

Magidor, Ofra, *Category Mistakes*, Oxford: Oxford UP 2013. Pp. xii + 171. Price £30.00 (hardcover), ISBN 978-0199572977.*

The toolkit of most philosophers features the notion of *category mistakes*. We can recognize clear cases (the three in “green ideas sleep furiously”) with quasi-perceptual alacrity. But I, for one, usually feel a tinge of self-conscious awkwardness when I deploy it. This has to do, I think, with the fact that we do not professionally acquire the notion together with a minimally adequate account – of the kind that usually comes, for instance, with introductions to the use/mention distinction: a philosophical explication that helps to make the phenomenon *prima facie* intelligible. If you are like me, this excellent short, clearly focused monograph will help you fill this gap in a few hours of engaging reading.

The first introductory chapter justifies the study of the phenomenon in itself but also instrumentally, in that it offers a conveniently isolated field in which views about foundational aspects of how linguistic representation works can be examined and tested. The chapter also provides helpful short summaries of the few previous discussions of the notion in philosophy, linguistics, and computer science. One learns here that the professional deficit noted above might originate in the lack of illumination provided on the nature of category mistakes by Ryle – the philosopher who coined the now standard term, gave us standard if unclear examples, and put the concept to ontological work in debates about the mind-body problem. Wisely, Magidor does not bring up Ryle’s famous example, the location of Oxford University; as a fairly standard metonym it is, I think, a borderline case that should not be used to introduce the topic, but left to be catalogued as *spoils for the victor*.

The first pages also frame the book's goal, which Magidor (citing examples such as Stalnaker on *possible worlds*, Williamson on *knowledge* and Campbell on *color*) declares not to be "to offer a precise analysis of the concept of 'category mistake', to provide informative necessary and sufficient conditions for being a category mistake, or to explain what makes category mistakes into a distinctive class of sentences" (3). The project is allegedly more modest: just "to explain what makes category mistakes infelicitous" (2). At the end of the book Magidor does provide some hypotheses that go beyond her declared goals, even if inconclusively – which is apt, because one might understandably find some tension in them. The philosophers she mentions to illustrate how similarly limited goals are compatible with a measure of philosophical illumination do not merely reject that one can provide conceptual analyses or *a priori* justified definitions of their topics; they also rule out illuminating *a posteriori* characterizations of their "real essences". Magidor, however, adopts "as a working hypothesis the assumption that it is possible to give a uniform account of category mistakes" (2). Moreover, she contends that category mistakes have a distinctive phenomenology, differentiating them from other sorts of linguistic infelicity; and she provides an account on which, in having these conscious impressions, we are being sensitive to common reasons for which category mistakes arise.

Chapter 2 presents and compellingly rejects the already *prima facie* suspect syntactic theory defended by Chomsky in *Aspects*. On Chomsky's proposal, lexical items come with "selectional features"; 'boy', say, is labeled *-abstract, +animate, +human*, and verbs and adjectives come with corresponding restrictions on the selectional features of their arguments. This is claimed to be part of syntax in that it allows us to classify the infelicity of, say, 'the boy which came earlier snores' as *ungrammatical*. After interestingly arguing that the view is not so easily refuted as one might have thought, Magidor shows it to be after all untenable. The arguments are related to the ones levelled against the much more popular "meaninglessness" view, discussed in chapter 3, on

which category mistakes make for meaningless sentences. The main worry about both views lies in considerations related to the Principle of Compositionality.

Influential compositional semantic theories assume a type-theoretic syntax (cf. Heim & Kratzer's (1998) presentation of the Montagovian approach). But, as Magidor shows (35), category mistakes may involve very specific categories (types of fruits, say) which hardly have the syntactic effects of types. They might also depend on empirical beliefs ('the priest is pregnant' does not sound infelicitous if you assume that women can be priests, or that a woman might recently have become a man, 40-1) or contextual facts ('John's best friend is pregnant' is infelicitous in a context in which John's best friend is assumed to be male but not if assumed to be female, 42). Other forceful considerations supporting the meaningfulness (and hence grammaticality) of category mistakes are that they have adequate translations (58), and that they can be embedded in true propositional attitude ascriptions ('Liza dreamt that John was pregnant', 38-9, 59-66).

A final point starts with the observation that many metaphors are generated by category mistakes: 'a pregnant poem'. Magidor's desire for compactness does not allow for any detailed discussion of a huge literature, but here as elsewhere she manages to muster compelling considerations in support of her points. She examines different theories of metaphor, divided into two groups. The first includes those allowing for the generation of metaphorical contents by meaningless phrases, such as the account on which they abbreviate similes ('a poem like someone pregnant'). These are compatible with the meaningless view; but all of them, she argues, supply erroneous accounts of metaphor. The second group includes the ones that require the metaphorical phrases to have a literal meaning, such as Davidson's theory on which metaphorical sentences produce their effects without having any other meaning than their literal

one. For her purposes, Magidor does not need to choose between them, because they are equally consistent with her point against the meaninglessness theory. These two chapters prove how sophisticated tools devised by philosophers of language and linguists allow us to put aside vague Wittgensteinian decrees stigmatizing discourses as “nonsense”.

I will discuss together the two final chapters, 4 and 5, for which I reserve my sparse critical remarks. They would have been even fewer had she applied here the strategy she uses in other parts of the book – when she deals with propositional attitudes, relative to their embedding category mistakes, or metaphors, vis-à-vis their generation by category mistakes. To wit: to establish that several plausible views on the issue allow for the claims she wants to make about her topic, and point out the weaknesses of the ones that are incompatible with them, while refraining from defending any one of them. Magidor convincingly defends in chapter 5 a presuppositional view of category mistakes. She could easily have followed this practice when it comes to presenting her positive view too, for different views on presuppositions are compatible with the proposal that category mistakes are presuppositional infelicities, and her declared aim was to explain this infelicity, as opposed to characterizing the nature of category mistakes. Chapter 4 should also have been rewritten in a more ecumenical way, instead of using it to defend the controversial, and, in my view, unconvincing views found there.

Presuppositions (say, *that someone robbed the bank* in the cleft construction ‘it was John who robbed the bank’) are pieces of background information that, although needing to be brought up for some purpose, are somehow taken for granted, and hence not the main point of the exchange – not what is asserted in literal uses of declarative sentences, questioned in yes-no questions, or requested in imperatives. Consistent with this background character, they *project* in that complex sentences such as negations, conjunctions, conditionals, or modals typically “inherit” those of the

simpler ones they embed: cf. ‘it was not John who robbed the bank’, ‘if Peter was at home, then it was John who robbed the bank’. However, there are special linguistic environments in which they are not “projected”; cp. ‘if someone robbed the bank, then it was John who did it’, ‘someone robbed the bank and it was John who did it’.

Some intuitive tests for presuppositions target this background, “taken for granted” character; thus, von Stechow (2004, 271) (citing Shannon) suggests a *hey, wait a minute* test, which contrasts a format for felicitous objections to presuppositions, “hey, wait a minute, I had no idea that someone robbed the bank” with the infelicity of a similar objection to assertoric content, “*hey, wait a minute, I had no idea that John robbed the bank”. Similarly infelicitous are “*I have no idea whether someone robbed the bank, but it was/wasn’t John who robbed the bank” and “*it was/wasn’t John who robbed the bank, and, what is more, someone robbed the bank”.

Magidor (131-146) convincingly shows that category mistakes meet these criteria for presuppositions. Take ‘two is green’. Using persuasive considerations that I cannot go into here (139-146), she argues that the presupposition that ‘being green’ triggers is weak – something like *being colored* or *being susceptible of having a color*. Check now:

- (1) This is/isn’t green.
- (2) Hey, wait a minute, I had no idea that that is susceptible of being colored.
- (3) *I have no idea whether this is susceptible of being colored, but it is/isn’t green.
- (4) *This is/isn’t green, and, what is more, it is susceptible of being colored.

Magidor’s main claim is thus well made. My qualms concern the foundational views on the nature of presuppositions that she espouses. An initial one has to do with her classification of these matters relative to the semantics/pragmatics divide; it is minor, but not because the issue is

merely terminological as it might be thought. She proposes to “take pragmatics to deal with those linguistic phenomena that go beyond the contributions of expressions to the truth-conditional content of sentences” (110). Now, although this view is not idiosyncratic, it is problematic. I take semantics to be a part of linguistic theory, whose main concern is the explanation of the phenomena that call for some form of the Principle of Compositionality: the fact that competent speakers are able to understand sentences they have never encountered before, that their understanding some involves understanding others, and so on. The evidence suggests that explanations of this kind interact in complex ways with syntax, and have proprietary cognitive underpinnings of the sort envisaged by Chomsky and his followers (cf. García-Carpintero 2012). Pragmatics includes instead explanations of the sort paradigmatically illustrated by Grice’s account of conversational implicatures, appealing to general principles of rationality.

The two illustrations that Magidor provides of what she counts as pragmatic exemplify what is wrong with it. One is the infelicity of ‘Today is Tuesday and today is not Tuesday’ (110). This is manifestly a pragmatic phenomenon in my sense, for from the standpoint of semantics the sentence is perfectly acceptable. The other is the infelicity of ‘Jane is happy but satisfied’ (110-1). Now, as several writers have convincingly argued (cf., e.g., Potts (2007)) conventional implicatures such as the one triggered by ‘but’ (that there is some conflict between the two relevant terms), on which the infelicity of the previous sentence depends, are semantic in my sense, and not just, nor crucially, because they are conventionally associated with the meaning of ‘but’. This is so mostly because contents of this kind compositionally interact with those of other expressions (for instance, when embedded under different operators) in complex ways that a semantic theory in my sense should acknowledge. Presuppositions are a semantic phenomenon for similar reasons, as we will see also according to the account that Magidor prefers.

Given their projection behavior – the fact in particular that they seem to be “entailments” both of a sentence and its negation – presuppositions were earlier thought to be a semantic phenomenon in Magidor’s sense: conditions for both the truth and the falsity of a sentence, which therefore render them neither true nor false when they fail. But this cannot be an adequate characterization. In the first place, there are “non-catastrophic” presupposition failures, which we have no problem in classifying as true (‘John is pregnant or he is not’) or false (‘I had breakfast with the king of France’) – some cases that Magidor discusses belong in this group, 135-7. Secondly, as we mentioned before, in some cases presuppositions are not projected to the whole sentence; hence, they might fail without generating truth-value gaps. In part for these reasons, other writers suggested pragmatic accounts. The most influential one is due to Stalnaker (1973). Stalnaker allows for the triggering of some presuppositions by the conventional meaning of lexical items, but it is clear that he favors pure pragmatic accounts, providing them whenever he can; and his account of projection is entirely pragmatic.

I have argued elsewhere (García-Carpintero 2013) that pragmatic accounts do not work, and that we need instead accounts of the sort of the “dynamic semantics” (DS) theories inspired by Stalnaker’s work developing from Heim (1983). An important consideration is that the cases of complex sentences in which presuppositions do not project require a compositional semantic treatment and cannot be dealt with pragmatically. Perhaps the assertion of a conjunction can be understood as a conjunction of assertions, and then these cases can be explained along Gricean lines in the way suggested by Stalnaker. But this cannot be applied to similar cases involving conditionals or disjunctions, as Schlenker (2009) notes. Now, precisely for this reason, Magidor does not opt for a pragmatic account of presuppositions such as Stalnaker’s (127, 130). But she does not like DS accounts either, mostly because it allows for truth-value gaps. A different problem she mentions (127-9) has been repeatedly voiced before. However, like Stokke (2013) I

do not see much in it. The objection is that the DS semantic rules for connectives are not explanatory, because there could be alternative rules that ascribe them a different projection behavior, while preserving their classical truth-conditions. But, so what? We can of course imagine different rules for the lexical items in a given natural language than the ones that an empirically adequate compositional semantics ascribes them. The objection just begs the question by prejudging a pragmatic account of presuppositional phenomena – one, moreover, of the kind that, Magidor concurs, does not seem to be forthcoming.

Magidor's discomfort with DS, as I said, lies in the fact that, although DS allows for true and false sentences with presupposition failures, it renders others truth-valueless. In chapter 4, Magidor examines what she calls *MBT* view, on which category mistakes are *meaningful but truth-valueless*. There she appeals to a variation on a specious argument by Williamson (86-9) to reject such views, and with them also presuppositional accounts relying on DS (130-1). She hence declares her preference for accounts of presuppositions compatible with bivalence such as Schlenker's (2009). As a follow up to the discussion above of the semantics/pragmatics divide, let me point out that, although Schlenker contends that his account follows "the spirit of Stalnaker's approach" (*ibid.*, 13), unlike Stalnaker's it is only pragmatic in Magidor's sense; exactly as in DS, the account assumes that presuppositions are calculated in a compositional way "locally", i.e., with respect to phrases, including predicates, that are proper parts of a full sentence. This is why Stalnaker (2010, 149-151) distances himself from it. Let me add that I only have a personal stake in rejecting pragmatic (in my sense) accounts of presuppositions. I feel some sympathy towards Schlenker's bivalent, non-dynamic proposal; this probably results from training-acquired prejudices, but it is certainly not based on Williamson's argument.

In closing I will now move on to this point. In chapter 4, Magidor admits that there might be meaningful sentences that are truth-valueless. She mentions sentences with empty referential terms, denotationless definite descriptions, or vague terms applied to borderline cases. Her contention, she says, is only that category mistakes are not among them (81). But it is unclear to me how accepting any truth-valueless sentence is compatible with her considerations. She says that, on an unproblematic conception, truth-value gaps result from meaningful sentences that fail to express propositions; and she persuasively argues that sentences involving category mistakes express partial propositions, truth-valueless only at some worlds. However, as far as I can tell the considerations she invokes could be used to make the same point about the cases she considers unproblematic. Take, for instance, ‘this is red’, in a context in which the demonstrative fails to refer. Some theorists may want to take it to express a Stalnakerian diagonal proposition that, although truth-valueless at the actual world, might be true or false at others.

Magidor then deploys the Williamsonian argument. Williamson’s argument is offered for utterances that say something; Magidor’s extension is given for propositions. The argument assumes that propositional versions of the Tarskian disquotational biconditionals are necessarily true, and derives a contradiction from this (87-8). In keeping with her desire for conciseness, she limits herself to following Williamson’s dismissal of rejections of such biconditionals, on the grounds that they are an essential component of our understanding of truth. But why should the reader grant any force to these admonitions, if they are left as being supported only by this rhetoric? Consider ‘necessarily, the proposition that this is red is true if and only if this is red’, with the demonstrative empty. Surely there are decent theoretical accounts of all relevant notions, including truth, compatible with the intuition that the biconditionals are true in core unproblematic cases, which count this one as on the one hand a proper instance, in that a proposition is expressed, and on the other false (having a truth-valueless right hand side and a

false left one). I have critically examined Williamson's argument elsewhere at some length (García-Carpintero, 2007); my point here is that it suffices to plug any problematic sentence in the Tarskian schemas to remove the force from Magidor's argument against MBT theories.

I find similarly weak Magidor's reply to the objection that any assignment of truth-value to category mistakes would be arbitrary: "a proponent of the view that category mistakes are genuinely truth-valued need not think of 'assigning truth-values' as a technical exercise, but rather as an attempt to describe the world as it is" (99). Suspiciously, however, when she has to assign some truth-value to sentences involving presupposition failure, including category mistakes, she just says that "presumably" they are false (126, 132). On what presumption, other than a stipulative one? She does not say. Plug any unproblematic sentence in the Tarskian biconditionals: necessarily, the proposition that Mallory reached the summit of Everest is true if and only if Mallory reached the summit of Everest. This is perfectly ok. We can thus take the claims that the proposition in question is true or false to be genuine attempts at describing the world as it is. They would be genuine, even if there is no way of deciding them. But what about the proposition that two is green, or that this is red, in the previously envisaged context in which the demonstrative fails? Can we think of any possible way, other than stipulation, of ever finding out, even if it is in fact not open to us? I do not think these questions betray any questionable verificationism, and hence I cannot find persuasive Magidor's claim that they are straightforward descriptions of the world that it is up to the world to settle.

References

- von Fintel, K. (2004): “Would You Believe It? The King of France is Back! (Presuppositions and Truth-Value Intuitions)”, in M Reimer & A Bezuidenhout (eds), *Descriptions and Beyond*, Oxford UP, Oxford, pp. 315-341.
- García-Carpintero, Manuel (2007): “Bivalence and What Is Said”, *Dialectica* 61, 167-90.
- García-Carpintero, M. (2012): “Foundational Semantics, I”, *Philosophy Compass* 7(6), 397-409.
- García-Carpintero, M. (2013): “Insinuating Information and Accommodating Presupposition”, in L. Goldstein (ed.), *Brevity*, Oxford University Press: Oxford, 214-229.
- Heim, I. (1983): “On the projection problem for presuppositions”, *Proceedings of the West Coast Conference on Linguistics* 2, 114-126.
- Heim, I., & Kratzer, A. (1998): *Semantics in Generative Grammar*, Oxford: Blackwell.
- Schlenker, P. (2009): “Local Contexts”, *Semantics & Pragmatics* 2, 1-78.
- Stalnaker, R. (1973): ‘Presuppositions’, *Journal of Philosophical Logic* 2, pp. 447-457.
- Potts, Christopher (2007): “Conventional Implicature: A Distinguished Class of Meanings”, G. Ramchand & Ch. Reiss (eds.), *Oxford Handbook of Linguistic Interfaces*. Oxford: Oxford U.P., 475-501.
- Stalnaker, R. (2010): “Responses to Stanley and Schlenker,” *Philosophical Studies* 151, pp. 143–157.
- Stokke, A. (2013): “*And* and *And**”, in L. Goldstein (ed.), *Brevity*, Oxford University Press: Oxford, 198-214.

Manuel García-Carpintero

LOGOS-Department of Logic, History and Philosophy of Science; University of Barcelona

* Thanks to Ofra Magidor and Teresa Marques for very helpful discussion of a previous version, and to Michael Maudsley for the grammatical revision.