

Title: Multi-Sentential Category Mistakes

Author: Paul Elbourne

Affiliation: Faculty of Philosophy
University of Oxford

Postal Address: Magdalen College
Oxford OX1 4AU

Telephone Number: 01865 276000

Email address: paul.elbourne@magd.ox.ac.uk

Acknowledgements It is an honour to be invited to comment on Magidor's groundbreaking book. I am grateful to Herman Cappelen for the invitation and to Ofra Magidor and Daniel Rothschild for valuable discussion of this and many other things.

Abstract

Magidor (2013) argued that category mistakes are infelicitous due to presupposition failure. The case for this position is strengthened by the consideration of a previously unnoted category of data, namely multi-sentence discourses in which category mistake phenomenology arises at the end of the last sentence, but arguably due to content contained in a previous sentence. This phenomenon is analysed in terms of the previous sentence giving rise to a presupposition that is shown to be false only in the last sentence.

Keywords: category mistakes, presupposition, ERP, N400

Multi-Sentential Category Mistakes

Paul Elbourne

June 2015

1 Introduction

The purpose of this paper is to present a novel argument in favour of Ofra Magidor's theory of the infelicity of category mistakes.

In her recent book, Magidor argues that category mistakes are infelicitous because they suffer from presupposition failure.¹ The presuppositions in question are supposed to be contributed by individual words ('presupposition triggers'), in the manner familiar from the literature on presupposition. Consider the following examples.

- (1) John stopped smoking.
- (2)
 - a. The number two is green.
 - b. My toothbrush is pregnant.
 - c. The chair is dreaming.
 - d. The theory of relativity is prime.

Just as the word 'stopped' is responsible for the presupposition that John used to smoke in (1), so the word 'green' in (2a) is supposed to introduce a presupposition to the effect that the denotation of the subject of the sentence is coloured; and the failure of this presupposition in all but the most outré contexts is supposed to be responsible for the sentence's status as a category mistake.² Similarly, in (2b) the presupposition trigger is claimed to be 'pregnant', and the presupposition is that the denotation of the subject can be pregnant; in (2c) the trigger is 'dreaming' and the presupposition is that the denotation of the subject has mental states; and in (2d) the trigger is 'prime' and the presupposition is that the denotation of the subject is either prime or composite.³

Magidor backs up her claim that presuppositions are involved by subjecting example (2a) to an array of tests for presupposition, including presupposition projection behaviour.⁴ For example, just as (3), like (1), still bears the presupposition that John used to smoke, so (4), like (2a), is also a category mistake.

- (3) John has not stopped smoking.
- (4) The number two is not green.

Magidor's claim is that (4) is a category mistake because 'green' introduces a presupposition, as before, and this presupposition projects through negation and becomes a presupposition of the whole sentence; (4) would work exactly like (3), then. This kind of similarity is shown to hold

1. Magidor, *Category Mistakes*, 111.

2. *Ibid.*, 131–2.

3. *Ibid.*, 140–44. Magidor admits freely that there is room for debate over the exact presuppositions contributed in each case but observes quite rightly that her theory does not stand or fall on the exact content of the presupposition of any one example.

4. *Ibid.*, 132–40.

over a wide variety of presupposition projection environments and constitutes an impressive case for Magidor's theory.

There is, however, another, previously unremarked, category of data with regard to which Magidor's theory fares well. I will describe it in the next section.

2 Multi-Sentential Category Mistakes

One characteristic of presupposition failure is that it is possible for a presupposition triggered in one sentence to be shown to be false only in a subsequent one. This means, of course, that the phenomenology of presupposition failure is experienced only after reading the subsequent sentence (or the relevant part of it). Consider the following examples, which use the presupposition triggers 'stop', 'the' and 'again'.⁵

- (5) John has stopped smoking. He never did smoke.
- (6) The Akond of Swat is a jolly old fellow. There is no Akond of Swat.
- (7) John has been swimming again. He never has before.

By analysing category mistakes as being infelicitous due to presupposition failure, Magidor predicts, then, that we will see analogues in the realm of category mistakes of the multi-sentential examples above. This prediction turns out to be correct, as we see with the following examples:

- (8) The number two has the property that John just mentioned. The property that John just mentioned is the property of being green.
- (9) a. The number two came first in the recent contest. The recent contest was the London Marathon.
b. The number two came first in the recent contest. The recent contest was the number popularity contest.

The distinctive phenomenology that Magidor uses to pick out category mistakes⁶ is clearly felt in the case of examples (8) and (9a). (If (9a) seems independently odd because of the talk of the number two coming first in something, which might be thought to have a slight anthropomorphizing flavour, it is nevertheless clear that (9a) is a whole lot odder than (9b).) And these examples readily lend themselves to analysis in terms of Magidor's theory. In the case of the first sentence of (8), we can say that 'has' triggers a presupposition to the effect that the number two is the kind of thing that could have the property that John just mentioned. (The suggested presupposition here is roughly modelled after the one suggested by Magidor for 'pregnant', which we saw above, and after some discussion on page 146 of her book.) Since this sentence is presented out of the blue, the hearer does not know yet what that property was and hence is not in a position to find the presupposition false. But after receiving more information about this property in the second sentence, the hearer finds the presupposition false and the characteristic category-mistake phenomenology ensues. Something very similar could be going on in (9a), where the predicate 'came first' would trigger a presupposition to the effect that the number two is the kind of thing that could come first in the recent contest. Upon hearing just the first sentence, the hearer is not in a position to know whether the presupposition is true (although suspicions might be aroused), and it is necessary to wait until the second sentence in order to find out that the presupposition is false, whereupon category mistake phenomenology is experienced. These examples, then, are plausibly very similar to (5)–(7), a fact that further bolsters support for Magidor's theory.

5. It is notable that discussions of presupposition generally assume without comment that the definite article is a presupposition trigger, even though Russellian analyses of it are popular elsewhere. For my own view, see Elbourne, *Definite Descriptions*.

6. Magidor, *Category Mistakes*, 1–2.

3 Consequences for Other Theories

Having said this, it is worth explicitly addressing the question of the consequences that these new data have for the rival theories that Magidor considers. They will turn out to present problems for all the rival theories.

3.1 The Syntactic Approach

The syntactic approach to the infelicity of category mistakes claims that category mistakes are syntactically ill-formed or ungrammatical.⁷ Let us adopt a charitable reading and assume that it predicts that at least one (as opposed to both) of the sentences in (8) is ungrammatical. It is clear, however, that neither of the sentences in (8) is ungrammatical. (Taken by themselves, they are not infelicitous in any way; and ungrammaticality is a property of individual sentences.) So the syntactic approach cannot be correct.⁸

3.2 The Meaninglessness View

The meaninglessness view maintains that category mistakes are meaningless.⁹ But it is perfectly obvious that all the sentences in (8) and (9a) are meaningful. (It is helpful in allaying doubts on this score to note that none of these sentences constitutes a category mistake by itself.) A friend of the meaninglessness view could conceivably maintain that although both sentences in (8) are meaningful, they do not make a meaningful unit when placed together in a discourse. But it hard to see any theoretical justification for saying that two meaningful sentences placed together do not make a meaningful discourse, especially when, as in this case, the sentences are bound together by an anaphoric link ('the property that John just mentioned... the property that John just mentioned'). So examples of the current kind constitute a grave objection to the meaninglessness view.

3.3 The MBT View

3.3.1 Introduction

That leaves the MBT view. This is the doctrine that category mistakes are meaningful but truth-valueless.¹⁰ Is this view consistent with (8) and similar data? If it is, or if it is to be minimally adapted so to deal with these data, it will presumably have to claim that at least one of the sentences in (8) is meaningful but truth-valueless. So which of the two sentences in (8) could be truth-valueless?

3.3.2 The second sentence as truth-valueless

Could it be the second, repeated here as (10)?

(10) The property that John just mentioned is the property of being green.

This is unlikely, since, intuitively, this sentence seems eminently susceptible of truth, provided that John did indeed just mention exactly one property. If this separate and here irrelevant pre-supposition is fulfilled, we just want to know whether the property in question is the property of being green. If so, the sentence is true. If not, it is false. Note that the contention that a crucial

7. Magidor, *Category Mistakes*, 25.

8. The problem that (8) and similar data raise for the syntactic approach is arguably a variety of the problem labelled 'interactions with context' by Magidor, *ibid.*, 42.

9. *Ibid.*, 43.

10. *Ibid.*, 81.

sentence is truth-valueless gains no traction in this case by any intuitive oddness that might be felt in the sentence taken by itself. One is tempted to leave the matter there.

However, a defender of the MBT view might point to the fact that the subject of (10) is a definite description and suggest that the number two finds its way into the interpretation of the subject by whatever means is used in adding extra content to incomplete definite descriptions. As is well known, Frege and Russell, followed by many others, maintained that, given a phrase of the form ‘the so-and-so’, there must be exactly one so-and-so for the containing sentence to be true and felicitous.¹¹ But Quine and Strawson noted that it is possible to use definite descriptions in true and felicitous sentences even when their descriptive content is not true of exactly one thing, as in (11).¹²

(11) The table is covered with books.

A large literature has grown up around this point, with one option commonly espoused being to claim that these *incomplete* descriptions are interpreted as providing more descriptive material than meets the eye.¹³ It could be maintained, then, that the occurrence of ‘the property that John just mentioned’ in the second sentence of (8) might be interpreted as if it were ‘the property that John just mentioned that the number two has’. (The occurrence of ‘that’ after ‘mentioned’ introduces a restrictive relative clause; it is not meant to be a complementizer introducing indirect speech.) So (8) would be equivalent to (12):

(12) The number two has the property that John just mentioned. The property John just mentioned that the number two has is the property of being green.

And this, it might be maintained, brings the number two and the property of being green together in one sentence (the second one) in such a way as to make it more like a conventional, one-sentence category mistake. The second sentence, then, would be truth-valueless on this view for whatever reason makes one-sentence category mistakes truth-valueless (or for a closely related reason).

But it is evident that we do not need to supply implicit content like *that the number two has* in order to interpret the second sentence of (8). For the definite description ‘the property that John just mentioned’ could very well be complete. Let us stipulate so; that is, let us stipulate that John has just mentioned exactly one property. Nothing changes about the category-mistake status of the example. But it is clear, now, that we do not need to add this implicit content in the second sentence. So the implicit content in question is optional at best in the second sentence. But the category-mistake phenomenology upon reading the second sentence is not optional. So the category-mistake phenomenology is not caused by the addition of this content.

It is also worth noting that we do not need to supply any such content in order to interpret the occurrence of ‘the property that John just mentioned’ in the *first* sentence of this example. If we suppose that this implicit content is present in the first sentence, we make it equivalent to (13):

(13) The number two has the property John just mentioned that the number two has.

(13) is quasi-tautologous in an interesting way: given that the second definite description has a referent that does indeed satisfy its descriptive content, the sentence has to be true. But the first sentence of (8) does not feel tautologous in any way. It does not include this implicit content, then. So this implicit content is not necessary in order to interpret the occurrence of the definite description in question in the first sentence; and so it is surely optional at best in the second sentence. The argument now proceeds as above.

11. Frege, “Über Sinn und Bedeutung”; Russell, “On Denoting.”

12. Quine, *Mathematical Logic*; Strawson, “On Referring.”

13. Sperber and Wilson, *Relevance*; Neale, *Descriptions*, and many others.

Likewise, if the subject definite description of the second sentence of (8) was interpreted by means of implicit content drawn from the first sentence, in such a way as to introduce the number two into the second sentence, which is the option we are now exploring, (8) would be equivalent to (14):

- (14) The number two has the property John just mentioned. The property John just mentioned that the number two has is the property of being green.

But (14) is slightly awkward in the following way: it seems to presuppose in the first sentence that there is exactly one property that John just mentioned. But then in the second sentence extra descriptive content is provided for ‘property John just mentioned’, as if we need to distinguish between various such properties. But (8) is not awkward in this way. So (8) is not equivalent to (14).

Our discussion so far has concentrated on examples involving overt definite descriptions. It is also notable that very similar examples can be made up that do not involve definite descriptions (or not overt ones, at least). Here is a pair involving proper names:

- (15) a. The number two came first in Mayfest. Mayfest is a philosophers’ drinking contest.
b. The number two came first in Mayfest. Mayfest is a number popularity contest.

Example (15a) produces category-mistake phenomenology, just as (8) and (9a) do. But there is little chance of smuggling the number two into the interpretation of the second sentence. While proper names have, of course, been analysed as covert definite descriptions, by Russell¹⁴ and many others, I am not aware of anyone who has taken it upon themselves to analyse them as *incomplete* definite descriptions. There are only so many burdens one can shoulder at once. Nor am I aware of any descriptivist theories of names that maintain that extra descriptive material can be added to names in any process similar to the completion of incomplete definite descriptions. Even on a descriptivist treatment of proper names, then, examples like (15a) will pose a problem for any version of the MBT view that tries to assimilate the second sentences of multi-sentential category mistakes to ordinary one-sentence category mistakes.

Here is one last example, in which the second sentence is, if anything, even more innocuous-seeming than those of the previous examples:

- (16) Everything John adores is green. John adores the number two, for example.

Again, this evokes category-mistake phenomenology. According to Magidor’s theory, the first sentence creates a presupposition to the effect that everything John adores is coloured; and the second one (assuming we believe it) shows this to be false. But it is difficult to see how the second sentence could be truth-valueless (assuming that John and the number two both exist). John either adores the number two or he does not.

3.3.3 The first sentence as truth-valueless

Let us turn to the first sentence of (8), repeated here:

- (17) The number two has the property that John just mentioned.

This sentence too might be claimed to be meaningful but truth-valueless.

As with the second sentence, it might be thought implausible on the face of it to claim that (17), on a suitable occasion of use, is truth-valueless. Assuming, once more, that there are such things as the number two and the property that John just mentioned, then it would seem that the former either has the latter (in which case the sentence is true) or it does not (which leads to falsity). But things are perhaps slightly more uncertain in this case. Given that the property

14. Russell, “On Denoting.”

that John just mentioned is, in fact, the property of being green, (17) in effect predicates of the number two the property of being green, and this is exactly the kind of thing that is claimed by the MBT view to give rise to lack of truth value. It is not clear, I would maintain, exactly how this kind of predication is supposed to produce this result; but it is worth granting this assumption provisionally in order to see if anything else can be said against the idea in this context.

Once more, then, we should investigate the possibility of somehow assimilating this sentence to common-or-garden, one-sentence category mistakes, since this might be thought to allow whatever mechanism prevents a truth value in the one-sentence cases to operate here too. I see two options. We could claim that the definite description ‘the property that John just mentioned’ is interpreted as *the property John just mentioned that the number two has*, by whatever means turns out to be involved in interpreting incomplete definite descriptions. This, of course, is just the strategy explored above for the second sentence. Or we could claim that something slightly more radical is going on, as outlined in the last paragraph: the sentence’s semantic value is recalculated, based on information gathered from the second sentence, with the property of being green put in the place, as it were, of the property that John just mentioned. I will explore these two options one at a time.

It can be seen, as I said, that the first option is basically a variant of the strategy used above in the case of the second sentence. And analogues of the objections to the earlier version of this strategy apply here too. Although it is perhaps possible that hearers go back to adjust the interpretation of definite descriptions in the first sentence after hearing the second one (but see below for doubts on this score), it is hard to see why this would be a necessity. This is particularly so given that, as stipulated, (8) does not use incomplete definite descriptions. So this postulated process of adjustment would be optional at best. But the category-mistake phenomenology is not optional. So the postulated process is not responsible for the category-mistake phenomenology.

Similarly, it is hard to see how this kind of strategy could make headway with examples like (15a) and (16) that do not use definite descriptions in the relevant places. I do not see any means for dealing with (15a) using the current strategy, for the reasons given above. As for (16), one might consider adding some implicit content to the descriptive material of the initial quantifier phrase, under the assumption that this option is available since quantifier domain restriction works like this, in such a way that the number two now makes an appearance there; this, once again, would be with a view to assimilating the current kind of case to the one-sentence category mistakes. It is hard to see how to do this, however, in any way that would not produce plainly unintuitive results. Certainly the most obvious option that springs to mind, using a property like *identical to the number two* to introduce this number into the restrictor, will not fly. For that would make (16) equivalent to (18):

- (18) Everything John adores that is identical to the number two is green. John adores the number two, for example.

It is obvious that ‘everything John adores’ is not equivalent, in (16), to ‘everything John adores that is identical to the number two’. So, once more, it is difficult to know how to proceed if one wishes to analyse these examples in accordance with the MBT view.

Let us turn, then, to the more radical second option mentioned above. The idea here is that we grant, for the sake of argument, the idea inherent in the MBT view that predicating the property of being green of the number two does not produce a truth value. We hypothesize that hearers of (8), after they have heard the second sentence and thus come to know that the property just mentioned by John is the property of being green, go back, as it were, and calculate a truth value for the first sentence based on this knowledge; and this means predicating the property of being green of the number two, which would, we are assuming, not lead to a truth value for the first sentence. (Recall that the MBT view has to say that one of the sentences in (8) does not have a truth value.) This, then, looks like a promising strategy for the MBT view. How can we evaluate it?

The answer, I believe, lies in noticing a particular point of contrast between this theory and Magidor's. The theory currently being explored claims that the hearer goes back and reprocesses the first sentence after hearing the second sentence. This will inevitably take a certain amount of time. Magidor's theory, on the other hand, uses Stalnaker's theory of presupposition and, in particular, his notion of a constantly updated common ground, with which presuppositions have to be consistent.¹⁵ According to this theory, then, the information contributed by the first sentence, to the effect that the number two has the property that John just mentioned, is added to the common ground and is part of the background of assumptions against which the second sentence is processed.¹⁶ When the hearer processes the second sentence and comes to know that the property that John just mentioned is the property of being green, the information that makes this odd is immediately to hand in the common ground, according to Magidor's theory. There is no need to undertake the relatively costly and slow business of going back and reprocessing the first sentence (assuming, indeed, that hearers will have memorized the first sentence). In short, we can expect that processing times for the second sentence will be much shorter if Magidor's theory is correct than they will be if the current version of the MBT view is correct.¹⁷

Remarkably, there is quite a lot of information in the psycholinguistics literature that bears on this question. While no-one has investigated the question using example (8), there has been enough investigation of relevantly similar sentences for us to be able to draw some preliminary conclusions. And the results do not look good for the MBT view.

The relevant evidence comes chiefly from neurolinguistic studies of the processing of multi-sentence discourses very like (8). The studies in question use electroencephalography, which is the measuring of electrical currents that flow across the scalp, caused by underlying neurological activity. In particular, they measure event-related brain potentials (ERPs), which can be defined as follows:¹⁸

An ERP is a patterned series of voltage deflections elicited by a critical stimulus, such as a tone or a word. It is obtained by averaging bits of raw electroencephalogram (EEG) data recorded at the scalp as the listener or reader encounters a large number of these stimuli.

To be more specific still, the studies in question measure an ERP response called the N400, which is 'a broad negative deflection of the ERP that starts 200–300 ms after a word has been presented auditorily or visually and peaks after approximately 400 ms'.¹⁹ It has been noted in numerous studies that the amplitude of the N400 is greater when the word that evokes it is unexpected or anomalous in context. The N400 would have a significantly greater amplitude after 'socks' in (19b), for example, than it would after 'sugar' in (19a).²⁰

- (19) a. I like my coffee with cream and sugar.
b. I like my coffee with cream and socks.

The stimuli used in the relevant experiments are not always category mistakes as opposed to anomalous in some other way. (I would not class (19b) as a category mistake, for example.) But

15. Magidor, *Category Mistakes*, 124. See Stalnaker, "Pragmatics"; Stalnaker, "Pragmatic Presuppositions."

16. In case any philosophical feathers are ruffled by my talk of sentence processing, I will note merely that by characterizing category mistakes in phenomenological terms and utilizing a Stalnakerian notion of presupposition, Magidor has already placed the discussion of category mistakes squarely on psychological turf. I agree with this placement.

17. Note, once more, that the MBT view needs a particular *sentence* to be without a truth value. Just adding a Stalnakerian theory of context to the MBT view will not help it, since there being a contradiction involving the common ground is just not the kind of thing that this theory claims is going on.

18. Van Burkum, "Understanding Sentences in Context," 376.

19. Lau, Phillips, and Poeppel, "A Cortical Network for Semantics," 920.

20. *Ibid.*, 921.

sometimes they are category mistakes. Nieuwland and van Berkum, for example, investigated the N400 response to animacy violations, with example sentences like those in (20).²¹

- (20) a. Once upon a time, a psychotherapist was consulted in her home office by a sailor with emotional problems.
b. Once upon a time, a psychotherapist was consulted in her home office by a yacht with emotional problems.

As one would expect, the N400 amplitude after ‘yacht’ in (20b) was significantly greater than that after ‘sailor’ in the otherwise identical (20a); and (20b) is a fine example of a category mistake. Similar results were obtained by Hagoort et al., who contrasted cases of world-knowledge violations, exemplified by (21b), and what they call ‘semantic violations’, which we can take to be category mistakes, exemplified by (21c). Both kinds of sentence were contrasted to similar true sentences, exemplified by (21a).²²

- (21) a. Dutch trains are yellow and very crowded.
b. Dutch trains are white and very crowded.
c. Dutch trains are sour and very crowded.

(All of the Dutch subjects in the experiment knew that all Dutch trains are yellow.) The N400 amplitudes after ‘white’ and ‘sour’ were both significantly greater than those after the correct ‘yellow’. Interestingly, the amplitudes after the category mistake ‘sour’ were greater than those after the merely incorrect ‘white’; but this difference, although statistically significant, was small; and the onset time of the N400 effects in the two cases were not significantly different. The impression we are left with from the neurolinguistics literature is that category mistakes and sentences that are implausible in other ways fall on a continuum and that the neurological reactions to them are basically the same in kind, differing only, as in the last experiment mentioned, in degree.²³

With this background in place, we are now in a position to appreciate the relevance of some electroencephalographic investigations of multi-sentence discourses that are relevantly similar to (8). Van Berkum, Hagoort and Brown used three-sentence discourses with a word in the final sentence that was anomalous (though not a category mistake in the examples we see) in the context of the discourse but fully acceptable in its own sentence in isolation.²⁴ An example is (22b), which was tested alongside the coherent (22a).²⁵ The first two sentences were read to the subjects and then the last sentence was presented to them visually, word by word.

- (22) a. As agreed upon, Jane was to wake her sister and her brother at five o’clock in the morning. But the sister had already washed herself, and the brother had even got dressed. Jane told the brother that he was exceptionally **quick**.
b. As agreed upon, Jane was to wake her sister and her brother at five o’clock in the morning. But the sister had already washed herself, and the brother had even got dressed. Jane told the brother that he was exceptionally **slow**.

Note the parallels with (8), which also has a last sentence that is anomalous in context but fully acceptable in isolation. They also tested these final sentences in isolation from the preceding discourse:

21. Nieuwland and van Berkum, “When Peanuts Fall in Love.”

22. Hagoort et al., “Integration of Word Meaning and World Knowledge.”

23. Similarly, Murray, “Pragmatic Plausibility Effects,” 2–3 reports that in eye-tracking experiments investigating the reading of category mistakes (which he calls examples of ‘anomaly’) and sentences that are merely highly implausible there is little or nothing to indicate that we are dealing with anything other than a continuum of plausibility. Magidor, *Category Mistakes*, 152–3 also seems sympathetic to this view.

24. Van Berkum, Hagoort, and Brown, “Semantic Integration in Sentences and Discourse.”

25. I give the literal translations provided by van Berkum, Hagoort and Brown of the Dutch originals. The target words are bolded here, but were not distinguished typographically or otherwise in the experimental materials.

- (23) a. Jane told the brother that he was exceptionally **quick**.
 b. Jane told the brother that he was exceptionally **slow**.

And they tested single sentences that were anomalous without the help of preceding discourse, and similar controls, as in (24):

- (24) a. Gloomily the men stood around the **grave** of the president.
 b. Gloomily the men stood around the **pencil** of the president.

The subjects showed N400 effects for the two kinds of anomaly (discourse, as in (22b), and single-sentence, as in (24b)) ‘with the same overall morphology, overall time-course, and distribution across the scalp’.²⁶ The effects started 200–300 ms after the appearance of the anomalous word. The authors conclude as follows:²⁷

The similarity of these effects, particularly in polarity and scalp distribution, is compatible with the claim that they reflect the activity of a largely overlapping or identical set of underlying neural generators, indicating similar functional processes [...] There is no indication whatsoever that the system is slower in relating a new word to the semantics of the wider discourse than in relating it to local sentence context.

In a follow-up study, van Berkum et al. used the same materials as the previous study but presented the target sentences aurally.²⁸ Exactly comparable results were obtained. Moreover, the subjects in this experiment related the target words to discourse-level representations within 150–200 ms after the acoustic onset of the words. The target words took about 550 ms to pronounce and contained an average of 6.7 phonemes. So they were related to discourse-level information after just two or three phonemes had been pronounced. This extremely early detection of discourse-level anomaly was also found in a similar study by Camblin, Gordon and Swaab.²⁹

The relevance of these studies to the current examples is clear. The theory currently being investigated suggests that the category-mistake phenomenology in examples like (8) is due to people revisiting the first sentence and subjecting it to a further round of processing after they have heard or seen the second sentence. But the experimental literature on this kind of question suggests that people in this kind of situation do nothing of the sort. They are capable of detecting discourse-level anomaly caused by a particular word before the speaker has even finished saying it, and no more slowly than they detect sentence-level anomaly. This strongly suggests that the current version of the MBT view is on the wrong track. Indeed, van Berkum, Hagoort and Brown even cite Stalnaker’s common ground as a likely explanation for the effects they observe.³⁰ This, of course, is fully consistent with Magidor’s theory.

It seems, then, that neither of the two versions of the MBT view that we have just considered can easily support the view that the first sentence of (8) is meaningful but truth-valueless. We earlier saw serious difficulties for the view that the second sentence had this status. So the MBT view is faced with a significant problem when it comes to accounting for examples of multi-sentential category mistakes. Magidor’s theory, on the other hand, accounts for these examples naturally, even when we take into account fine temporal details of the processing of relevant examples.

4 Conclusion

Magidor’s theory makes the prediction that there will be discourses like (8). The other theories do not. Likewise, Magidor’s theory accounts for these data naturally. The others do not. So

26. Van Berkum, Hagoort, and Brown, “Semantic Integration in Sentences and Discourse,” 663.

27. Ibid.

28. Van Berkum et al., “When and How do Listeners Relate a Sentence to the Wider Discourse?”

29. Camblin, Gordon, and Swaab, “The Interplay of Discourse Congruence and Lexical Association.”

30. Van Berkum, Hagoort, and Brown, “Semantic Integration in Sentences and Discourse,” 665.

Magidor's theory of the infelicity of category mistakes has a clear advantage over its rivals when it comes to multi-sentential category mistakes.

In the light of the neurolinguistics literature reviewed above, we can also conclude that category mistake phenomenology is an N400 response. Category mistakes constitute a variety of a wider class of anomalous sentences that have been studied in this literature. The rich empirical literature on N400 effects will hopefully offer us further insights into this phenomenon in the future.

Acknowledgements It is an honour to be invited to comment on Magidor's groundbreaking book. I am grateful to Herman Cappelen for the invitation and to Ofra Magidor and Daniel Rothschild for valuable discussion of this and many other things.

References

- Camblin, C. Christine, Peter Gordon, and Tamara Swaab. "The Interplay of Discourse Congruence and Lexical Association during Sentence Processing: Evidence from ERPs and Eye Tracking." *Journal of Memory and Language* 56 (2007): 103–28.
- Elbourne, Paul. *Definite Descriptions*. Oxford: Oxford University Press, 2013.
- Frege, Gottlob. "Über Sinn und Bedeutung." *Zeitschrift für Philosophie und Philosophische Kritik* 100 (1892): 25–50.
- Hagoort, Peter, Lea Hald, Marcel Bastiaansen, and Karl Magnus Petersson. "Integration of Word Meaning and World Knowledge in Language Comprehension." *Science* 304, no. 5669 (April 2004): 438–41.
- Lau, Ellen, Colin Phillips, and David Poeppel. "A Cortical Network for Semantics: (De)constructing the N400." *Nature Reviews Neuroscience* 9 (December 2008): 920–33.
- Magidor, Ofra. *Category Mistakes*. Oxford: Oxford University Press, 2013.
- Murray, Wayne. "The Nature and Time Course of Pragmatic Plausibility Effects." *Journal of Psycholinguistic Research* 35, no. 1 (2006): 79–99.
- Neale, Stephen. *Descriptions*. Cambridge, Mass.: MIT Press, 1990.
- Nieuwland, Mante, and Jos van Berkum. "When Peanuts Fall in Love: N400 Evidence for the Power of Discourse." *Journal of Cognitive Neuroscience* 18, no. 7 (2006): 1098–111.
- Quine, Willard van Orman. *Mathematical Logic*. Cambridge, Mass.: Harvard University Press, 1940.
- Russell, Bertrand. "On Denoting." *Mind* 14 (1905): 479–93.
- Sperber, Dan, and Deirdre Wilson. *Relevance: Communication and Cognition*. Revised second edition in 1995. Oxford: Blackwell, 1986.
- Stalnaker, Robert. "Pragmatic Presuppositions." In *Semantics and Philosophy*, edited by Milton K. Munitz and Peter Unger, 197–214. New York: New York University Press, 1974.
- . "Pragmatics." *Synthese* 22 (1970): 272–89.
- Strawson, Peter. "On Referring." *Mind* 59 (1950): 320–44.
- Van Berkum, Jos, Peter Hagoort, and Colin Brown. "Semantic Integration in Sentences and Discourse: Evidence from the N400." *Journal of Cognitive Neuroscience* 11, no. 6 (1999): 657–71.

- Van Berkum, Jos, Pienie Zwitserlood, Peter Hagoort, and Colin Brown. "When and How do Listeners Relate a Sentence to the Wider Discourse? Evidence from the N400 Effect." *Cognitive Brain Research* 17 (2003): 701–18.
- Van Burkum, Jos. "Understanding Sentences in Context: What Brain Waves Can Tell Us." *Current Directions in Psychological Science* 17, no. 6 (2008): 376–80.