Una Stojnić, *Context and Coherence: The Logic and Grammar of Prominence*. Oxford: Oxford University Press, 2021. ix + 219 pp.

Here is a picture of the relationship between natural-language semantics and pragmatics that many theorists would accept: semantics seeks to reverse-engineer the database of word meanings and composition rules by means of which we encode and decode the meanings of sentences. This is going pretty well, because our encoding-decoding algorithms are sufficiently discrete and well behaved that they can be treated as an autonomous grammar in this way. But this approach doesn't work in pragmatics, which is fueled not by a proprietary database of rules but by a mess of domain-general reasoning that resists computational tractability. Recent debates about the semantics-pragmatics boundary have tended to be about which parts of language-use can be treated as part of the grammatical algorithm and which must be consigned to the much blacker box of pragmatic inference.

In *Context and Coherence*, Una Stojnić argues that far more of our explanation of language-use should be relocated into the neatly rule-governed grammar than nearly anyone would have guessed. The implications of this view are far reaching, but the book revolves around a formal model of discourse that treats the process of assigning specific contents to pronouns and modals as case studies. If someone says, "He should run for office," what makes it the case that "He' refers to one person rather than another and that 'should' is being used to say something about this person's obligations as opposed to something about what is likely? Stojnić's formal model includes representations of contexts that evolve over the course of conversations and that determine the contents of the pronouns and modals that are uttered along the way. She takes the crucial components of contexts to be rankings of entities and possibilities according to their "prominence." The content of an occurrence of 'he' is the most prominent male at that moment in the conversation, and whether 'should' is being used as a deontic or epistemic modal depends on what sort of possibility is most prominent.

Stojnić innovates by grammaticalizing phenomena that others have taken to be part of the extralinguistic background to speech. Paralinguistic gestures, topic situations, intonational contours, and discourse-coherence relations are all inserted into logical form, which for Stojnić is a grammatical representation of a whole discourse rather than of an individual sentence. Each of these phenomena plays a distinctive role in manipulating the context's prominence rankings. Some of the empirical predictions of this model are also innovative, many of them organized around Stojnić's attempts to defend classical logic and propositional semantic contents from recent threats posed by dynamic semanticists and expressivists.

The most surprising aspect of Stojnić's view is her conventionalist interpretation of her model of discourse context. She argues that the context "is a running record of *linguistic* information that is contributed by discourse-internal, linguistic cues" and that it, "in turn, fully determines the interpretation by fixing the content expressed by the discourse" (10). As Stojnić puts it, "My conception of context is thus *linguistic*, rather than extra-linguistic" (10).

This is an iconoclastic idea. Nearly all recent work takes context to be a body of information that boils down to the interlocutors' beliefs or other extralinguistic states of mind. These states inform how hearers interpret speech, and since speakers anticipate this, they also inform how speakers design speech. Some would say that these states themselves fix the contents of context-sensitive expressions. Others would insist that the speaker's intentions do this but that the context is an important part of the evidence that a speaker expects their addressee to rely on when inferring their intentions. These mental states can be changed by linguistic utterances but also in other ways. Robert Stalnaker (1999: 86) illustrates this point by imagining a goat walking into the room where a conversation is happening. By way of perceptual rather than linguistic channels, this event allows interlocutors to felicitously refer to the goat with a pronoun or otherwise speak in a way that presupposes its presence and familiarity.

This looks like a problem for the idea that the context is fully under linguistic control. One way that Stojnić responds is by positing a coherence relation called *Summary*, which connects an utterance to a "perceptually present" topic situation and "makes the central entity in the situation described prominent" (174). Stojnić illustrates this idea with an example in which we see a video of Julia Child uttering (1) while cooking an omelet:

(1) That's your omelet.

Most theorists would think of this situation and its omelet as figuring in the extralinguistic beliefs or perceptual representations that lead interlocutors to triangulate on a referent. But Stojnić sticks the situation right into the logical form of (1) and has the *Summary* relation manipulate its "central entity" into prominence to serve as the referent of 'that'.

Stojnić uses a similar strategy to understand deictic pronouns that are accompanied by pointing gestures or other demonstrations. Most theorists think of demonstrations as extralinguistic evidence of the speaker's intentions. But Stojnić treats them as further elements in logical form, whose function is to raise the prominence of their demonstrata, thus effectively serving as grammatical antecedents of the pronouns that they accompany.

Consider also Stojnić's defense of modus tollens against Seth Yalcin's (2012) apparent counterexample, which revolves around a scenario in which the premises of (2) seem true but the conclusion false:

(2) If the marble is big, then it is likely red.
The marble is not likely red.
So, the marble is not big.

Stojnić argues that (2) is not a genuine instance of modus tollens because the two occurrences of the modal "likely" are interpreted relative to different prominent possibilities, which gives them different contents. Her explanation for this is that there are coherence relations in the logical form of (2) that make different possibilities prominent at different stages of the discourse. An *Elaboration* relation in the consequent of the first premise forces us to read the first instance of "likely" as saying what is likely in the initial context *modified* with the assumption that the marble is big, which has been raised by the first premise's antecedent. But the two premises are linked by the *Contrast* relation, which forces them to be interpreted relative to the same context. This means that the second instance of "likely" must be read as saying what is likely in the initial context, *unmodified*.

What are coherence relations, and why must we posit them in logical form just where Stojnić's explanations need them to be? Some previous theorists have thought of coherence relations as assumptions that interpreters make for broadly extralinguistic reasons, as part of a strategy for understanding how a speaker's communicative intentions fit with their broader plans. On this view, they flow from our tendency to attribute coherent plans to others as part of a general strategy for understanding behavior.

But again, Stojnić finds linguistic convention where others have found pragmatic inference. She argues that coherence relations are part of the grammatical machinery that we must learn to use when acquiring a language. Speakers use subtle but rule-governed signals—sometimes involving intonation or gesture—to indicate which relations go where in logical form. In some cases there are multiple options, in which case interpreters have to guess which arrangement makes the most sense. For example, although *Contrast* isn't grammatically required in (2), the two premises would turn out to be inconsistent without it.

Stojnić defends the conventionality of both coherence relations and demonstrations, as well as her strategy of putting them in logical form, by citing cross-linguistic variation in how they work (48, 69–71). But presence in logical form does not follow from conventionality.[1] And pure conventionality does not follow from a degree of cross-linguistic variation. Different languages sometimes include different grammatical affordances to serve universal communicative needs. For example, every language gives us clause-types that are specialized for making assertions, issuing directives, and asking questions,

1. Justin Khoo (2022) makes this point in detail in his excellent review of Stojnić's book.

respectively, but the syntactic implementation details of these clause-types vary greatly (Portner, Pak, and Zanuttini 2019). It does not follow that, for example, the existence of assertion itself is a matter of convention. By contrast, our most deeply conventional activities are conventional not just in their implementation but in their very ontology. Different societies differ not only in how one gets married but also in the nature and purposes of marriage. Surely the fact that we converse in coherent ways and track entities' prominence during conversation is conventional at most in its implementation details, and not in its raison d'être. To me, this suggests that our linguistic tools for signaling coherence relations may be tools for doing something that we would still be doing even if we didn't have specialized grammatical tools with which to do it.

It remains to be seen whether Stojnić's technical innovations will be widely taken up. Even if we do possess a surprisingly rich grammatical system of the kind that Stojnić describes, I am still tempted to reject Stojnić's view of it as an autonomous determiner of content that floats free of our extralinguistic psychology, and to instead think of it as a system for providing our addressees with finely crafted but still merely partial and defeasible evidence of our intentions.

Linguistic evidence is *partial* in that our addressees still need some nonlinguistic sources of information in order to figure out what we're saying. It seems to me that Stojnić's strategy of putting topic situations and demonstrations into logical form merely leaves us with a new version of this old problem. In virtue of what, after all, does a demonstration itself refer to one thing rather than another? What selects and individuates a given topic situation and its central entity? And what about cases in which there is no perceptually available topic situation? (For example, suppose that we witness a horrific car accident and, after several hours and some unrelated conversation, I see a disturbed look on your face and say, "Are you still thinking about that?" What determines the referent of "that" in this case?) Presumably, the answers to these questions aren't narrowly grammatical but have something to do with the interlocutors' nonlinguistic psychology.

When faced with the inescapable partiality of linguistic evidence, Stojnić sometimes says that this shows only that interpreters must sometimes disambiguate between admissible logical forms. But this strikes me as relabeling the problem, since in these cases disambiguating is a matter of selecting the referent of a demonstration or a topic situation and central entity from among many options—the same kinds of tasks, involving the same extralinguistic psychology, that Stojnić's opponents need to posit.

Linguistic evidence is *defeasible* because successful interpretation sometimes involves recognizing that you have been given misleading linguistic evidence of a speaker's intentions. Suppose that John intends to tell you that he and his husband practice monogamy, but utters, "My husband and I are monotonous." Or suppose that Sue wants you to know that she loves a certain Fabergé egg but sloppily points to the wrong thing when uttering "I love

that." What would you have to take John or Sue to have said in order for successful communication to happen in these cases? A conventionalist like Stojnić predicts that if someone carelessly invokes the wrong conventions when they speak, that's too bad, because it's the conventions rather than the intentions that set the terms of successful communication. But this seems like the wrong prediction. The best communicative outcome would be for you to somehow realize that the speakers have provided misleading evidence of what they were trying to say, and infer what they intended instead.

I am ultimately unconvinced by Stojnić's defense of conventionalism, then. Nonetheless, I do think that it is the most ingenious such defense that has been so far articulated. And even beyond this foundational issue, this book's innovative technical and empirical advancements make it required reading for anyone interested in the semantics and pragmatics of pronouns, modals, gesture, intonation, discourse coherence, and context sensitivity in general.

References

Khoo, Justin. 2022. Review of *Context and Coherence: The Logic and Grammar of Prominence*, by Una Stojnić. *Mind*, February 28. https://doi.org/10.1093/mind/fzab098.

Portner, Paul, Miok Pak, and Rafaella Zanuttini. 2019. "The Speaker-Addressee Relation in Imperatives." In *Proceedings of the Fourteenth Workshop on Altaic Formal Linguistics (WAFL14)*, edited by Tatiana Bondarenko, Colin Davis, Justin Colley, and Dmitry Privoznov. MIT Working Papers in Linguistics 90, 1–36. Cambridge, MA: MIT Working Papers in Linguistics.

Stalnaker, Robert. 1999. *Context and Content*. Oxford University Press.

Yalcin, Seth. 2012. "A Counterexample to Modus Tollens." *Journal of Philosophical Logic* 41, no. 6: 1001–24.

Daniel W. Harris
Hunter College and the Graduate Center, City University of New York

Philosophical Review, Vol. 133, No. 1, 2024
DOI 10.1215/00318108-10935394