

Situating Rich Demonstrations in Discourse*

Christian De Leon

University of California, Los Angeles
cdeleon@humnet.ucla.edu

Abstract

I discuss a phenomenon I call *rich demonstration*, wherein a demonstrative gesture is seemingly used to communicate an entire thought. I present an analysis based in discourse coherence theory on which rich demonstrations have their functions in virtue of the interpretive effects of discourse-structuring mechanisms.

1 Introduction

Standard cases of demonstrative gestures in conversation are deictic—they serve to fix the reference of some word or phrase (McNeill, 1992; Kendon, 2004). (1) is an example of such a case; the speaker employs a co-speech demonstrative gesture to fix the reference of “this”. (2), by contrast, is a case in which the speaker uses a slightly different combination of speech and gesture to accomplish the same communicative goal as in (1) without any reference-fixing.

- (1)
 - a. X: [This]_b is going to make my parents furious.
 - b. [X turns their arm, showing a fresh tattoo]
 - c. Y: But they have tattoos themselves!
- (2)
 - a. X: My parents are going to be furious.
 - b. [After speaking, X turns their arm, showing a fresh tattoo]
 - c. Y: But they have tattoos themselves!

I call gestures like the one in (2b) *rich demonstrations*. This is because they seem to be the causal source of crucial discourse interpretation—in this case, a) that there is some reason for the speaker’s parents to be upset, and b) that the reason is the speaker’s new tattoo. What is puzzling about rich demonstrations is that they are formally identical to simple demonstrations—those gestures that merely supply an object for reference, e.g. (1b)—yet a more complex discourse function is attributable to them. *Prima facie*, the discourse function of rich demonstration is like that of assertion. That is, rich demonstrations seem to be intentionally used to inform, and the information they communicate can be responded to by an interlocutor, typically in the same way.

The purpose for this paper is to investigate the discourse function of rich demonstration. I begin by filling in details of the phenomenon and arguing that general models of discourse focused on common ground (Stalnaker, 1978) fail to capture its features. I present an analysis based in discourse coherence theory (Hobbs, 1979; Kehler, 2002; Asher and Lascarides, 2003) on which a rich demonstration’s apparent complex discourse contribution is actually attributable to the semantic effect of the coherence relation structuring the discourse. On my view, there is no need to assign specific propositional content to perceptual scenes or to rich demonstrations themselves. This contrasts with the analysis given by Hunter et al. (2018) who conceive of the

*I’m grateful to Josh Armstrong, Sam Cumming, Gabe Greenberg, Jessica Rett, and Matthew Stone for discussion and encouragement. Thanks also to the Amsterdam Colloquium’s anonymous reviewers for helpful feedback.

mechanisms of discourse coherence as guiding the ascription of content to gestures and to visual scenes more generally.

2 Puzzle

2.1 More Cases and Details

Grice (1957) famously took the case of Herod's presenting the head of John the Baptist to Salome to not be a case of non-natural meaning (or, for the purposes here, of communication). He (1969) also discussed the related case of a squash player communicating specifically that they are unable to play but not that their leg is bandaged. An adapted version is given in (3):

- (3) a. A: Do you want to go play squash?
 b. [*B displays their bandaged leg*]

It was important to Grice that one cannot communicate what is perceptually apparent. He allowed for, but did not focus on, the possibility that one can use what is perceptually apparent to communicate information that is not. When this is done via a demonstrative gesture, it is a case of rich demonstration. To further illustrate the phenomenon, I will quickly discuss some other cases.

- (4) a. I moved the table into the living room this morning.
 b. [*The speaker nods toward some scratches on the wall*]
 c. I had to buy some new paint.¹

The speaker of (4) does not nod to merely inform the hearer that the wall is scratched. Instead, they intentionally use the nod to inform the hearer of what happened while moving the table—the speaker scratched the wall with it. That information is necessary for understanding the point of (4c) and how the need to buy paint arose from the events earlier in the day.

- (5) a. [I went to the grocery store earlier]_b
 b. [*While talking, the speaker motions to an unopened box of cookies with an open hand*]
 c. ...if you want any.

In (5), the speaker does not merely present two situations, the shopping and the presence of the cookies. Instead, they successfully communicate that they bought that box of cookies while on their recent shopping trip. Recognizing this is necessary for properly interpreting the biscuit conditional in (5c).

- (6) *Maria is helping Carl hunt for a new apartment. While walking up to a new complex:*
 a. [*C taps M's shoulder, then points to the building's pool*]
 b. M: You would get such a bad sunburn.

Maria comments on what would happen in a very specific hypothetical scenario, one in which Carl rents the apartment and decides to frequent the complex's pool. She can only felicitously make that comment given that she understands Carl to be raising that as a live option with his gesture.

¹From Hunter et al. (2018)

These cases illustrate that rich demonstrations do not correspond to any particular formal properties of gestures. Rich demonstrations can take the form of pointing with the index finger, an open hand, nodding the head, and so on. What matters is that the speaker directs the hearer’s attention to something in the conversational vicinity. Rich demonstrations can, but need not, be co-speech, and can even be discourse-initial.

These cases also help illustrate two related reasons for believing that rich demonstrations are functionally assertion-like. First, they exhibit hallmark properties of assertions—they are informative, can be followed up, responded to, etc. Second, as [Hunter et al. \(2018\)](#) observe, rich demonstrations can frequently be replaced by actual assertions without issue. For example, (2b) could be replaced by “I got a new tattoo”, (3b) by “I hurt my leg”, (4b) by “. . . and scratched the wall”, etc., and the same communicative goals would be accomplished. If this is right, an account is needed of how a simple demonstrative gesture can fill the same role as (or a similar role to) assertion.

2.2 Beyond Noticing

Here I want to quickly elaborate on why rich demonstrations are puzzling from the perspective of traditional approaches to the semantics and pragmatics of discourse. It is important to first emphasize that it is not puzzling how the demonstrated eventuality comes to be mutually known in the conversation. In the tradition of [Stalnaker \(1978\)](#), we can conceive of the semantic content of a sentence s as (the characteristic function of) a set of worlds—those worlds in which s is true. Discourse is modeled as a sequence of updates to the Context Set (CS), which represents the body of information compatible with what the discourse participants mutually believe. For now I will represent a discourse as a double consisting in a list of moves and the CS . Assertions update the conversation by being added to the list of moves and by restricting what is mutually believed.²

- (7) For a sentence s , a proposition $\llbracket s \rrbracket = \{w : s = 1 \text{ in } w\}$
- (8) A conversation state $K = \langle M, CS \rangle$:
- a. $M = \langle m_1 \dots m_n \rangle \approx$ an ordered list of moves in the discourse
 - b. $CS = \bigcap \{\llbracket s \rrbracket : \llbracket s \rrbracket \text{ is mutually believed by interlocutors}\}$
- (9) Assertion update: $K_1; [A(s)] = K_2$:
- a. $M_2 = M_1 \cup \{A(s)\} \approx$ add the assertion to the list of moves
 - b. $CS_2 = CS_1 \cap \llbracket s \rrbracket$

Importantly, given its definition in (8b), the CS can be updated by non-linguistic events. [Stalnaker \(2002, 2014\)](#) calls such events *manifest events*; they are events that happen in the conversational vicinity that all parties are mutually aware of. Because they are mutually aware of them, the fact that they happened becomes mutually believed, and so the CS adjusts accordingly. To use his famous example, were a goat to walk suddenly into the room, we would be able to use anaphoric expressions to refer to it, presuppose its presence, and so on (*how did that get in here?*). This means that part of what is informative about a rich demonstration can already be accounted for. That the demonstrated scene is the way it is can be incorporated into the discourse via the same mechanism as all manifest events—via update to mutual belief.

²Technically for Stalnaker an assertion $A(s)$ only directly updates the common ground, which models mutual belief, $CG = \{\llbracket s \rrbracket : \llbracket s \rrbracket \text{ is mutually believed}\}$, $CG_1[A(s)] = CG_2 = CG_1 \cup \{\llbracket s \rrbracket\}$. The CS is then defined in terms of the CG , $CS = \bigcap CG$. So assertions indirectly update the CS . But I follow most in focusing on the CS , and so fold the definition of the CG into the CS , per (8b).

What cannot be accounted for by this mechanism is the communicative purpose a speaker has in richly demonstrating. In (2), that the speaker has a fresh tattoo can be handled via traditional means because that fact is mutually salient to the interlocutors (because the speaker made it salient via their gesture). What is missing is the fact that the tattoo constitutes the reason why the parents will be upset. It is not perceptually available, of course, that the tattoo is the reason, and the speaker failed to explicitly mention that there exists any reason at all. Similarly, (3)-(6) all involve interpretations above and beyond what is manifestly available, and those interpretations are crucial for understanding what has been communicated in each discourse.

3 Coherence and Demonstration

What is missing from a purely *CS*-focused approach to rich demonstration is the conceptual connection between the demonstrated scene and whatever else the speaker is talking about. For this reason, discourse coherence theory (Hobbs, 1979; Kehler, 2002; Asher and Lascarides, 2003) is well-suited for explaining the phenomenon. In pursuing this approach, I follow others in explaining phenomena involving situated discourse by appeal to coherence mechanisms (Lascarides and Stone, 2009a,b; Stojnić et al., 2013; Stone and Stojnić, 2015). Before discussing the details of my proposal, though, I should first discuss an analysis of rich demonstration from Hunter et al. (2018) that is similar in spirit to mine, but diverges at key points.

The project undertaken by Hunter et al. is specifically aimed at modelling how non-linguistic events in a *Settler's of Catan* game can come to be semantically relevant in a conversation among the game's players. But more generally they are motivated by the same kind of data that motivate me here. They present an expansion of Segmented Discourse Representation Theory (SDRT) (Asher and Lascarides, 2003) that allows for the assignment of content to non-linguistic events on the basis of reasoning about coherence relations. That content can interact with standard linguistic SDRT content upon incorporation into a discourse.

What is distinctive about their analysis is that it represents two related conceptual shifts, one from coherence theory specifically and the other from standard discourse modelling generally. With respect to coherence theory, Hunter et al. shift from conceiving of coherence relations as relating discourse segments and affecting discourse on that basis to conceiving of them as holding between events generally, some of which are discourse events, some of which are not. Conceiving of coherence relations in this way leads naturally to the second shift: from taking the mechanisms of discourse interpretation to model how communicative actions have the contents they do to taking them to model how inferences are made about the world generally, even when no communication is involved. These two shifts are made clear when they say,

In our view, semantic structures composed entirely of what are traditionally classified as discourse moves (including, perhaps, discourse dependent nonlinguistic moves) are just a subclass of the kinds of structures that we can use such moves to build. In fact, we think that the kinds of semantic structures built up from coherence relations need not involve any discourse moves at all. Suppose Peter looks out into the garden and sees his cat, Lupin, staring at a pile of leaves. The leaves suddenly move, and Lupin pounces. Peter goes to investigate and finds a baby whipsnake. He now understands why Lupin was staring at the leaves and why the leaves rustled; he also understands that Lupin's pounce was a result of the leaf movement. Yet, neither the snake nor the cat intended to communicate anything, and certainly the snake didn't intend its presence to explain Lupin's behaviour and Lupin didn't

intend his pounce to be a result of the leaf movement. Nevertheless, both the result and meta-level explanation are inferred. (Hunter et al., 2018, pp. 18-19)

The authors take this position regarding discourse for two broad reasons. The first is that we rely on conceiving of external events in particular ways in planning and interpreting discourse events. The second is that specific data (some of which is rich demonstration data) seems difficult to explain without explicitly representing external events as having structured content.

Regarding the first broad reason, while I am generally friendly to the kinds of motivating factors for Hunter et al., I do not think that they necessitate a total reconceptualization of the theoretical role of discourse modelling. But whether this is correct is far beyond the scope of this paper. Here, I merely hope to touch on the second broad reason. In the next section, I present my analysis of rich demonstration and argue that the data can be explained with minimal assumptions from coherence theory and *CS*-focused approaches without the need to model the psychological conceptualization of external events.

4 Analysis

The core of my analysis is simple. Per usual in coherence theory, there are some coherence relations such that when they relate two discourse moves, there is an additional semantic effect on the discourse. Rich demonstrations should be considered full, complete, discourse moves, so this kind of semantic effect is operative in cases of rich demonstration. That is what accounts for their assertion-like quality. So rich demonstrations are functionally like assertions in that they feed some object or scene to be utilized by the semantic effect of the operative coherence relation. The intuition, then, is that the richness of rich demonstrations is accounted for by the very mechanisms that account for the non-explicit content in (10), from Kehler (2002).³

- (10) a. John took a train from Paris to Istanbul.
b. He has family there.

Following Kehler (2002), the two sentences in (10) are related by the *Explanation* relation. When *Explanation* holds, the discourse is updated with the further information that the semantic content of the former obtains because the semantic content of the latter obtains.

4.1 Adding Coherence

To reflect this formally, I update the definitions of *conversation state* and *update* from (8)-(9):

- (11) A conversation state $K = \langle M, R, CS \rangle$:
- $M = \langle m_1 \dots m_n \rangle \approx$ an ordered list of moves in the discourse
 - $R = \{ \langle r, m_i, m_j \rangle \}, \forall m_i \in M, \exists m_j \in M, \exists r \in \mathcal{R} \approx$ a record of coherence relations between moves
 - $CS = \{ w \dots w_n \}$
- (12) Assertion update: $K_1; [A(s)] = K_2$:
- $M_2 = M_1 \cup \{ A(s) \} \approx$ add the assertion to the list of moves
 - $R_2 = R_1 \cup \{ \langle r, A(s), m \rangle \}, \exists r \in \mathcal{R}, \exists m \in M_1 \approx$ add the new move's (contextually-determined) coherence relation with some prior move to the record

³In Kehler (2002), the case is meant to contrast with the classic case of incoherence from Hobbs (1979): "John too a train from Paris to Istanbul. #He likes spinach."

$$c. \quad CS_2 = CS_1 \cap \llbracket s \rrbracket \cap \llbracket \{ \llbracket r(A(s), m) \rrbracket : \langle r, A(s), m \rangle \in R_2 \} \rrbracket$$

M , the record of moves, and how it's updated remain unchanged. I add to K a record of coherence relations between moves R that requires that each move in M coherently relate to some other move(s).⁴ The specific relation r is pulled from \mathcal{R} , the set of all coherence relations.⁵ The determination of which $r \in \mathcal{R}$ obtains is, per usual, a matter of the morpho-syntactic features of the utterance and surrounding discourse, as well as general pragmatic reasoning on the basis of world knowledge. The relations that matter for me here will have semantic constraints, reflected as a set of worlds $\llbracket r(m_n, m_m) \rrbracket$.⁶ Those semantic constraints restrict what the new CS can be in order to capture content that is not explicit in any assertions. The LF of (10) is in (13).⁷

$$(13) \quad [A(\langle 10a \rangle = \text{took-train}(\text{john}))]; [A(\langle 10b \rangle = \text{lives}(\text{john's-fam}, \text{istanbul}))]$$

To see how this captures the basic intuition that the speaker of (10) communicates that John went to Istanbul in order to visit family even though that information is not explicit in the discourse, we can use the following simple model:

$$\begin{aligned} \llbracket \text{took-train}(\text{john}) \rrbracket &= \{w_1, w_2, w_3\} & \llbracket \text{lives}(\text{john's-fam}, \text{istanbul}) \rrbracket &= \{w_1, w_2, w_4\} \\ \llbracket \text{visit}(\text{john}, \text{john's-fam}) \rrbracket &= \{w_2, w_5\} \end{aligned}$$

The first update adds the assertion to the list of moves M and requires that $CS \subseteq \llbracket \text{took-train}(\text{john}) \rrbracket$. $R = \emptyset$ because $A(\text{took-train}(\text{john}))$ is discourse-initial, and so cannot coherently relate to any prior discourse moves. The second adds the new assertion to M , and adds to R that the new assertion relates to the first via *Explanation*. Following Kehler (2002), the reason for explanation involves the tense/aspect of the two sentences as well as the general world knowledge that a normal explanation for taking trips is to visit family. The CS is updated not only with the information that $\text{lives}(\text{john's-fam}, \text{istanbul})$ but also that $\text{visit}(\text{john}, \text{john's-fam})$, which falls out of the semantics of *Explanation*. The evolution of the discourse can be seen in the following representations of K :

$K_1 = [A(\text{took-train}(\text{john}))]$			$K_2 = K_1; [A(\text{lives}(\text{john's-fam}, \text{istanbul}))]$		
M_1	R_1	CS_1	M_2	R_2	CS_2
$\langle A(\langle 10a \rangle) \rangle$	$\{\}$	$\{w_1, w_2, w_3\}$	$\langle A(\langle 10a \rangle), A(\langle 10b \rangle) \rangle$	$\{\langle \text{Exp}, \langle 10b \rangle, \langle 10a \rangle \rangle\}$	$\{w_2\}$

4.2 Adding Demonstration

So far the simple formalism is able to handle conversation update with the semantic effects of operative coherence relations. What is still missing is a way for demonstrated scenes to factor in. To account for this, I make one final revision to the definition of discourse update from (12):

$$(14) \quad \text{Discourse update: } K_1; [\Phi] = K_2 : \\ a. \quad M_2 = M_1 \cup \{\Phi\} \approx \text{add the move to the list of moves}$$

⁴I gloss over discourse-initial assertions for simplicity. The definition also allows for the possibility that new moves coherently relate to multiple prior moves.

⁵I remain mostly agnostic about the composition of \mathcal{R} since my analysis does not essentially depend on which or how many relations exist. I intend this sketch to be compatible with approaches in the style of Hobbs/Kehler, SDRT, RST (Mann and Thompson, 1988), etc.

⁶For purely structural relations like *Parallell*, we can simply assume their semantic value is identical to the set of all worlds \mathcal{W} .

⁷I am ignoring many factors in the representation of (10), e.g. pronoun resolution, that this system could easily be adapted to explain using coherence mechanisms (cf. Stojnić et al., 2017).

- b. $R_2 = R_1 \cup \{\langle r, \Phi, i \rangle\}$, $\exists r \in \mathcal{R}, \exists m \in M_1 \approx$ add the new move's (contextually-determined) coherence relation with some prior move to the record
- c. $\Phi = A(s) \Rightarrow CS_2 = CS_1 \cap \llbracket s \rrbracket \cap \bigcap \{\llbracket r(A(s), m) \rrbracket : \langle r, A(s), m \rangle \in R_2\}$
- d. $\Phi = D(\delta) \Rightarrow CS_2 = CS_1 \cap \bigcap \{\llbracket r(D(\delta), m) \rrbracket : \langle r, D(\delta), m \rangle \in R_2\}$

I distinguish between two kinds of update, assertion $A(s)$ and demonstration $D(\delta)$, where δ is a perceptual scene. Any discourse move Φ updates M and R in the same way—by being added to M and by resolving to an appropriate coherence relation r . Assertions and demonstrations update the CS in slightly different ways. Assertions restrict the CS to the content of the asserted sentence as well as the content of the operative coherence relation(s). Demonstrations restrict the CS only with respect to the operative coherence relation(s). The important result of this is that the semantic content of coherence relations accounts for all of the communicative informational updates triggered by the gesture. That is, I do not employ coherence mechanisms to interpret or determine the content of the gesture or visual scene, as do Hunter et al.. Instead, I employ them to account for the general communicative strategy used by the speaker. To see how this explains the intuitions surrounding rich demonstration, I will use a simple model paired with the LF of (2) in (15).

$$(15) \quad [A(\langle 2a \rangle = \text{furious}(\text{parents}))]; [D(\langle 2b \rangle = \text{TAT})]; [A(\langle 2c \rangle = \text{has-tattoo}(\text{parents}))]$$

$$\begin{aligned} \llbracket \text{furious}(\text{parents}) \rrbracket &= \{w_1, w_2, w_3, w_4\} & \llbracket \text{cause}(\text{tattoo}, \text{furious}) \rrbracket &= \{w_2, w_3, w_4\} \\ \llbracket \text{has-tattoo}(\text{speaker}) \rrbracket &= \{w_1, w_2, w_3\} & \llbracket \text{has-tattoo}(\text{parents}) \rrbracket &= \{w_2\} \end{aligned}$$

$K_1 = [A(\text{furious}(\text{parents}))]$			$K_2 = K_1; [D(\text{TAT})]$		
M_1	R_1	CS_1	M_2	R_2	CS_2
$\langle A(\langle 2a \rangle) \rangle$	$\{\}$	$\{w_1, w_2, w_3, w_4\}$	$\langle A(\langle 2a \rangle), D(\langle 2b \rangle) \rangle$	$\{\langle \text{Exp}, \langle 2b \rangle, \langle 2a \rangle \rangle\}$	$\{w_2, w_3\}$
$K_3 = K_2; [A(\text{has-tattoo}(\text{parents}))]$					
M_3		R_3		CS_3	
$\langle A(\langle 2a \rangle), D(\langle 2b \rangle), A(\langle 2c \rangle) \rangle$		$\{\langle \text{Vio}, \langle 2c \rangle, \langle 2b \rangle \rangle, \langle \text{Exp}, \langle 2b \rangle, \langle 2a \rangle \rangle\}$		$\{w_2\}$	

K_1 is the result of the first update, the assertion of `furious(parents)`. The assertion is added to M and the CS shrinks such that $CS \subseteq \llbracket \text{furious}(\text{parents}) \rrbracket$. The demonstration in (2b) is then added to M and recorded in R as coherently relating to the prior assertion via *Explanation*. I elaborate on what goes into the selection of the appropriate coherence relation with respect to rich demonstrations in §4.3, but for now the fact that the assertion of (2a) heavily biases *Explanation* should suffice.⁸ The CS is updated so that it is a subset of $\llbracket \text{cause}(\text{tattoo}, \text{furious}) \rrbracket$, which is the semantic fallout of $\text{Exp}(A(\text{furious}(\text{parents})), D(\text{TAT}))$. But that by itself would merely result in $CS = \{w_2, w_3, w_4\}$. The reason that $w_4 \notin CS_2$ is that as a byproduct of the gesture, it is manifestly obvious to all interlocutors that `has-tattoo(speaker)`. That is, typical Stalnakerian mechanisms for CS -updates are perfectly capable of reflecting what is mutually known in a conversation, as discussed in §2.2. Finally, the third discourse move $A(\text{has-tattoo}(\text{parents}))$ updates M in the normal way, records in R that the assertion stands in the *Violated Expectation* relation with $D(\text{TAT})$, and shrinks the CS in the appropriate way.

I have so far said little about the visual scene δ (in this case, $\delta = \text{TAT}$) that is the argument to the demonstration $D(\cdot)$. One might worry that without an account of how δ is conceptualized in the minds of the speaker and hearer, the general account of rich demonstrations is lacking. After all, it will matter for the discourse whether δ is the tattoo on the speaker's arm, the

⁸This is a large part of why *Explanation* is operative in (1), which involves the same discourse-initial assertion, followed by another assertion instead of by a rich demonstration.

image it depicts, or something behind the arm entirely. This is the kind of consideration that leads Hunter *et al.* to discuss visual interpretation beyond discourse. I choose to remain mostly agnostic about the nature of δ 's content for two reasons.

First, while it is undoubtedly important to investigate how visual signals are interpreted, I consider such an investigation peripheral to the question of the discourse function of rich demonstration. So long as δ is capable of encoding information of the same type as is normally used in coherence-based explanations, the account here captures how and why rich demonstrations fulfil the task they intuitively do. I take the question of δ 's resolution to be analogous to the question of how phonetic items are resolved to particular semantic items; such questions are deeply important, but their answers are not answers to questions about how semantic items themselves serve the communicative functions they do. The analysis presented here is intended to (begin to) answer questions regarding the semantic and pragmatic functions of rich demonstration.

Second, filling in the content of δ on the basis of reasoning about what the communicative goals of the speaker are risks limiting the discourse representation unnecessarily, and thereby wrongly predicting that some followups are unavailable. For example, it may be genuinely unclear which aspect of the tattoo will upset the parents. My intuition is that what the speaker should be taken to say is, roughly, nothing more than that their parents will be upset because of the tattoo. Were the interlocutor to ask why the parents would be upset, they should be taken to ask a question about the psychology and motivations of the parents, not a metalinguistic question about what the speaker in fact said. This means that leaving the details of δ 's content to the perceptual faculties and not to the mechanisms of discourse interpretation will suffice for investigating how rich demonstrations accomplish their tasks.

This analysis has consequences for the theoretical question about the meanings of gestures. My preferred interpretation of this analysis is that rich demonstrations are themselves no more meaningful than are simple demonstrations (i.e. gestures that resolve referential expressions, as in (1b)); they merely designate a perceptual scene. What is interesting about our gestural conventions is that such a designation can suffice for accomplishing a complex communicative task—the same task accomplished by assertion. The reason this is possible has its roots in discourse structure. On a coherence-theoretic approach, explicit moves must relate in various ways to one another, and the specific ways they relate result in informational updates above and beyond the updates mandated by the meanings of the moves. That rich demonstrations exist shows *a*) what kind of actions can be counted as full discourse moves in their own right, and *b*) that some gestures, *qua* full discourse moves, need not have propositional meanings in order to function. Assertions are of propositional contents and also cause further propositional updates; rich demonstrations can cause those same further updates despite not being meaningful in the way assertions are.

4.3 Constraints on Coherence

The analysis I have presented relies on demonstrative discourse moves being related for appropriate coherence relations. I would like to conclude by briefly discussing the constraints on the selection of available coherence relations for connecting rich demonstrations to surrounding discourse moves. The determination of the operative coherence relation connecting two assertions is a matter of pragmatic reasoning on the basis of the grammatical features of each assertion, and the same is true when connecting an assertion with a rich demonstration.

First, the preceding assertion will bias or require resolution to some coherence relations, per usual. This will be a function of the particular phrases used and the sentence's tense/aspect.

I claim further that rich demonstrations require (or at least strongly bias) subordinating coherence relations to prior discourse. That is, a rich demonstration must maintain the current discourse topic or goal. Particularly common relations, as is obvious at this point, are *Explanation* and *Elaboration*. The former is operative in cases such as (2), the latter in cases like (4). It might seem at first glance that some coordinating relations, namely *Result* (which can be considered the dual of *Explanation*), are possible candidate relations. I contend, however, that *Result* readings are strange at best. Contrast the following cases:

- (16) *Explanation*:
- a. [*The speaker gestures toward a broken vase*]
 - b. Alex threw a football earlier.
- (17) *Result*:
- a. Alex threw a football earlier.
 - b. ?[*The speaker gestures toward a broken vase*]

The intuitions about these cases are admittedly subtle, but instructive. It is clear what the speaker intends to convey in (17), but it is difficult to force the gesture to come after the completion of the utterance. It is much more natural for the gesture to occur during the utterance or before it, as in (16). But such an occurrence would permit *Explanation*.⁹ The connection between throwing a football and breaking a vase is easy enough to make, but it feels like the speaker would more naturally point as a followup, not to introduce a new event to be discussed. That (16) is arbitrarily more natural than (17) is explained by the fact that (17) only makes sense with a coordinating relation, which rich demonstrations resist. Nevertheless, there are variants where the gesture's being after the utterance is perfectly natural:

- (18) *Elaboration*:
- a. Alex was throwing a football around,
 - b. [*The speaker gestures toward a broken vase*]

Changing the aspect of the sentence to the progressive makes the trailing gesture more natural. The reason for this is that the progressive biases *Elaboration*. The eventuality described in (16) and (17) involves one throwing event, but the eventuality described in (18) involves many throwing events. Expanding the eventuality to Alex's general activity means that the broken vase eventuality can be contained within it. That is, more details can be given about the general activity, which the speaker does via the rich demonstration. This is, of course, a first gloss. More research should be done into the constraints gestures impose on discourse structure.

5 Conclusion

Broadly, the purpose for this paper was to expand on current research into gestural communication and situated conversation more generally within a coherence-theoretic framework. Specifically, I have focused on the phenomenon of rich demonstration and argued for an analysis on which its complex discourse function is accounted for by being a discourse move in a full sense, and thereby being capable of relating to surrounding discourse via coherence relations.

⁹Note, however, that my claim is that rich demonstrations ought to relate to *prior* discourse via subordinating relations. I make no claim about posterior discourse. So (16) is not an instance of my prediction; its purpose is merely to provide a more natural way of using a gesture to communicate the intended thought in (17).

This analysis makes clear several paths for further research into multimodal communication. First, other phenomena should be investigated and situated with respect to their expressive content and their discourse-structural function. Other gestures, facial expressions, gaze, and more all have deep communicative effects to be studied. Second, further research can be done with respect to the nature of coherence relations specifically and discourse interpretation generally. As I touched on in §3, I have approached this topic with an eye toward maintaining the distinction between discourse interpretation and general reasoning about world events. I have not argued for the distinction, but take myself to have provided an analysis of a phenomenon that is compatible with it.

References

- Nicholas Asher and Alex Lascarides. *Logics of Conversation*. Cambridge University Press, 2003.
- Herbert Paul Grice. Meaning. *The Philosophical Review*, 66(3):377–388, 1957.
- Herbert Paul Grice. Utterer’s meaning and intention. *The Philosophical Review*, 78(2):147–177, 1969.
- Jerry R. Hobbs. Coherence and Coreference. *Cognitive Science*, 3(1):67–90, 1979.
- Julie Hunter, Nicholas Asher, and Alex Lascarides. A formal semantics for situated conversation. *Semantics and Pragmatics*, 11(10), 2018.
- Andrew Kehler. *Coherence, Reference, and the Theory of Grammar*. Center for the Study of Language and Information, 2002.
- Adam Kendon. *Gesture: Visible Action as Utterance*. Cambridge University Press, 2004.
- Alex Lascarides and Matthew Stone. Discourse coherence and gesture interpretation. *Gesture*, 9(2):147–180, 2009a.
- Alex Lascarides and Matthew Stone. A Formal Semantic Analysis of Gesture. *Journal of Semantics*, 26(4):393–449, 2009b.
- William C. Mann and Sandra A. Thompson. Rhetorical structure theory: Toward a functional theory of text organization. *Text*, 8(3), 1988.
- David McNeill. *Hand and Mind*. The University of Chicago Press, 1992.
- Robert Stalnaker. Assertion. In Peter Cole, editor, *Pragmatics*, pages 35–332. Academic Press, 1978.
- Robert Stalnaker. Common ground. *Linguistics and Philosophy*, 25(5/6):701–721, 2002.
- Robert Stalnaker. *Context*. Oxford University Press, 2014.
- Una Stojnić, Matthew Stone, and Ernie Lepore. Deixis (Even Without Pointing). *Philosophical Perspectives*, 27:502–525, 2013.
- Una Stojnić, Matthew Stone, and Ernie Lepore. Discourse and Logical form: Pronouns, Attention, and Coherence. *Linguistics and Philosophy*, 40:519–547, 2017.
- Matthew Stone and Una Stojnić. Meaning and Demonstration. *Review of Philosophy and Psychology*, 6(1):69–97, 2015.