (2002). Demonstratives as Definites. In K. van Deemter, & R. Kibble (Eds.), *nformation Sharing: Reference and Presupposition in Language Generation and Interpretation* (pp. 89–196). Stanford: CSLI Press.

Sidner, C. (1983). Focusing in the Comprehension of Definite Anaphora. In M. Brady, & R. Berwick (Eds.), *Computational Models of Discourse,* 267–330. MIT Press.

Smyth, R. (1994). Grammatical Determinants of Ambiguous Pronoun Resolution. *Journal of Psycholinguistic Research , 23* (3), 197–229.

Stanley, J., & Szabó Gendler, Z. (2000). On Quantifeier Domain Restriction. *Mind and Language , 15* (2-3), 219–261.

Tonhauser, J., Beaver, D., Roberts, C., & Simons, M. (2013). Toward a Taxonomy of Projective Content. *Language , 89*, 66–109.

Walker, M. A., Masayo, I., & Cote, S. (1994). Japaease Discourse and the Process of Centering. *Computational Linguistics , 20* (2), 193–232.

Webber, B. L. (1988). Tense and Discourse Anaphor. *Computational Linguistics, 14* (2), 61–73.

Webber, B. L., Stone , M., Joshi, A. K., & Knott, A. (2003). Anaphora and Discourse Structure. *Computational Linguistics , 29*, 545–587.

Wilkins, D. (2003). Why Pointing With the Index Finger is Not a Universal (in Sociocultural and Semiotic Terms). In S. Kita (Ed.), *Pointing: Where Language, Culture and Cognition Meet* 171–215. Mahwah, NJ: Lawrence Erlbaum.

Winograd, T. (1972). *Understanding Natural Language.* New York: Academic Press.

Wolf, F., Gibson, E., & Desmet, T. (2004). Discourse Coherence and Pronoun Resolution. *Language and Cognitive Processes , 19*, 665–675.

Appendix: Formal definitions

We conclude with complete definitions for a logical language that formalizes our Attention–Coherence approach to pronouns. We assume a set of individual constants C; a set of predicate symbols $\mathcal{P}$, each taking a specified number of arguments; and variables x_i for each natural number i. The interpretation of constants is set up in terms of frames and models in the usual way.

- **A Frame** is a tuple $F = \langle D_w, R, D_e, D_t \rangle$ where D_w is a domain of possible worlds, R is a (transitive and reflexive) accessibility relation on D_w, D_e is a domain of individuals, and D_t is a domain of truth values ($D_t = \{0, 1\}$). We require that the domains be disjoint: $D_t \cap D_w = D_w \cap D_e = D_t \cap D_e = \emptyset$.

- **A Model** is a pair $\mathcal{M} = \langle \mathcal{F}, I \rangle$, where $\mathcal{F}$ is a frame and I is an interpretation function, which assigns to each individual constant an element of $\mathcal{D}_e$ and each n-place predicate constant a set of pairs $\langle w, \sigma \rangle$ with $w \in \mathcal{D}_w$, and σ an n-tuple of elements of $\mathcal{D}_e$.

The Attention–Coherence approach captures interpretive dependencies across formulas using dynamic semantics. Operations on sequences of individuals play a key role in the semantics. Given the individuals defined by a Frame F, we use the following notation to specify these operations.

- i_m
  If m is an integer, then i_m is the mth 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 $u \in D_t \cup D_w \cup D_e$ 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 u of i such that $u \in D_w$.

Now we can define the expressions of the language and their interpretations in a model $\mathcal{M}$.

- 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 $\langle w, i_0 \rangle \in I(p)$.
  $[\![@p]\!]^{i,w} = [\![@p]\!]^{i_{1...},w}$ otherwise.
  (Find most prominent referent that agrees with anaphor.)

- $[\![@p^t]\!]^{i,w} = i_0$ if $\langle w, i_0 \rangle \in I(p)$ and $i_0 \neq [\![t]\!]^{i,w}$.
  $[\![@p^t]\!]^{i,w} = [\![@p^t]\!]^{i_{1...},w}$ otherwise.
  (Find most prominent free referent that agrees with anaphor.)

Conditions:

- If r is an n-place predicate symbol and t_1 through t_n are individual expressions, then $r(t_1, \ldots, t_n)$ is a condition.

The interpretation of conditions:

- $[\![r(t_1, \ldots, t_n)]\!]^{i,w}$ is true if and only if $\langle w, \langle [\![t_1]\!]^{i,w}, \ldots, [\![t_n]\!]^{i,w} \rangle \rangle \in I(r)$
  (interpret atomic conditions by making sure the specified entities are in the specified relation at the world of evaluation)

Dynamic updates:

- $\langle \alpha k \rangle$ is an update, for $k \in \mathbb{N}$.
  (push new indefinite assignment for variable x_k)

- $\langle \pi k t \rangle$ is an update, where t is an individual expression, and $k \in \mathbb{N}$.
  (push new assignment of t as a value of a variable x_k)

- $[\varphi]$ is an update if φ is a condition.
  (restrict the values of variables)

- $H; K$ is an update if H and K are updates
  (composition—conjunction)

- $\Box K$ is an update if K is an update
  (metaphysical necessity)

The interpretation of dynamic updates: At each possible world w, the interpretation of a dynamic update is a relation on sequences:

- $[\![\langle \alpha k \rangle]\!](w, i, j)$ if and only if $j = i_{0.k} + o.i_{k...}$ for some individual $o \in D_e$.

- $[\![\langle \pi k t \rangle]\!](w, i, j)$ if and only if $o = [\![t]\!]^{i,w}$ and $j = i_{0.k} + o.i_{k...}$ for some individual o of type e.

- $[\![[\varphi]]\!](w, i, j)$ if and only if $j = i$ and $[\![\varphi]\!]^{i,w}$ is true.

- $[\![H; K]\!](w, i, j)$ if and only if there is some sequence h such that $[\![H]\!](w, i, h)$ and $[\![K]\!](w, h, j)$.

- $[\![\Box K]\!](w, i, j)$ if and only if $j = i$ and for all worlds v accessible from w, there is some k such that $[\![K]\!](v, i, k)$.

Following Kaplan, we can define an initial context of a model M as any sequence (a, x, y) where $a \in D_w$ (representing the actual world of the context), $x \in D_e$ (representing the speaker of the context) and $y \in D_e$ (representing the addressee of the context). (In a more general language, this could be extended by whatever parameters are appropriate for the interpretation of relevant indexical elements.) Then we can define truth in a context and a model as abstractions over the basic dynamic updates:

- H is true in a model M and initial context J for M, if and only if there is some sequence i such that $[\![H]\!](a, J, i)$.

- H is valid if and only if H is true in every model M for every initial context for M.

- H entails$_S$ K if and only if for any model M and initial context J for M, if H is true at M and J, then K is true at M and J.
  This is the "static" sense of entailment: K is a summary of H.

- H entails$_D$ K if and only if for any model M and initial context J for M, and any assignment i such that $[\![H]\!](a, J, i)$, there is an assignment k such that $[\![K]\!](a, i, k)$.
  This is the "dynamic" sense of entailment: K doesn't add information to H. This version of entailment 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.