

Castagnoli, Luca. 2010. *Ancient Self-Refutation: The Logic and History of the Self-Refutation Argument from Democritus to Augustine*. Cambridge: Cambridge University Press (394 pages, ISBN 978-0-521-89631-3 (hardback)).

In this fascinating and well-written book, Dr. Luca Castagnoli draws on nearly a decade of research on the topic of self-refutation and self-refuting arguments in ancient and early medieval Western philosophy (he leaves aside cases of similar argumentation in the Eastern tradition; see pg. 2, fn. 4 for references to that tradition). Both philosophical and historical in its aims, Castagnoli's book expands on and critiques previous research on the topic by Mackie, Burnyeat, and others, and provides a new interpretation of what it means for an argument to be self-refuting.

The book begins – as all good books do – with outlining and then motivating its scope. The first challenge that the book addresses, in its short introduction, is the question of “what should count as an ancient instance of *self-refutation*?” (p. 2). As a basic first response, Castagnoli defines a ‘self-refutation argument’ as “any argument which aims at showing that (and how) something is ‘self-refuting’, i. e., refutes itself” (p. 3); but this, of course, is not very satisfactory as a definition for we do not know what it means for something to be self-refuting or to refute itself. Indeed, Castagnoli stresses throughout his book the difficulty in providing a uniform definition of ‘self-refutation’ or ‘self-refuting’. In the introduction, instead of trying to give a clear-cut definition in advance of discussing examples, he focuses on the criteria which he will use for selecting out these examples: (1) the occurrence of relevant key terms or phrases, such as ‘refutes itself’, ‘eliminates itself’, ‘overturns itself’, etc., in the texts of these arguments, (2) arguments which have been called ‘self-refuting’ in modern literature, and (3) the level of similarity between these arguments and the ancient arguments which display the relevant terms and phrases in (1) (pp. 4–5). He also distinguishes between arguments which are self-refuting and those which are self-contradictory, explaining that the two notions must be kept distinct and that it is the former, not the latter, in which he is interested.

It is expanding on this point – the distinction between self-refutation and self-contradiction – in the first chapter, on “Preliminaries” that the first surprise is met; many people, before reading this book, are perhaps like me, and when hearing ‘self-refuting argument’ think immediately of the Liar Paradox, the contradictory statement “This sentence is false”. This chapter is intended, among other things, to explain why the Liar Paradox is *not* the focus of the book, and why it will not be discussed further. Two divisions that Castagnoli makes are relevant for removing the Liar Paradox from the scope of the book; the first is the division between self-refutations and self-contradictions, and the second is between arguments and theses (p. 14). As Castagnoli points out,

We should be very careful not to confuse ‘Everything is false’ and ‘Everything is true’ with the Liar and the Truth-Teller. The latter were *arguments* (often branded as σοφίσματα, ‘sophisms’, or ἀποροι λόγοι, ‘insoluble arguments’), which ancient logicians perceived as a serious menace to the foundations of their logic, whereas, however surprising this might appear to us, the former were advanced as genuine philosophical *theses*, never making their appearance in any ancient list of sophisms or *insolubilia* (p. 14).

The ancient analyses of these two groups also distinguish them; ancient solutions to the Liar Paradox are nowadays neither very well-known nor understood, and they are quite complicated. On the other hand, ancient responses to theses such as “Everything is true” or “Everything is false” are, Castagnoli argues, simple and well-understood. The Liar Paradox, he suggests, offers a genuine challenge for ancient logical foundations; but to hold the thesis that “Everything is false” is a mere embarrassment. Ancient analyses of self-refutations

display the same type of simplicity that analyses of theses like “Everything is false” did; and furthermore, these analyses often served the same purposes (pp. 14–15).

We now begin to know what a self-refutation is: First, it is not a self-contradiction. Self-contradictions are false (because they are contradictory), but self-refutations are merely *undefensible*; they need not be false, but even if they are true, they are still not positions that can be coherently defended in a dialectical context – about which we will discuss more below. This undefensibility, as opposed to falsity, is one of the defining traits of self-refutations on Castagnoli’s analysis.

The second trait of self-refutations that Castagnoli discusses is their ellipticality. Many of our textual sources for ancient self-refutations are extremely compressed in their presentation, and the analyses must be carefully reconstructed on the basis of analyses of similar arguments and theses in other contexts; such reconstruction is masterfully done by Castagnoli throughout the book. However, while all of the self-refutations that he discusses in the rest of the book do show this trait in varying degrees, it is not clear that this trait is either necessary for a self-refutation or that it is unique to them, and thus it cannot serve as a *differentia* for the species “self-refuting”.

After the preliminaries are set forward, the rest of the book is devoted to a tour through self-refutation arguments, in (mostly) historical order. Before the ancient arguments are met, however, Castagnoli first discusses the typology of self-refutation given in Mackie’s seminal paper “Self-Refutation: A Formal Analysis”, which he calls the “gold standard, on account both of its undeniable merits and of the relative scarcity of previous and subsequent attempts” (p. 17). Mackie distinguishes absolute self-refutation from pragmatic self-refutation, a division which Castagnoli will adopt. Despite his obvious approbation of Mackie, Castagnoli points out many places where he thinks Mackie’s typology can be revised and improved. He criticizes Mackie’s analysis of the self-refutation of “Nothing is true”, pointing out that the argument requires both one to read the operator T sometimes as a sentential operator, attaching to obtaining states of affairs, and other times as a truth-predicate, attaching to whatever it is that are the bearers of truth, be they propositions, sentences, or something else, and one to read \neg sometimes as ‘It is not the case that ...’ and sometimes as ‘...is false’. This ambiguity leads to four different ways that the conclusion to the argument can be read, none of which, Castagnoli argues, follows from the premises in an unequivocal manner. As Castagnoli shows, rather conclusively in my opinion, self-refutation is not, unlike self-contradiction, something that can be analysed in a wholly *formal* way; the context of evaluation provides crucial material to the analysis of the viability of the self-refutation, and these contextual issues cannot be completely formalized.

Perhaps the biggest lesson to learn from the first chapter, and indeed the whole book, is the importance of the dialectical context. Castagnoli calls a ‘dialectical context’

loosely, any situation in which two opposing parties – either individuals or groups, not necessarily facing each other in the flesh – advance and support contradictory views and agree to try to settle their dispute through arguments (typically structured in the form of question and answer) responding to some shared rational standards or rules (p. 27).

The fact that charges of self-refutation occur in dialectical contexts has important consequences for their evaluation, and this fact plays a central role in Castagnoli’s analyses. One of the most crucial consequences of the dialectical context is that certain moves will be disbarred due to the fact that they occur in a dialectical context. Unlike the Liar Sentence, which is self-contradictory due to syntactic and semantic features inherent in the sentence and thus present in any context of evaluation, it is possible for some sentences to be self-

refuting on some contexts but not in other, or for some theses not to be coherently assertible even if there is nothing which prevents them from being *de facto* true.

Castagnoli divides the refutations into three different types: Absolute self-refutations (Part I); pragmatic, *ad hominem*, and operational self-refutations (Part II); and self-refutations related to skepticism (Part III). In the first section, in Chapter 3 Castagnoli discusses the arguments in the anonymous *Δισσοὶ λόγοι* – standardly dated to around 400 BC, and hence one of the earliest texts he considers – that ‘Every *λόγος* is true and every *λόγος* is false’, and in Chapter 4 he looks at how this thesis and related show up in the *Euthydemus* in consideration of whether ‘it is impossible to speak falsely’. In this chapter, Castagnoli tries to show that “there are *some* Platonic arguments that cannot be bared of their ‘conversational clothing’ without ruinous logical or philosophical losses” (p. 32), another point in support of his continual contention of the importance of the dialectical context to understanding ancient self-refutations. This chapter also contains an analysis of that well-known view of Protagoras, the ‘Measure Doctrine’ that “Man is the measure (*μέτρον*) of all things, of those which are, that they are, of those which are not, that they are not” (p. 41).

Aristotle is the focus of the next chapter, which opens with an acknowledgment of the “scarce interest he [Aristotle] appears to show in [self-refutation], both in comparison with his great teacher and in absolute terms” (p. 68); in fact, in much of the chapter, Castagnoli considers passages in Aristotle which have been called self-refuting, or identified as containing self-refutations, by modern commentators, and argues that they have been mislabeled. He uses these examples to point out features which he believes are crucial for the label ‘self-refutation’ or ‘self-refuting’ to be appropriate; this methodological maneuver is the focus of § 5.3.

Chapter 6 covers a theme of central importance in the analysis of ancient self-refutation, the introduction of a technical vocabulary designed specifically for self-refutation, in the form of the verb *περιτρέπειν* ‘to reverse, to turn around’ and its substantive form *περιτροπή* ‘reversal, turnabout, turning the tables’. Castagnoli argues that these words, while used metaphorically in Sextus, had their origin in quite literal applications (Chapter 9 contains – complete with pictures! – an interesting discussion of these and related terms in Greek wrestling vocabulary), and came, in “Sextus, and more generally in Hellenistic and post-Hellenistic philosophical jargon ... to indicate not simply a reversal or refutation of a thesis into its contradictory, but a *self*-reversal or *self*-refutation which has that very thesis as its own premiss” (p. 96). It is also in this chapter that the relation between the so-called *Consequentia Mirabilis*, the thesis $(p \rightarrow \neg p) \rightarrow \neg p$, and self-refutation arguments is considered; Castagnoli wants to reject any suggestion that ancient self-refutations are a type of *Consequentia Mirabilis*, or that ancient authors had any notion of this type of argument structure in mind when advance their charges of self-refutation.

The *Consequentia Mirabilis* returns in Chapter 7, whose subject matter differs from the previous in both timing and content. Here, Castagnoli discusses, for purposes of contrasting illustration, Augustine’s proof of the imperishability of truth in the *Soliloquies*. Castagnoli points out that “Augustine seems to follow an argumentative pattern different from all those we have encountered so far in the ancient landscape” (p. 121) and that by considering the differences we can gain a better understanding of self-refutation. While Castagnoli argues in many places in the book that ancient self-refutation arguments are *not* to be understood as a type of or invoking a *Consequentia Mirabilis*, he says that Augustine’s proof of the imperishability of truth “while not making explicit reference to CM, seems to rely on some analogous inferential step” (p. 123); the appearance of a genuine *Consequentia Mirabilis* in Augustine’s writings is the topic of Chapter 12, which discusses his argument *Si fallor, sum* (“If I am mistaken, I am”). The argument, as Castagnoli notes, does “not use the vocabulary of self-refutation” (p. 197), but he argues that it is not an attempt to prove his own existence

in a constructive fashion but rather it “is meant to provide a dialectical refutation of the Sceptics inasmuch as it answers their suspicion ‘What if you are mistaken?’” (p. 201), and thus it is not unreasonable to discuss it in terms of self-refutation.

In Part II, the focus switches from the absolutely self-refuting arguments of the previous section to arguments which Castagnoli classifies as pragmatic, *ad hominem*, or operational self-refutations. He considers “a quite heterogeneous group of ancient self-refutation arguments which have been classified by commentators (or appear to be liable to be classified) as early instances of what, following Mackie’s taxonomy, are commonly dubbed ‘pragmatic’ and ‘operational’ self-refutation” (p. 145), with a goal of showing that Mackie’s categories need to be revised to provide a proper understanding of these arguments. These arguments include Epicurus against the determinists (Chapter 9), anti-sceptical dilemmas (Chapter 10), Aristotle’s argument in the *Protrepticus* that ‘one must philosophize’ (Chapter 11), and the *Si fallor, sum* argument of Augustine mentioned in the previous paragraph. This group of self-refutations is distinguished from the previous, the absolute self-refutations, because the refutation arises because of a conflict “between what one says (λόγῳ) and what one does (ἔργῳ) when saying it”, which Castagnoli says is “exactly the core of pragmatic self-refutation” (p. 172). In defining *ad hominem* self-refutation, Castagnoli follows Passmore, who distinguished this type from strictly pragmatic self-refutation as a refutation “in which one person’s admission that he is speaking or thinking – as distinct from *the fact that* he is speaking or thinking ... – is used as an argument to show that what he is speaking or thinking cannot be in fact the case” (p. 163); both types are contrasted with operational self-refutation, in which “a proposition cannot be coherently asserted because what is implied by making an assertion contradicts the asserted content” (p. 120, fn. 82). Thus, in Part II we have arguments of three different, but related flavors, each involving the act of assertion. A pragmatic self-refutation arises when the fact that an assertion is made conflicts with the content of the assertion; an *ad hominem* refutation when the admission of assertion causes the conflict; and an operational refutation when the implications of making an assertion cause the conflict.

The final Part is devoted to self-refutation in the context of scepticism, looking at arguments from Sextus Empiricus (Chapter 14), the Atomists, the Academics, Stoics, and the Pre-Sextans (Chapter 15). What is interesting here is that self-refutation is apparently sometimes *embraced* as a productive argumentative strategy, rather than a purely destructive one; in particular, in § 14.2.3 it is clear that Castagnoli thinks the self-refuting nature of the sceptical arguments actually *helps* the Sceptic. A second interesting discussion in this section is the distinction he makes between arguments which are (absolutely) self-refuting and those which are merely self-bracketing (cf. § 14.2), which distinction is based on a complex linguistic argument concerning the history and usage of the words περιτροπή and περιγραφή, appealing to paleographic evidence and ancient scribal practices to physically indicate text which should be deleted.

The book ends with a brief summary and conclusions. The theses which are charged with self-refuting status all, “despite their variety, [share] a common trait: they express radically revisionist positions, diverging not only from what was, or at least what would soon become, philosophical orthodoxy, but also, and more basically, from pre-philosophical, commonsense views” (p. 353). Self-refutations are not *consequentiae mirabiles*, and the self-refutation method is not solely about proving that a thesis is *false*, but can also be used to show that it is (dialectically) untenable. The importance of the dialectical context and the act of assertion is re-emphasized (pp. 355–356), as well as the importance of self-refutations not just defeating the dialectical opponent but also giving cause to ridicule him.

However, the book is not without its various unanswered questions, both of interpretation and execution. Regarding the latter, it is strange that Latin and Greek texts are translated

throughout, but quotations in French, Italian, and other languages are not, given that the audience of such a book is more likely to know the former two languages than the latter. While of course one could desire that any well-trained philosopher or historian would be versed in all of the relevant languages, this is unfortunately unrealistic and thus the book could have been made more accessible by providing translations of all texts not in English. With respect to the former, there are two points I wish to highlight. First, while Castagnoli does an admirable job in his attempt to provide a definition, by ostension if by nothing else, of self-refutation, it is not completely clear how self-refutations are to be distinguished from *reductio ad absurdum* arguments (cf. p. 96, fn. 4). One might be tempted to say that in order for a proposition to be self-refuting, that single proposition alone “must be responsible for its own refutation” (p. 98); but Castagnoli points out that in argument T23 it is not apparent that only one proposition is involved; in such a case, it is both not clear how the thesis is then *self*-refuting, if the refutation requires more than one proposition, and the distinction between a self-refuting thesis and a *reductio ad absurdum* argument, which requires the presence of additional premises, is further blurred. Second, when discussing the different types of pragmatic self-refutations, Castagnoli stresses the presence of locutions such as “one must *say*” or “if he *says*” or “those who *say*”, indicating the importance of a particular action in the context of the self-refutation. However, it is not clear how one can report the arguments of others, whether self-refuting or not, without some such locution. Without being provided with an alternative account of how one can report someone else’s philosophical views without using such a locution, it seems potentially question-begging to include this type of ascription within the definition of pragmatic self-refutation.

Notwithstanding these points, the book is fascinating, well-written, and a joy to read. The development of the definition of self-refutation from the initial pages where the idea is murky and undefined to the concluding chapter where a detailed discussion of the unique properties of self-refutations are summarized makes the book read like a mystery, where one is given clues along the way and only at the end is the final story unveiled and made clear. Accessible to both the scholar and the layman, it would serve well as supplementary material on a course on ancient dialectics as well as a resource for the historian and philosopher of the ancient Greek and Latin tradition.

Sara L. Uckelman