

Discourse and Method

Ethan Nowak and Eliot Michaelson*

Abstract

Stojnić et al. (2013, 2017) argue that the reference of demonstratives is fixed without any contribution from the extra-linguistic context. On their ‘prominence/coherence’ theory, the reference of a demonstrative expression depends only on its context-independent linguistic meaning. Here, we argue that Stojnić et al.’s striking claims can be maintained in only the thinnest technical sense. Instead of eliminating appeals to the extra-linguistic context, we show how the prominence/coherence theory merely suppresses them. Then we ask why one might be tempted to try and offer such a view. Since we are rather sympathetic to the motivations we find, we close by sketching a more plausible alternative.

1 Introduction

If you point at the V-notch couloir on the North Palisade and utter (1), most people will take you to have said something false.¹ If you point at the U-notch couloir instead, while uttering the same sentence, people will take you to have said something true:

(1) That is the U-notch couloir.

Linguists and philosophers generally agree that the two utterances should be associated with different truth conditions, which result from differences in the extra-linguistic contexts in which the two utterances take place. Some think that the speaker’s gesture determines which object the demonstrative picks out. Some think her

*Forthcoming in *Linguistics and Philosophy*. This work is entirely collaborative and names are listed in reverse alphabetical order.

¹Nathaniel Hansen reports the intuition: ‘What the fuck is a V-notch couloir?’ We are confident that the arguments we offer here are consistent with this response. By way of information, a couloir is a steep and narrow channel on the side of a mountain, typically filled with ice and snow. The V-notch is a couloir on the North Palisade in the Sierra Nevada mountains of California that looks like the letter ‘v’.

referential intentions are what does this work. Some approaches weigh these two things against one another, and others invoke further considerations still.

In the Golden Age of philosophical work on demonstratives, debates between theorists advocating one or another of these positions were treated as semantic debates.² It was generally assumed that an adequate compositional semantic theory would take a demonstrative sentence and a context and return a determinate truth condition. More recently, philosophers have been attracted to analyses on which, semantically speaking, demonstratives are represented simply as variables. On this way of thinking, the classic disputes turn out to be disputes about which principles should be used to link contexts to particular variable assignments.³

Stojnić et al. (2013, 2017) have recently attracted significant attention by rejecting the framework of the classic debates entirely. According to what they call the ‘prominence/coherence’ theory of demonstratives, the question of which feature of a context fixes the referent of a demonstrative is fundamentally confused. On their view, the context-independent linguistic meaning of ‘that’ fixes its reference without *any* contribution from speaker intentions, demonstrations, or other features of the speech situation in which the expression is used.

Our primary aim here is to show that Stojnić et al.’s striking claim can be sustained only in the thinnest technical sense. We will argue that instead of eliminating appeals to extra-linguistic context, the prominence/coherence theory simply relocates them. This relocation, in our view, fails to resolve any of the thorny questions that arise regarding the way in which the extra-linguistic context helps to determine the intuitive truth conditions of a particular use of a demonstrative sentence. Properly understood, Stojnić et al.’s proposal raises exactly the same issues—and is susceptible to exactly the same sorts of challenges—as every other extant theory of demonstratives of which we are aware.

Having established this negative thesis, we turn our attention to a discussion of its significance. We structure that discussion around two connected questions. First, why might one be tempted to offer a picture along the lines of Stojnić et al.’s? And second, if we are right that the authors fail to deliver the theoretical goods promised, might there be some other way of obtaining them? To preview: we suspect that Stojnić et al. are eager to avoid invoking the extra-linguistic context because they worry that fractious debates about which variable assignment should count as ‘the one true assignment’ in a context lead nowhere.⁴ We share this concern; what once appeared

²Artifacts from this Golden Age include Kaplan (1977, 1978, 1989), Bertolet (1980), McGinn (1981), Schiffer (1981), Wettstein (1984), Reimer (1991, 1992), and Bach (1992).

³For discussion, compare Rabern (2012a,b), Yalcin (2014), Nowak (2016), and Nowak (forthcoming).

⁴For a representative sample, see Neale (2004), Gauker (2008), Mount (2008), King (2013, 2014a,b), and Speaks (2016, 2017).

to be a healthy ecosystem of competing views has, in certain regards, begun to look like a degenerative research program.⁵ To really move away from the old picture, however, we suspect that something more radical than Stojnić et al.’s suggestions might be required. One response, which we take to have at least some appeal, would be to give up the idea that demonstrative sentences in a context really have canonical truth conditions in the first place.

2 The prominence/coherence theory

Here is the thesis that we want to dispute:

The semantic value of a pronoun is never determined, even partly, by extra-linguistic cues; it is fixed, invariably and unambiguously, by features of its context of use governed entirely by linguistic rules. (Stojnić et al., 2017, pg. 519)

Except for the part about extra-linguistic cues, this thesis sounds perfectly consonant with longstanding philosophical tradition. Most theorists endorse one or another variation on the following template, according to which the semantic value of a demonstrative is fixed invariably and unambiguously by some specific feature (or specific combination of features) of its context of use:⁶

- (2) $[\text{that}]^{c,w}$ = the object ostended/intended/etc. by the speaker of c

For their part, Stojnić et al. claim that the semantic value of a demonstrative in a context is the object that is at the *center of attention* in the context. They do not think their positive proposal amounts to just another way of filling out the standard template from (2), however. By ‘center of attention’, they do not mean something determined by any actual or hypothetical person’s psychological state, and by ‘context of use’, they do not mean a representation of the concrete speech situation in which a demonstrative is or might be tokened (Stojnić et al., 2017, pp. 524-526). It is worth pausing here to elaborate on these potentially-confusing claims.

Stojnić et al. frame their view using dynamic semantic machinery on which contexts are simply sequences of objects. In their terms, the object that occupies the first position in a sequence is said to be the *most prominent*, or to be at the *center of attention*. Their key claim is that the object that occupies that position is determined by linguistic rules which require no supplementation from the speech situation.

⁵Compare Lakatos (1974).

⁶Read (2): the extension of ‘that’ with regard to any context of utterance and circumstance of evaluation is the object ostended by the speaker of the context. Compare semantic theories from Kaplan (1977, 1978), Reimer (1991, 1992), Bach (1992), Salmon (2002), and Caplan (2003).

Some of those rules, which are characterized in terms of functions from contexts into contexts, have recognizable analogues in existing work on dynamic semantics. The rule Stojnić et al. associate with the existential quantifier, for example, takes an arbitrary input context and returns a variation on which the order of the objects is preserved, but on which they are all shifted down one place in the sequence to allow the introduction of a new object in the first position. This means that demonstrative expressions that occur after an indefinite will refer to it anaphorically, assuming nothing else intervenes. Witness:

(3) A man walked in. He sat down.

Stojnić et al. model this discourse by means of the following series of updates, where ‘ α ’ represents the existential quantifier and ‘@’ a pronominal element that picks out the first object in the sequence with regard to which it is evaluated:⁷

(4) $\langle \alpha 0 \rangle$; [*man*(@)]; [*walk.in*(@)]; [*sit.down*(@)]

The operation associated with the existential quantifier takes an input context and adds an unspecified new object to the first (the 0-th) position, moving all the other objects down one spot (Stojnić et al., 2017, pg. 527).⁸ The remaining updates specify conditions on the new object. Analyzed as per (4), sentence (3) is true with respect to a model and an input context just in case there is some object in the model that is a man who walked in and sat down. Intuitively, that is the right result.

Most of the update operations that Stojnić et al. define implement coherence relations of the sort described by Kehler (2002), and are used to provide an account of the resolution of pronominal anaphora. Instead of relying on speakers’ and interpreters’ general pragmatic competence, Stojnić et al. claim that coherence relations encoded in the grammar modify prominence relations, which in turn provide referents for pronouns. While this is an intriguing idea, we will focus here primarily on those instances where—even according to Stojnić et al.—coherence relations play a fairly minimal role. This will simplify the dialectic to follow.⁹

For now, we want to focus on a pair of operations that have a specific role to play in the interpretation of deictic uses of demonstratives. Both of those operations im-

⁷For present purposes, no harm will come from ignoring differences between the sets of features (e.g. number, grammatical gender, person) that characterize various pronouns and demonstrative elements. We have no wish to contest that Stojnić et al.’s system is capable of handling these.

⁸We must confess that we are unsure what exactly an ‘unspecified new object’ amounts to here. Perhaps this is just shorthand for a set of instructions, i.e. ‘select an arbitrary object from the model’. If not, we worry that this sounds a great deal like Russell (1903)’s claim that the indefinite description ‘a man’ denotes an ambiguous man—a claim which has proved notoriously difficult to make sense of, and which Russell himself subsequently abandoned for just these reasons.

⁹We take up the role played by coherence relations in §3.2, as they are helpful for understanding how Stojnić et al. (2017) think about demonstratives unaccompanied by pointing gestures.

port objects into the linguistic context from the extra-linguistic context.¹⁰ One of them, π , is used when a demonstrative occurs together with a pointing gesture. The other, σ , is used in cases where a demonstrative is tokened without an accompanying gesture.

Consider (1) again:

- (1) That is the U-notch couloir.

Intuitively, this sentence should be true in a context in which the speaker is pointing at the U-notch couloir. Stojnić et al. derive that result by offering the following analysis:¹¹

- (5) $\langle \pi 0 U\text{-notch} \rangle; [is.the.U.notch.couloir(@)]$

When a demonstrative is uttered together with a pointing gesture, Stojnić et al. say that the gesture contributes a π -update to the LF associated with the demonstrative sentence (Stojnić et al., 2013, pg. 516). Like the update associated with the existential quantifier, π -updates modify the input context by adding a new object to the top slot. Unlike the existential update, however, π -updates involve a determinate object—the object ostended by the person uttering the demonstrative.¹²

If we use (5) to analyze (1), we make the correct prediction about the intuitive truth value of the sentence in the situation described.¹³ The π -update modifies the input context so that the U-notch couloir is the first object in the sequence; every other object moves down one position. The demonstrative element, @, picks out the top object in the context, which is then checked against the condition of being the U-notch couloir. We end up with the result that (1) is true with regard to a model and an input context just in case the U-notch couloir is in the domain of the model and is identical with the U-notch couloir.¹⁴

¹⁰The operations in question, that is, import objects from the context, construed after the fashion of Kaplan (1977), into the context, construed à la Heim (1982).

¹¹Where ‘*U-notch*’ is a meta-language name for the U-notch couloir. Read (5) as an instruction to first put the object designated by ‘U-notch’ into the 0-slot of the input context, whatever it may be, and then to test the result against the condition that the top object be the U-notch couloir.

¹²Actually, Stojnić et al.’s proposed operator is more flexible than this, being able to push objects to other slots in the sequence as well. $\langle \pi 1 U\text{-notch} \rangle$, for instance, would push *U-notch* to the second slot in the sequence, the 1-slot. We elide this additional complexity, however, as it is used to deal with sentences involving multiple pronominal elements and we will consider no such sentences below.

¹³We would note, however, that (5) is true in all possible worlds; in other words, it is necessary. This derives from the fact that being a U-notch couloir is an essential property of the referent of the meta-language name *U-notch*, given that we have stipulated the referent of that name. It strikes us as highly implausible that any utterance of (1) will be necessary. After all, in some nearby possible worlds, the couloir will, due to slightly different patterns of erosion, resemble a ‘v’ more than a ‘u’.

¹⁴We leave it to the reader to set up the case in which (1) is uttered by someone pointing not at the U-notch, but at the V-notch or at something else entirely.

The same basic strategy is used to explain the intuitive data in cases that do not involve gestures. If we are standing on the Palisade Glacier, looking southwest towards a prominent channel of snow and ice that breaks up a rock face,¹⁵ you might use (1) to say something true without pointing at anything. According to Stojnić et al., the LF associated with (1) in such a case would be:

(6) $\langle \sigma 0 s_0 \rangle; [is.the.U.notch.couloir(@)]$

Whereas the π operation involved an individual constant, the σ operation features a term denoting a situation, in the sense of Barwise and Perry (1983) and Kratzer (2002).¹⁶ Stojnić et al. do not address the question of how this situation is determined in great detail, but their discussion of examples suggests that it will typically be a part of the situation in which the speech act occurs. A σ operation has the effect of taking the central individual, the central location, and the central event from the relevant situation and putting them into the top three spots of the context, pushing all other objects down accordingly (Stojnić et al., 2013, pg. 516).¹⁷

In the case where we are standing on a glacier and you speak without pointing, they will say that the central object of the situation in question is the U-notch couloir. So, the input context is modified in such a fashion that the new top object is the U-notch couloir. When the demonstrative element is interpreted, then, it picks out the U-notch, which of course satisfies the condition of being the U-notch.

3 Old problems in a new guise

We accept that there is a perfectly clear sense in which the semantic value of a demonstrative, on Stojnić et al.’s account, does not depend on the extra-linguistic context in which it is uttered. We submit, however, that there is just as clear a sense in which the extra-linguistic context plays as significant a role for them as it does for everyone else. While there may well be genuine work for their machinery to do—this seems especially likely where the treatment of anaphora is concerned—it

¹⁵Michaelson (2013) points out that directed glances might, in fact, be considered a sort of gesture. If you are inclined to agree, you should analyze this example using a π update instead of σ , and reframe the example using an object that could be picked out by way of a demonstrative in the absence of any obvious gazing, like a loud noise, a smell, or a tremor.

¹⁶Hat tip to Locke (1689) for providing the substrate of minimal situations that today’s theoretical edifice adheres to.

¹⁷Again, we are eliding some of the complexities of Stojnić et al.’s system here. In its fullest framing, that system also includes more complex pronominal elements than ‘@’, elements which are, for instance capable of accessing the 1-slot or 2-slot. This proves useful in modeling the behavior of demonstratives like ‘there’ or ‘then’, which are used to refer to places or times rather than the individuals which typically inhabit the 0-slot (Stojnić et al., 2013, pg. 518). Effectively, ‘@’ abbreviates ‘@that’.

is important to recognize that by itself, that machinery does nothing to address the fundamental questions raised by the basic demonstrative data.

At the end of the day, what we expect from a theory of demonstratives is an answer to the question of why two different truth conditions should be associated with a demonstrative sentence when it is uttered by someone pointing at two different objects. Historically, this has usually been taken to be a semantic question; demonstratives were assumed to be semantically sensitive to context, and the challenge was to identify the particular contextual feature or features they track. More recently, the idea that the question is a post-semantic one has gained traction; demonstratives are treated as variables in the compositional semantics, and the challenge of the intuitive data is to identify the principles that link a context of utterance to a particular variable assignment. If Stojnić et al. are right about the semantics (‘that’ picks out the top object in a context) and metasemantics (the top object is determined by π/σ -updates) of demonstratives, the intuitive challenge might manifest in yet another guise, but it does not simply dissipate. Now we must ask: in virtue of what should we associate a particular π/σ -update with a particular speech situation? We see no way of answering that question without invoking details of the extra-linguistic context.

3.1 Deixis by pointing

As we have just seen, analyzing (1) as per (5) gets us the expected truth conditions for cases where the speaker is pointing at the U-notch:

- (1) That is the U-notch couloir.
- (5) $\langle \pi 0 U\text{-notch} \rangle; [is.the.U.notch.couloir(@)]$

As Stojnić et al. emphasize, (5) has a perfectly determinate meaning that is constant across both concrete speech situations and also across contexts, taken in the authors’ technical sense. We fail to see, however, how this fact makes (1) interestingly insensitive to context, understood in the more traditional, Kaplanian sense. If we accept Stojnić et al.’s proposal, the challenge of explaining the intuitive data takes the form of the following question: why should we analyze (1) along the lines of (5)? Why should we not understand (1) as involving some other possible π -update instead—for instance, $\langle \pi 0 V\text{-notch} \rangle$, or an update that puts something else altogether into the top spot of the sequence?

At this point, the available answers start to look like all the old familiar ones. We might say that (5) is the right representation because the speaker intended to refer to the U-notch. Or we could say that it is because the speaker’s gesture was directed

at the U-notch. Or we could blend these two responses together, or say something more sophisticated still.

As far as we can tell, however, any non-deflationary response to this question is going to involve an appeal to elements of the extra-linguistic context. Even if one refuses to provide any details, saying only that ‘Context determines which π -update is involved here,’ we see no way to avoid the conclusion that it is the *extra-linguistic* context that does the work. Where (1) appears discourse initially, there is simply no linguistic context to appeal to. And even when (1) appears within an extended discourse, one of the interesting things about deictic uses of demonstratives is that they can shift the linguistic context by introducing new discourse referents drawn from the extra-linguistic context.

Stojnić et al.’s discussion makes clear that they think gestures bear significant explanatory weight in cases like these. They offer plausible reasons—like intralinguistic stability and interlinguistic variability—for treating gestures as full-blown linguistic expressions (2017, pp. 528-529). Remarkably, however, the treatment they offer for the semantics of gestures does not invoke the extra-linguistic context; gestures, on their view, are semantically complete expressions like proper names, which pick out a particular referent regardless of which context they are used in.

The basic thought here is that two physically indiscernible acts of pointing can count as different gestures. So, one gesture might be a *pointing at Bill*, while another, involving what appear to be exactly the same bodily motions, might be a *pointing at Bill’s shirt*. According to Stojnić et al., it is a pre-semantic matter, a matter of disambiguation, to determine which gesture some ambiguous physical motion is in fact an instance of (2017, pg. 530).

We have no wish to quarrel with this way of individuating gestures. It is worth noting, however, that there is a fairly well-developed alternative: a theorist might agree with Stojnić et al. that gestures deserve to be counted as full-blown linguistic expressions, but disagree about a gesture’s contributing a determinate object independent of context. Kaplan (1977) himself took gestures to be underdetermined expressions that denote ranges of possible referents rather than any specific individual or property. In fact, one of the co-authors of both Stojnić et al. (2013) and Stojnić et al. (2017) previously argued for just this sort of picture (specifically, in Lascarides and Stone 2009). So, we find it curious that no argument is offered for the much stronger metaphysics of gestures that Stojnić et al. now endorse. These metaphysical-semantic assumptions are crucial to the smooth running of the semantic machinery the authors rely on, yet they go entirely unsupported.

While we are ourselves inclined to follow Kaplan in thinking of gestures more like definite descriptions than proper names, for the sake of argument let us simply accept Stojnić et al.’s treatment. On their picture, there is no question of why a *pointing at Bill* has Bill as its target; that is simply part of what it is to be a *pointing at Bill*.

Instead, the relevant question becomes: what makes it the case that such-and-such a bodily motion (something that can be common between two gestures, on this way of individuating things) counts, in the relevant context, as a *pointing at x*? In other words, what specifically makes this particular bodily motion into a gesture? As far as we can tell, any reasonable answer to this question is going to have to appeal to extra-linguistic context.

Consider a context in which the speaker has uttered (1) while pointing in such a way that—holding fixed her location, surroundings, and the physical form of her gesture—it is metaphysically possible for her pointing to instantiate either a *pointing at the V-notch couloir*, a *pointing at the U-notch couloir*, or a *pointing at the unnamed couloir to the looker's right of the U-notch at 13,658'*. Now we are faced with three different interpretive options for the logical form of the speaker's utterance of (1):

- (5) $\langle \pi 0 U\text{-}notch \rangle; [is.the.U.notch.couloir(@)]$
- (7) $\langle \pi 0 V\text{-}notch \rangle; [is.the.U.notch.couloir(@)]$
- (8) $\langle \pi 0 couloir\ at\ 13,658' \rangle; [is.the.U.notch.couloir(@)]$

We take it that the natural thing to say here is that the speaker's referential intentions determine one from among the candidates as the proper disambiguation of (1).¹⁸ That line is blocked for Stojnić et al., however, since it would entail that the extra-linguistic context plays a meaning-determining role with respect to (1). Similarly, Stojnić et al. cannot appeal to either salience or to all-things-considered judgments rendered on the basis of various aspects of the extra-linguistic context.¹⁹

In fact, things are even worse than this. As we have presented the example, the three possible interpretive options were that the speaker's gesture targets the U-notch couloir, the V-notch couloir, or the unnamed couloir to the right of the U-notch at 13,658'. The idea that the speaker's gesture might target, say, Barack Obama, was not considered. But why? The natural explanation would seem to run: Barack Obama is not present in the extra-linguistic context, and thus cannot be targeted by the speaker's gesture. One cannot gesture at what is not there.²⁰

¹⁸ Compare Kaplan (1978), Reimer (1991, 1992), and Michaelson (2013).

¹⁹ See Mount (2008) for the former and Wettstein (1984) and Gauker (2008) for the latter.

²⁰ One might attempt to challenge this claim on the basis of so-called deferred reference, such as the pointing at a picture of Carnap, saying 'he', and purportedly referring to Carnap himself. We are unsure whether to treat such reference as a semantic or a pragmatic phenomenon (cf. Nunberg 1995). Regardless, our principle can be weakened to say something like: one cannot gesture at what is neither there nor represented, in context, by something which is there. This weakened principle should suffice for our purposes, as would any number of yet weaker alternatives; although Carnap is represented by the picture of him, no feature of the scene from the Palisade Glacier can plausibly be treated as a representation of Obama.

Stojnić et al. cannot endorse such an explanation, however. The presence or absence of an object is a fact about extra-linguistic context, and according to their theory, extra-linguistic context plays no role at all in determining the meaning of a demonstrative. So, according to Stojnić et al., extra-linguistic context cannot constrain the possible interpretations of the physical form of a gesture—for that would be to allow that extra-linguistic context partially determine the meaning of a demonstrative.

We cannot discern any direct response to this problem of disambiguation in Stojnić et al.’s work. In their most detailed discussion of the issue, they compare the kind of disambiguation that demonstratives require, on their view, with the kind of disambiguation required in cases of lexical ambiguity:

For instance, if you say ‘I am at the bank’ while standing at the river-bank, this very fact might serve as a cue towards disambiguating one way; if you say ‘I put some money in the bank’, the plausibility of one content over another might serve as a cue that prompts a different interpretation. (Stojnić et al., 2017, note 23, pg. 529)

We find this analogy puzzling. We expect that most linguists and philosophers will say that cases involving lexical ambiguity are cases in which the phonological form of a sentence underdetermines its logical form. In this regard, there is a clear parallel with Stojnić et al.’s proposal. Where lexical ambiguity is concerned, however, we imagine that everyone who thinks that there are facts of the matter about which LF is expressed by a sentence on an occasion will agree that those facts are determined by some concrete features of the speech situation, like facts about the speaker’s intentions, or facts about what a reasonable hearer would assume the speaker meant. We take lexical ambiguity, then, to show that constituents that are not themselves semantically sensitive to context can nevertheless be implicated in the production of contextual effects at a higher level of description.²¹

Even on Stojnić et al.’s description, extra-linguistic context plays a key role in distinguishing between the possible LFs that could be associated with a sentence formed

²¹Stojnić et al. continue: ‘[W]hile we recognize that disambiguation *can* exploit a broad range of epistemic cues, we point out that, in the usual case, it exploits a set of precompiled solutions that obviate the need for open-ended reasoning about the speakers’ mental states, that would require the interlocutors to construct a broad range of potential interpretations on the fly. In most cases, the cues single out the correct interpretation from a set of possible ones the speakers know in advance’ (2017, note 23, pg. 529). But this seems merely to change the subject from the question of what, metaphysically, serves to disambiguate a sentence like (1) to the question of how listeners typically recover that disambiguation. Metaphysics has no need for precompiled solutions; it deals in facts, not reasoning. On the other hand, listeners, trying to identify what the speaker is talking about, might well make use of such solutions. Perhaps what Stojnić et al. mean is that there is a metaphysical correlate of these ‘solutions’. But it seems to us that such correlates will need to appeal to extra-linguistic context if they are to identify objects and properties present in those contexts as referents. So, once more, the suggestion fails to deliver what has been advertised.

from an ambiguous expression. It is the fact that the speaker is standing at a river-bank, in their example, that favors one interpretation of ‘bank’ over the other. How should we reconcile this with their hostility to the idea that context plays a meaning-determining role? One strategy might be to take the authors to be endorsing a view that posits a gap between the resources available to help interpreters in arriving at referential hypotheses and the resources that in fact determine reference.²²

There is clearly logical space open for a kind of error theory along these lines, on which interpreters use context to help them form hypotheses about what ambiguous sentences mean, despite the fact that in the background, as it were, the real meanings are fixed without any contextual input. We think such a view would be *prima facie* unattractive, however, as it threatens to make meanings epiphenomenal. In other words, on this sort of theory, the meanings of sentences like (1) might turn out to be completely detached from what speakers can expect to succeed in conveying by means of them. In fact, the meanings of such sentences could turn out to be something that one will *never* be able to convey by means of an utterance of that sentence. This strikes us as worrisome. Although we do not intend to defend a particular claim about which features of a context determine the reference of a demonstrative (indeed, we have doubts about whether there can be a single correct theory in this connection), and although we take it that both speakers and listeners are surely capable of being mistaken about what an ambiguous sentence means in a context, we would hesitate to endorse a theory that allows that meaning in a context can systematically come apart from what sentences are typically used in such a context to communicate.

Even leaving aside these worries, however, positing a systematic gap between the referents of demonstrative expressions and the objects people take to be the referents does nothing to advance us towards an actual theory of disambiguation—something that we still stand in need of if we are to fully assess the viability of Stojnić et al.’s proposed semantics and metasemantics for demonstratives. Effectively, by treating gestures as determinate linguistic items, Stojnić et al. have introduced a new and highly ambiguous item into the lexicon. That, in turn, makes the need for a theory of disambiguation all the more pressing for them, particularly if we are supposed to be able to evaluate their proposal in part on the basis of the predictions it makes.

In short, then, we take Stojnić et al.’s semantics and metasemantics for sentences like (1), sentences containing a demonstrative and accompanied by a pointing gesture, to turn the question ‘What determines the reference of a demonstrative (accompanied by a pointing gesture)?’ into the question ‘What determines the target of the accom-

²²We think the idea that such a gap looms in at least some cases is unobjectionable. That differences can arise between what an interpreter takes to have been expressed and what was actually expressed is presumably what allows us to intentionally mislead people; I know that you will think I meant ϕ , because I know about your epistemic state, but in fact I stated ψ . For the record, an earlier time-slice of Ernie Lepore, i.e. the co-author of Fodor and Lepore (2004), agrees with us about this.

panying gesture?’ We have argued that all of the natural answers somewhere appeal to extra-linguistic context. Since Stojnić et al. fail to provide an alternative, we conclude that instead of offering a theory on which the reference of demonstratives is fixed independently of the extra-linguistic context, what they have offered is more accurately conceived of as a theory on which any explicit mention of extra-linguistic context has simply been suppressed.

3.2 Deixis without pointing

Stojnić et al.’s approach to cases in which the use of a demonstrative goes unaccompanied by any pointing gesture raises essentially the same worries as those which involve pointing. In many such gestureless cases, the authors appeal to discourse relations to fix the semantic values of demonstrative expressions. However, even if we grant that discourse relations might sometimes have a role to play in resolving questions of reference, as far as we can tell, extra-linguistic context still ends up doing the heavy lifting. This comes out particularly clearly in connection with the most prominent example of deixis without pointing from Stojnić et al. (2013).

In that example, (9) is uttered by someone watching a cooking show in which an omelet is being fried in a pan:

(9) That’s an omelet.

Intuitively, the omelet in question should turn out to be the referent of the demonstrative. But here, as Stojnić et al. set up the case, there is no pointing gesture. This means that the strategy we considered earlier, of using a π update to make a certain object prominent (and thus available to serve as an antecedent for ‘@’) will not work.

Instead of π , then, it might seem more promising for Stojnić et al. to appeal to their σ update, which requires no accompanying demonstration. On this strategy, they might offer an LF along the following lines:

(10) $\langle \sigma 0 s_0 \rangle; [is.an.omelet(@)]$

On the face of things, this LF seems like it should do the right sort of work. Recall that the σ operation puts the central individual from situation s_0 into the 0-slot of the sequence representing the linguistic context. Supposing that the central individual from s_0 is an omelet, this maneuver will deliver the intuitive truth conditions for an utterance of (9).

The maneuver, however, raises the following question: what makes the omelet the ‘central individual’ in s_0 (which might be interpreted as, *inter alia*, the speech situ-

ation, the scene viewed on the television, or the scene that is at issue in some conversation)? As we noted earlier, Stojnić et al. never offer a straightforward answer to this question. Some obvious possibilities might include the omelet’s being maximally salient or being at the center of the speaker’s attention. But neither of these options will do for them, since both rely on an obvious appeal to the extra-linguistic context.

Is there another avenue by which Stojnić et al. might secure the referent for the demonstrative from (9)? Given how prominently their theory features discourse relations, a piece of theoretical machinery they frequently use to mimic the kind of work gestures do,²³ and given the fact that the formalizations they apply in cases of deixis without pointing all feature discourse relations, it would be natural to think that such relations must do substantial work in determining the referents of demonstratives in examples like (9).

Crucially, however, the particular discourse relation Stojnić et al. apply to the paradigmatic omelet case is not one that updates the attentional register. To analyze (9), Stojnić et al. (2013, pg. 518) offer the following LF:

$$(11) \quad \langle \alpha 0 \rangle; [Summary(s_0, e_0)]; \langle \sigma 1 s_0 \rangle; \langle \pi 1 @that \rangle; [omelet(e_0, x_1)]^{24}$$

This LF is substantially more complicated than any we have considered so far. This is because above, we presented LFs in terms of the formalism of Stojnić et al. (2017), which simplifies the system from Stojnić et al. (2013) in several respects. Since the fullest explanation Stojnić et al. offer of deixis without pointing is in their (2013), however, we turn here to the full complexity of that system in order to see whether it can offer an answer to our question about reference.

In the 2013 system, Stojnić et al. assume that predicates like *being an omelet* pick out relations between events (or, more technically, and following Kratzer 2002, *eventualities*) and objects, locations, etc.²⁵ Discourse relations pick out relations between events, in this case s_0 and e_0 . Some discourse relations update the attention register in various ways, while others leave the register unchanged, imposing other semantic or pragmatic constraints on the discourse.²⁶

²³In their (2017) at pg. 539, for example, the authors claim that the *Elaboration* relation systematically puts the direct object of a sentence into the top spot in the attention register. Other relations, like *Narrative*, put the subject of the sentence into that spot instead. Such discourse relations serve to make certain objects prominent (in the sense of prominence their theory is based on), and thus available to serve as antecedents for demonstratives. A sentence that encodes one of these relations will constrain the possible interpretations of anaphora in sentences to follow, without leaving any role to be played by the extra-linguistic context.

²⁴So far as we can tell, this LF involves a redundant π -update; we will essentially ignore that update in our discussion below.

²⁵This is why the predicate *omelet* takes both a situation, e_0 , and an individual, x_1 as arguments.

²⁶Compare note 23 above.

The contribution made by the discourse relation Stojnić et al. offer in (11), *Summary*, is characterized by them as follows:

Semantically, e_0 must be part of s_0 . Following Kratzer (2002), this entails that the information in the accompanying sentence, which is fleshed out in terms of constraints on e_0 , all winds up true in s_0 . Pragmatically, $Summary(s_0, e_0)$ holds only if the information the speaker uses to characterize e_0 provides a good answer about “what’s happening” in s_0 . (Stojnić et al., 2013, p. 517)

Whatever claim *Summary* might have to being an important element of our theoretical toolkit, it clearly does not provide an answer to the question of why one omelet, as opposed to some other, or as opposed to any other thing at all, should end up counting as the referent of the demonstrative from (9) when the sentence is used on a particular occasion.²⁷

If gestures do not fix the referent of our demonstrative, however, and if *Summary* doesn’t either, then what does? We suspect that Stojnić et al. will say: ‘The linguistic context by itself. When there is no pointing, and when discourse relations do not shape the attention register, then a demonstrative will just pick out whichever individual is in the top slot of the context’.²⁸ By our lights, however, this is an unsatisfying answer. We see no clear way of determining *which* individual will be in the top slot of any particular linguistic context updated as per (11). Nearly any Heim context will support any number of different possible summaries, each with a different object at the center.

To bring this problem into even sharper relief, notice, notice that the first operation specified in (11) is $\langle \alpha 0 \rangle$. This operation works roughly like a Heim-style indefinite: it puts an ‘unspecified entity’, e_0 , into the 0-slot of the attention register.²⁹ *Summary* requires that this event be a part of a broader event, s_0 , and that it in some sense provide a summary thereof.

But what events, exactly, should count as s_0 and e_0 ? At first glance, it might seem like different answers to this question will provide different referents for the demonstrative from (9). If that demonstrative ends up picking out the central individual from e_0 , then treating different events as e_0 will yield different objects as the referent; the center of my omelet cooking is one omelet, and the center of your omelet

²⁷In personal communication, Stojnić et al. have expressed agreement with us on this point; *Summary*, they say, is not meant to play a reference-fixing role. We are not sure what to make of their claim that ‘[I]t is the effect of a *Summary* that the central entity of a situation the summary is about is rendered most prominent.’ (Stojnić et al., 2013, pg. 518)

²⁸In the Stojnić et al. (2013) formalism, this will be the 1-slot, as the 0-slot is occupied by an event.

²⁹This is why, in the Stojnić et al. (2013) formalism, the top-ranked individual in an attentional context must occupy not the 0-slot, but the 1-slot; otherwise we would lose track of the relevant event.

cooking is another. So how, when we set about evaluating a particular utterance, should we know which situation to treat as s_0 , and which as e_0 ?

In point of fact, Stojnić et al. have a way of dodging this question. Recall that, on their analysis of (9), namely (11), each instance of s_0 and e_0 is embedded under the existential quantifier associated with α . So in the context of this LF, s_0 and e_0 are not names for particular situations; rather, s_0 and e_0 are variables that range over events.³⁰ If an utterance of (9) expresses the LF given in (11), then it will be true with regard to any variable assignment that maps s_0 and e_0 to a pair of events in which the second summarizes the first and in which the central object in the second event is an omelet. So, on this analysis, demonstrative sentences like ‘that is F ’ which are not accompanied by a gesture turn out to have the form of existentially quantified sentences: they will be true with regard to every assignment that returns a pair of events such that the second is a summary of the first and which is such that the central object of the second is an F .

This audacious proposal reminds us of Russell (1905)’s line with regard to proper names. Stojnić et al. ‘solve’ the problem of having to pair a Kaplan context up with the referent of a demonstrative by saying that demonstratives, when unaccompanied by gestures, do not in fact refer. To repeat: deixis without pointing, according to Stojnić et al., does not involve reference. Rather, non-gestural deictic sentences, like (9), are to be understood as truth conditionally equivalent to rather complicated descriptive sentences.

As with Russell’s theory, this theory not only counterintuitively claims that a certain class of intuitively referential terms does not really involve reference at all, it also runs into trouble where folk intuitions about truth conditions are concerned. Suppose you utter (9) while we are watching Julia Child cook an omelet on TV. We would contend that you have said something true. The theory described by Stojnić et al. correctly predicts this. On that theory, your utterance is analyzed using (11), and since there is a pair of situations that meets the constraints that LF imposes, the theory predicts that you have said something true. Now suppose, however, that we are standing on the street and there are no omelets anywhere nearby. You say (9). Intuitively, you say something false or otherwise defective. But, on the theory described by Stojnić et al., this is not the result we get. Somewhere in the universe there is almost certainly a pair of events such that one summarizes the other and has an omelet at its center. So, their account is very likely to predict that you have said something true.

This strikes us as an unacceptable result. We take it to be a minimum condition on the adequacy of any story about demonstratives that it make what you say in the

³⁰Stojnić et al. (p.c.) confirm that this was their intended reading of the sentence, and that the point of adding the existential quantifier to the front of the LF was indeed to avoid the sort of worry raised in the last paragraph.

case just described false. In order to get that result out of Stojnić et al.’s theory, we would need a way to rule out all of the possible evaluations of (9) that involve pairs of omelet-involving situations that are nowhere near the place and time of the speech act in question. We see no way of pulling off that trick without invoking the place and time of the speech act. Regardless of whether you build domain restriction into the quantifier associated with the α operation or limit peoples’ judgments about truth to some subset of the model, all of the available avenues would appear to run through the Kaplan context.

The proposal Stojnić et al. offer also seems to conflict with their own claim that utterances of sentences like (9) are ‘situated utterances’ (Stojnić et al., 2013, pp. 502, 505, 517–18). Informally, they often write as though there is a relevant situation for the evaluation of the utterance: either the one in which the speech act occurs or else one which is somehow made clear by that speech act, as when we are trying to interpret a sentence uttered on television. But we can see no way of squaring this claim with the LF that Stojnić et al. propose to associate with utterances like (9). The problem is that this LF embeds s_0 under what is effectively an existential quantifier—since it is anaphorically linked to e_0 , which is explicitly quantified over, via the *Summary* relation—meaning that there is no way for it to be matched to any one situation in particular. In other words, there is no sense in which the LF reflects the ‘situated-ness’ of the utterance.

Of course, Stojnić et al. could modify their proposed LF, dropping the quantifier and treating terms like s_0 as meta-language names for particular events. But we fail to see how that would amount to anything other than a concession that the truth conditions associated with a sentence like (9) depend on the extra-linguistic context in which (9) is uttered or evaluated. That, however, is precisely sort of view that Stojnić et al. promised to break with—for this is just to admit that the meanings of demonstratives are context-sensitive in precisely the way that philosophers have long taken them to be.³¹

To underscore the strangeness of Stojnić et al.’s proposal, consider a case in which two omelets are being prepared side-by-side. One is a vegetarian omelet, the other filled with ham. Imagine that a vegetarian speaker says to her listener, who knows that she is a vegetarian:

(12) That’s a lovely-looking omelet.

³¹If all that Stojnić et al. mean by claiming that demonstratives are not sensitive to context is that no element corresponding to the contribution made by the demonstrative to the LF associated with a demonstrative sentence is sensitive to permutations of the context parameter, then their original claim becomes significantly less striking. Indeed, as we understand the state of play in the literature, this is now a very prominent view, if not the standard one: context resolution takes place not in the semantics, on many leading views, but in the post-semantics. Stojnić et al.’s proposal, on this reading, would differ in locating the effects of context in the pre-semantics instead, but this would not be the radical revision we took the authors to have been advancing.

We take it that it will be extremely natural for the listener to take the speaker to be talking about the vegetarian omelet—though, obviously, there are situations in which that might be overridden. Absent a gesture from the speaker, however, there is no way for Stojnić et al. to explain that natural interpretation in terms of the LF to be associated with (12). The problem isn't that the relevant LF will associate the demonstrative here with some other referent. Rather, the problem is that, on Stojnić et al.'s reading, the LF associated with (12) is purely quantificational, and lacking in any referential properties at all. So any natural interpretation of the utterance will actually amount to a misunderstanding of its logical form.

We have focused here on the most prominent example of deixis without pointing from Stojnić et al. (2013). But lest the reader think that our focus has been somehow uncharitable, let us consider another of their examples. Imagine that someone utters the following sentence while standing in front of a telescope that is aimed at Jupiter:

(13) That is Jupiter.

Stojnić et al. want their theory to associate the following update with (13):

(14) $\langle \alpha 0 \rangle$; [*Summary*(s_0, e_0)]; $\langle \sigma 1 s_0 \rangle$; $\langle \pi 1 @that \rangle$; [*jupiter*(e_0, x_1)]

What's being summarized here, according to Stojnić et al. (2013), is 'a scene viewed through the telescope' (pg. 519).³² So the truth or falsehood of (13) will depend on whether there is a pair of events, one of which summarizes the other and has Jupiter as its central object. For the purpose of considering this example, we are going to assume that Stojnić et al. can offer us some story about how the domain of quantification for α is to be restricted to scenes viewed through the telescope present in the actual speech situation. Even granting this, we find the proposed analysis problematic in several ways.

First of all, we think it is important to stress that, as a matter of empirical fact, it will only very rarely be the case that people who look through telescopes, microscopes, and similar devices converge on the same object. Learning to identify objects in a microscope takes significant practice, and except where the most powerful telescopes and the most obvious targets are involved, the most common refrain at astronomy demonstrations is 'what am I supposed to be looking at here?'³³ In other

³²We leave aside the question of whether a scene can legitimately count as an event (or even an 'eventative'). Clearly, one can construct a corresponding event easily enough: an event of someone's *looking* through the telescope. We take it that this must be what Stojnić et al. have in mind.

³³There has, in fact, been a great deal of work in the philosophy of science aimed at trying to understand how it is that we see through these sorts of devices. No one, to the best of our knowledge, takes things to be nearly so simple as Stojnić et al. assume. For a helpful discussion of the sorts of issues that arise with respect to microscopes, for instance, see Hacking (1981).

words, we reject the thought that it is appropriate to treat notions like ‘whatever is centrally imaged by a telescope’ as basic in a theory that aims to answer questions like: under what conditions are utterances of sentences like (13) true or false? Not only are we skeptical that there is a well-behaved property of *being what is centrally imaged in a telescope*, a theory like Stojnić et al.’s holds our overall understanding of what makes an utterance like (13) true or false hostage to our understanding of this property. This hardly strikes us as a desirable feature for a theory of demonstratives.

To illustrate the point, imagine a fuller description of the situation described above: one that specifies that, in addition to Jupiter’s being visible in the telescope, Ganymede, its largest moon, is visible as well. However, suppose that, although Ganymede is visible, it is mostly occluded—such that only an expert astronomer would be able to recognize that it is visible at all. In this case, we can imagine such an expert uttering either (13) or (15):

(13) That is Jupiter.

(15) That is Ganymede, Jupiter’s largest moon.

Either utterance seems perfectly acceptable to us. Now, for Stojnić et al. to successfully predict this, it looks like they’ll need to accept that Ganymede can count as being centrally imaged in the telescope in this situation. Since they don’t want to appeal to extra-linguistic context beyond the event of the imaging, we take it that it should not matter who is uttering (15). Regardless of this, there is *an* event that summarizes the scene viewed through the telescope with Ganymede at its center.

We find this prediction implausible. To our ear, when (15) is uttered by a novice who would be unable to look through the telescope and identify Ganymede, the utterance is at the very least infelicitous. Of course, we don’t mean to commit ourselves to anything like the claim that, for demonstratives to refer, the speaker must be occurrently perceiving her intended referent. We do, however, think that there is likely to be some potentially complex relationship between the discriminatory capacities of the speaker and the objects that we take her to be capable of referring to with demonstratives. Whatever this relationship is, the speaker—and hence the extra-linguistic context—looks to be an essential part of it.

Even if we set these worries aside for the sake of argument, we again see no way to make Stojnić et al.’s picture work without invoking the extra-linguistic context. Imagine a case in which the view through a nearby telescope is of a large, uniform, white circle against an undifferentiated black background. Is an utterance of (13) true or false in this context? Well, that presumably depends on what the telescope is actually imaging. The scene through the telescope might appear identical to any desired level of detail with the telescope either imaging Jupiter or a cleverly illuminated sheet of paper that happens (randomly) to have marks on it that look identical

to Jupiter’s surface, up to any desired level of detail. Intuitively, (13) is true in the first case but false in the second. But since the scene is visually indiscriminable in both cases, it would seem that truth and falsity must depend on the scene’s relational properties. As best we can tell, that’s just to say that the truth or falsity of (13) depends here on the extra-linguistic context.

4 Conclusion

In this essay we have argued that, while it is indeed possible to offer a semantic picture that exchanges the challenge of explaining indexicality or assignment-sensitivity for the challenge of explaining disambiguation, this by no means eliminates appeals to extra-linguistic context. On the contrary, even in their brief remarks on the nature of disambiguation, Stojnić et al. routinely appeal to extra-linguistic context. Absent further argument, we cannot see how moving appeals to extra-linguistic context into the ‘pre-semantics’ of one’s theory constitutes a real step forward in our understanding of the meanings of terms like pronouns and demonstratives.

Instead, we are inclined to think that, in trading questions like ‘On what does the reference or value of a use of e.g. a demonstrative depend?’ or ‘On what does the truth or falsity of a demonstrative sentence depend?’ for the question ‘What determines which of any number of unpronounced items should be used to represent the meaning of a demonstrative sentence at the level of logical form?’ we may have taken a step backwards in terms of the clarity of our inquiry. Moreover, we suspect that, if one could answer either of these first two questions in a satisfying way, it would be possible to implement the resulting theory, formally, at the level of pre-semantics, semantics, or metasemantics—however one might be inclined by other sorts of background considerations.

This does not mean that nothing is at stake in deciding whether, semantically speaking, demonstratives are highly ambiguous, as Stojnić et al. propose, or whether their meaning can be captured by a unified character, or whether they should be represented simply using free variables. In fact, we think that quite a lot is at stake here. For instance, one might think that Grice (1989)’s dictum that we should not multiply linguistic items beyond necessity is very much at issue.³⁴ Equally, one might worry that the sorts of languages described by Stojnić et al. might be unlearnable for finite beings like ourselves. We doubt that either of these issues will prove at all straightforward, and we have no ambitions of trying to settle either here.

Although this essay has been mostly critical, we would like to close on a constructive

³⁴We are ourselves unsure what the force of this ‘should’ this is supposed to be. At the end of the day, our best guess would be that this should be understood as an aesthetic principle.

note. Although we disagree with the details of the position Stojnić et al. describe, we are sympathetic to the thought that it may well be a mistake to appeal to extra-linguistic context in order to fix the reference of demonstratives. Both of us come from the West Coast tradition of semantics—a tradition which has long been associated with the quest to locate the one true metasemantics for demonstratives—but we increasingly find ourselves worrying that recent theories about how context determines demonstrative reference have become implausibly baroque. Perhaps, like us, Stojnić et al. have been moved by the niggling sense that there has to be a better way forward.

Consider a recent example from King (2013), who is modifying an earlier proposal in response to a bevy of arguments later to appear in Speaks (2016):³⁵

A speaker S's use δ of a demonstrative expression in context c has o as its semantic value iff 1. o is the object of S's controlling intention in using δ in c ; and 2. a competent, reasonable, attentive hearer H who knows the common ground of the conversation at the time S utters δ , and who has the properties attributed to the audience by the common ground at the time S utters δ would recognize that o is the object of S's controlling intention in using δ in c in the way S intends H to recognize her intention. (King 2013, pp. 300-01)

The notion of a 'controlling intention' here is a highly technical one, one which depends on understanding intentions on the model of plans (cf. Bratman 1987) and positing, further, that those plans are to be individuated hyperintentionally. Basically, reference to o succeeds if the speaker intends for her listener to identify o , specified *de dicto*, as her intended referent—and so long as a suitably-idealized version of the listener would in fact identify, via this very same *dictum*, this object as the speaker's intended referent. Otherwise, reference fails.

We have no desire to argue against the specifics of King's proposal here.³⁶ Rather, we want merely to suggest that, when faced with a proposal like this one, one reasonable reaction might well be to ask whether we should return to our starting assumptions to see whether we took a wrong turn somewhere. Stojnić et al. suggest that we went wrong in appealing to extra-linguistic context to determine the reference of demonstratives. Instead, we should appeal only to linguistic conventions. We take our arguments to have been directed against the second of these suggestions, not the first. Here is an alternative which is also compatible with the first: reject the claim that sentences containing demonstratives have canonical truth conditions, strictly speaking. Instead, we might suppose that demonstratives serve as some-

³⁵Even more confusingly, the earlier proposal appears in King (2014b), which finally went to press only after it had already been supplanted.

³⁶See, however, Michaelson (2013), Nowak (forthcoming), and Nowak and Michaelson (2018a,b) for such arguments.

thing like place-holders—unassigned variables, if you like—which get assigned to a values only when we focus in on some more specific questions about the relevant speech act.³⁷ Standard questions might include: what was the speaker trying to communicate? What would a reasonable listener have likely recovered in these circumstances? And how might we, as external observers knowing all the relevant facts, settle a bet on the truth or falsity of the utterance? While the answers to these questions will themselves require appeals to extra-linguistic context, those appeals will plausibly turn out to be more straightforward and direct than when we try to ask about reference *simpliciter*.

Our aim here is not to defend this suggestion.³⁸ Rather, what we hope to have clarified is that, if one is motivated to try to avoid appealing to the extra-linguistic context to determine *the* reference of a demonstrative, that may indeed prove possible. We have argued that it is not made possible, in any deep sense at least, by appealing to linguistic conventions. But one can also give up this assumption by rejecting the uniqueness presupposition that has driven so much of our inquiry into the nature of reference. Rather than offering a theory of how *the* reference of a demonstrative depends on certain aspects of the extra-linguistic context, one can aspire to offer a theory of how the *sort of reference relevant to understanding successful communication or a listener's expected reactions or betting behavior or judgments of sincerity* might depend on extra-linguistic context. And then one can hope to start to map out what relationships there might be, if any, between these different types of reference.

Given the sorts of theories to which this traditional assumption has led us, we think that it is worth at least considering this sort of alternative. Not the superficial alternative of exchanging context-sensitivity for ambiguity, but the more radical alternative of embracing truth conditional indeterminacy as a more substantial fact of our linguistic lives than we might previously have thought.³⁹

³⁷This suggestion is, in fact, compatible with Stojnić et al.'s semantics with only minor modifications. What we would need is to modify the update functions so as to add variables to the top position(s) in the input context. So, for example, $\langle \pi 0 U\text{-notch} \rangle$ would become $\langle \pi 0 x_1 \rangle$. In keeping with the present proposal, the assignment of x_1 will depend on what question we are asking.

³⁸For a defense of the thesis, see Nowak (2016), Nowak (forthcoming), and Nowak and Michaelson (2018a).

³⁹Our thanks to Jonathan Cohen and Nathaniel Hansen for comments and discussion, and to Una Stojnić, Matthew Stone, and Ernie Lepore for their gracious and critical feedback. Thanks as well to the editors of *Linguistics and Philosophy* and to our two anonymous referees for their very helpful comments.

References

- Bach, K. (1992). Intentions and demonstrations. *Analysis*, 52(3):140–146.
- Barwise, J. and Perry, J. (1983). *Situations and Attitudes*. MIT Press, Cambridge, MA.
- Bertolet, R. (1980). Demonstrations and intentions. *Philosophical Studies*, 38(1):75–78.
- Bratman, M. (1987). *Intention, Plans, and Practical Reason*. Harvard University Press, Cambridge, MA.
- Caplan, B. (2003). Putting things in contexts. *Philosophical Review*, 112(2):191–214.
- Fodor, J. and Lepore, E. (2004). Out of context. *Proceedings and Addresses of the American Philosophical Association*, 78(2):77–94.
- Gauker, C. (2008). Zero tolerance for pragmatics. *Synthese*, 165(3):359–371.
- Grice, H. P. (1989). Further notes on logic and conversation. In *Studies in the Way of Words*, pages 41–57. Harvard University Press, Cambridge, MA.
- Hacking, I. (1981). Do we see through a microscope? *Pacific Philosophical Quarterly*, 62(4):305–322.
- Heim, I. (1982). *The Semantics of Definite and Indefinite Noun Phrases*. PhD thesis, University of Massachusetts, Amherst.
- Kaplan, D. (1977). Demonstratives. In Almog, J., Perry, J., and Wettstein, H., editors, *Themes from Kaplan*, pages 481–563. Oxford University Press, Oxford.
- Kaplan, D. (1978). Dthat. In French, P., Uehling, T., and Wettstein, H., editors, *Contemporary Perspectives in the Philosophy of Language*, pages 383–400. University of Minnesota Press, Minneapolis.
- Kaplan, D. (1989). Afterthoughts. In Almog, J., Perry, J., and Wettstein, H., editors, *Themes from Kaplan*, pages 565–614. Oxford University Press, Oxford.
- Kehler, A. (2002). *Coherence, reference, and the theory of grammar*. CSLI, Stanford.
- King, J. C. (2013). Supplementives, the coordination account, and conflicting intentions. *Philosophical Perspectives*, 27(1):288–311.
- King, J. C. (2014a). The metasemantics of contextual sensitivity. In Burgess, A. and Sherman, B., editors, *Metasemantics: New Essays on the Foundations of Meaning*, pages 97–118. Oxford University Press, Oxford.

- King, J. C. (2014b). Speaker intentions in context. *Noûs*, 48(2):219–237.
- Kratzer, A. (2002). Facts: particulars or information units. *Linguistics and Philosophy*, 25(5-6):655–670.
- Lakatos, I. (1974). Science and pseudoscience. In Vesey, G., editor, *Philosophy in the Open*. Open University Press.
- Lascarides, A. and Stone, M. (2009). A formal semantics of gesture. *Journal of Semantics*, 26(4):393–449.
- Locke, J. (1689). *An Essay Concerning Human Understanding*.
- McGinn, C. (1981). The mechanism of reference. *Synthese*, 49(2):157–186.
- Michaelson, E. (2013). *This and That: A Theory of Reference for Names, Demonstratives, and Things in Between*. PhD thesis, UCLA.
- Mount, A. (2008). Intentions, gestures, and salience in ordinary and deferred demonstrative reference. *Mind and Language*, 23(2):145–164.
- Neale, S. (2004). This, that and the other. In Reimer, M. and Bezuidenhout, A., editors, *Descriptions and Beyond*, pages 68–182. Oxford University Press, Oxford.
- Nowak, E. (2016). *Two Dogmas About Demonstratives*. PhD thesis, University of California, Berkeley.
- Nowak, E. (forthcoming). No context, no content, no problem. *Mind and Language*.
- Nowak, E. and Michaelson, E. (2018a). Meta-metaseantics, or the quest for the one true metaseantics. (Manuscript).
- Nowak, E. and Michaelson, E. (2018b). Who’s your ideal listener? (Manuscript).
- Nunberg, G. (1995). Transfers of meaning. *Journal of Semantics*, 12(2):109–132.
- Rabern, B. (2012a). Against the identification of assertoric content with compositional value. *Synthese*, 189(1):75–96.
- Rabern, B. (2012b). *Monsters and communication: the semantics of contextual shifting and sensitivity*. PhD thesis, The Australian National University, Canberra, Australia.
- Reimer, M. (1991). Demonstratives, demonstrations, and demonstrata. *Philosophical Studies*, 63(2):187–202.
- Reimer, M. (1992). Three views of demonstrative reference. *Synthese*, 93(3):373–402.
- Russell, B. (1903). *Principles of Mathematics*. Cambridge University Press, Cambridge.

- Russell, B. (1905). On denoting. *Mind*, 14:479–493.
- Salmon, N. (2002). Demonstrating and necessity. *Philosophical Review*, 111(4):497–537.
- Schiffer, S. (1981). Indexicals and the theory of reference. *Synthese*, 49(1):43–100.
- Speaks, J. (2016). Speaker and hearer in the character of demonstratives. *Mind*, 125(498):301–399.
- Speaks, J. (2017). A puzzle about demonstratives and semantic competence. *Philosophical Studies*, 174(3):709–734.
- Stojnić, U., Stone, M., and Lepore, E. (2013). Deixis (even without pointing). *Philosophical Perspectives*, 27(1):502–525.
- Stojnić, U., Stone, M., and Lepore, E. (2017). Discourse and logical form: pronouns, attention, and coherence. *Linguistics and Philosophy*, 40(5):519–547.
- Wettstein, H. K. (1984). How to bridge the gap between meaning and reference. *Synthese*, 58(1):63–84.
- Yalcin, S. (2014). Semantics and metasemantics in the context of generative grammar. In Burgess, A. and Sherman, B., editors, *Metasemantics*. Oxford University Press, Oxford.