Introduction

The starting point for this investigation was the almost trivial observation that the output of research in logic is fundamentally different depending on whether it is conducted with or without formal languages. Of course, some might say that research in logic conducted without formal languages is not even 'logic' properly speaking. But the study of the history of logic suggests that such a view is at best limited and at worst misguided: many logical traditions which do (did) not make extensive use of formal languages often display a level of conceptual sophistication that leaves nothing to be desired relative to modern developments. And yet, they *are* very different from current research in logic, and this fact itself calls for an explanation. After all, what is so special about formal languages and formalisms more generally? Whence the magic?

Typical views on formal languages are based on the premise that they are, above all, *mathematical objects*: a greater level of precision and technical complexity in logical investigations arises because they are precisely defined mathematical objects. Yet it would seem that viewing formal languages exclusively from this point of view offers a very partial and limited explanation of the impact that their use (and uses of formalisms more generally elsewhere) actually has. In the present inquiry, the idea is to adopt a much wider conception of formal languages so as to investigate more broadly what exactly is going on when a reasoner puts these tools to use. Furthermore, most of the arguments presented here generalize to uses of formalisms in other disciplines, including the empirical and social sciences.

Originally, my idea for this project was to look more carefully into the role played by formal languages in the social interactions of logicians, that is, the public sphere of logic as a discipline practised by a community of researchers. However, and as is often the case with research, it turned out that a very different approach, one that I stumbled upon almost by chance, offered a more promising perspective on what is so special about formal languages and their uses in logic: formal languages viewed as *cognitive*

artefacts enhancing and modifying an agent's reasoning processes.[1] So the key question became: what happens on a *cognitive* level when an agent reasons with formal languages and formalisms in general, as opposed to reasoning without such devices at her disposal?

To address this question, I rely extensively on empirical data on human reasoning as amassed by research within the empirical sciences of the mind. To be sure, there is still a fair amount of disagreement and many open questions in this body of research; it seems we have not yet fully 'cracked the code' of how human beings reason. Nevertheless, many of the results so far obtained do shed light (as I attempt to show throughout the book) on what exactly happens, on the cognitive level, when a reasoning agent operates aided by these special devices, namely formal languages and formalisms.

However, empirical data from psychology and cognitive science are not sufficient for the formulation of an encompassing philosophical understanding of the role of formalisms in reasoning. As it turns out, another essential element in the puzzle is the *historical development* of this cognitive technology. My earlier work as a historian had taught me that many of the assumptions that we take for granted in our conceptualizations of a given phenomenon are in fact substantive commitments, corresponding to significant theoretical steps made along the way. Now, given that the goal is to question what is generally (and uncritically) assumed to be the case about formal languages so as to attain a deeper philosophical understanding of the phenomenon, historical analysis does offer a privileged perspective. It allows us to re-evaluate each of the steps made in the developments towards a given status quo – in this case, the current situation of ubiquity of formal languages and formalisms for research in logic and elsewhere.

To be sure, one of the conclusions to be reached here is that there is a sense in which they *are* indispensable (or in any case irreplaceable), but this will be argued for on the basis of empirical data rather than uncritically accepted. Indeed, one of the main theses of the book is that the rationales that are usually attributed to formal languages – increased precision and clarity, counterbalancing the imperfections of ordinary languages – fail to capture the real impact of using formal languages in practice. I will argue that, rather than expressive devices, formalisms are, above all, calculative, computing devices.

[1] By this I do not want to suggest that the public, social perspective becomes unimportant; in fact, my own account of the emergence of logic and deductive reasoning is inherently tied to public situations of dialogical interaction (see section 4.3). Rather, the change of focus is perhaps best viewed as simply postponing the project of adopting the social perspective to a later stage.

Hence, two key components of this study are, first, the analysis of the history of formal languages, and, second, the empirically informed investigation of the cognitive impact of operating with formal languages and formalisms when reasoning. These two elements will be essential in the search for an answer (even if not *the* answer) to the question of why we do or should use formal languages for research in logic, and more generally why (logical and mathematical) formalisms are such powerful epistemic tools.

The historical and the cognitive perspectives also represent two somewhat opposed yet complementary poles in the investigation: the historical perspective emphasizes what is *contingent* in the developments in question, while the cognitive perspective brings in the biological, *necessary* constraints upon the cognitive make-up of human beings. In effect, the historical development of formal languages can be viewed as a process of *cultural evolution* through which humans developed tools that would allow them to perform certain tasks and solve certain problems more efficiently, against the background of the possibilities afforded by human cognition as biologically determined.[2] Naturally, the search process can only take place within the realm of constraints inherent in the human cognitive apparatus, but, just as in processes of biological evolution, a substantial amount of chance and contingency influences these developments. Hence, it would be a mistake to view the developments towards the current situation of ubiquity of formal languages in logic as inexorable; many contingent, cultural, and historical factors played a significant role, as we will see in Chapter 3.[3] At the same time, this does not entail a form of cultural/scientific relativism; indeed, the view to be defended throughout the book is that the development of formal languages and complex systems of notation constitutes *real progress*.[4] A formalism is a powerful technology that allows humans to reason in ways that would be otherwise virtually beyond their reach;[5] explaining why this is so on a cognitive level is the main goal of the book.

[2] In other words, the point is a search for tools leading to an optimalization of human cognition given its inherent constraints. Other examples of such processes to be discussed in the book are the evolution of (spoken) languages towards increasing compositionality (Kirby 2012) and the evolution of writing systems towards a better trade-off between expressivity and learnability (Dehaene 2009).

[3] Staal (2006) also offers an overview of some of these factors.

[4] In many senses, the current approach follows Netz's (1999) idea of a 'cognitive history' (Netz also opposes the social constructivist (relativist) perspective of, e.g., Steven Shapin).

[5] However, a word of caution seems to be called for here: it must also be stressed that, as with virtually all technologies, there is always a trade-off involved. Precisely because they are a powerful technology, the use of formal languages also entails certain risks and dangers. In other words, it does not come 'for free', and things *can* go wrong (some such cases will be discussed in Chapter 3).

In what follows, I present an overview of the book, chapter by chapter. The aim is to provide a bird's-eye view of the overall purposes of the project, and of how its different elements all tie together.

Part I is composed of three chapters, and focuses predominantly on the history and philosophy of formal languages. It departs from the conviction that the conception of formal languages as mathematical objects only 'scratches the surface' of their actual cognitive impact.[6] Therefore, a new reconceptualization of formal languages is required in order to dig deeper. As already suggested, I propose to view formal languages and formalisms as a *cognitive technology*, i.e., as specific devices that may enhance and even modify the reasoning and cognitive processes of human agents in which they are involved. For this purpose, I discuss the following topics: two meanings of 'formal' that seem particularly relevant for the conception of formal languages as *formal*; the very concept of a formal language; the history of the development of this technology, which is thus presented as a *cultural product*; and reasons why one should or should not use formal languages and formalisms (specifically, but not exclusively, when doing logic).

Chapter 1 presents two senses of the term 'formal' that seem particularly relevant for the attribution of a formal character to formal languages and formalisms. Insofar as 'formal language' is not simply a set phrase, it is important to investigate more closely what is *formal* about formal languages and formalisms in order to improve our conceptual understanding of them. The two notions discussed are the formal as *de-semantification* and the formal as *computable*, and they remain central throughout the book.

Chapter 2 lays down the conceptual grounds for the whole investigation. It first presents a discussion of the very concept of a *language* (essentially adopting an evolutionary perspective), and analyses a series of crucial distinctions: spoken v. written languages, natural v. artificial languages, languages as practices v. languages as objects. In the second part, I discuss in which sense(s) formal languages are *languages* and in which sense(s) formal languages are *formal*. Then, I present an outline of the conceptualization of formal languages as a cognitive technology, which will guide the analysis throughout.

Chapter 3 presents a condensed history of the development of this technology. Focusing exclusively on the current stage of these developments (which we have no reason to believe have come to completion!)

[6] But, of course, the fact that formal languages *are* mathematical objects remains a crucial aspect of their uses in logic, in particular for metalogical investigations (see Chapter 3).

obscures the fact that the formal languages that we now have at our disposal are the product of very long and complex historical processes. Essentially, the development of formal languages is a specific chapter in the development of mathematical notations more generally, which in turn is a specific chapter in the development of writing systems. For the most part, the development of formalisms (including what we could describe as proto-formalisms) was driven by the search for more efficient tools for calculation/computation.

Chapter 3 also surveys the main views defended so far concerning the rationale for using formal languages when doing logic (by Leibniz, Frege, Wittgenstein, Carnap, etc.). I argue that none of them gets to the bottom of the phenomenon, even though there is some truth in each of them. From these proposals, three main functions for formal languages emerge: as expressive devices, as iconic devices, and as calculating devices. In the second part of the chapter, I argue that there are costs, risks, and limitations that must not be overlooked when using formal languages (for logic and other kinds of inquiry); it is, as always, a matter of trade-offs. This is important, as the essentially positive picture of the impact of using formal languages and formalisms presented here must not blind us to the fact that there are downsides too.

Part II is the truly empirically informed part of the investigation, where I offer extensive empirical support for the hypothesis of formal languages and formalisms as cognitive technologies. As a first step, I present an outline of the spontaneous reasoning patterns in humans, so as to be able to clarify in which ways formal languages represent a device that enhances, complements, modifies, and in some circumstances *corrects* these spontaneous patterns. It will also prove useful to review the recent empirical results on the neuroscience of reading, given the trivial but nevertheless often forgotten observation that formal languages are *written* languages. Neural mechanisms involved in reading processes can ultimately offer us insight into why formal languages and formalisms can make us reason in ways that we would otherwise not be able to (or with much difficulty); I describe this effect as the 'debiasing effect' of reasoning with formal languages.

Chapter 4 surveys some familiar results of research in psychology of reasoning, which strongly suggest that our spontaneous reasoning patterns are nothing like the precepts of 'logical reasoning' as traditionally construed (as also surveyed in, e.g., Evans 2002). I discuss in particular what has been described (Stanovich 2003) as a 'fundamental computational bias of human cognition', namely the systematic tendency to bring prior beliefs to bear when reasoning and solving problems. I also argue for a pluralistic

conception of human rationality, one which makes room for different canons of reasoning in different life situations.

Chapter 5 presents an examination of formal languages from the point of view of *extended cognition* (Clark 2008; Menary 2010b). The chapter begins with a discussion of the neuroscience of reading and writing, essentially following the results reported in Dehaene 2009. I then turn to the concept of extended cognition and apply it to formal languages and formalisms. I examine the results reported in Stenning 2002 on how different formal languages (discursive v. diagrammatic) have an impact on a student's process of learning logic (in an institutional setting). Equally important are the results by Landy and Goldstone (2007a; 2007b; 2009) suggesting that working with formalisms crucially relies on sensorimotor processing. Finally, I claim that formal languages not only increase our computational power, but actually allow us to 'run a different software' when reasoning, as it were;[7] in this sense, my analysis of formal languages from the point of view of extended cognition falls squarely within Sutton's (2010) 'second-wave extended mind' framework and Menary's (2007a) 'cognitive integration' conception of extended cognition.

Chapter 6 examines the crucial concept of 'de-semantification', as introduced by Krämer (2003). The concept is also discussed in Chapter 1, but now I return to it against the background of the experimental results presented in previous chapters, and connect it with experimental research on the phenomenon of *semantic activation* (i.e., the cognitive processing taking place when an agent hears or reads words). I claim, however, that the concept of de-semantification by itself is not sufficient to account for the processes involved in reasoning with formal languages (or formalisms, more generally). A closely related concept is then introduced, namely the concept of 're-semantification'; it concerns the possibility of applying a given formalism, which is developed against a specific background, to a *different* problem, phenomenon, or framework. In such cases, rather than being considered as 'meaningless', the formalism is given a new meaning, but the debiasing effect of de-semantification (to be discussed in Chapter 7) also occurs.

Chapter 7 explains in more detail what I call the 'debiasing effect' of reasoning with formal languages and formalisms. It begins with a philosophical, schematic discussion of the very concept of formalization, which will pave the way towards an explanation of how de-semantification and sensorimotor manipulation of the notation are combined for the debiasing

[7] The reference to software here should be understood as metaphorical, and thus not as an endorsement of computational theories of the mind.

effect of reasoning with formal languages and formalisms to come about. However, it is important to notice that the debiasing effect is particularly prominent in cases of *applications* of formalisms, i.e., when logical or mathematical apparatuses are employed for the investigation of particular (extra-logical) phenomena – that is, when formalisms are *tools* rather than *objects* of investigation in themselves. Because it allows us to counter the 'computational bias' of systematically bringing prior beliefs into the reasoning process, reasoning with formal languages and formalisms in fact increases the chances of obtaining surprising results, i.e., results that go beyond previously held beliefs. This, I think, is the real magic, the *generosity*[8] of formal languages. I also discuss the work of Houdé and collaborators on inhibition training (e.g., Houdé and Tzourio-Mazoyer 2003), which represents a very different approach to debiasing, and draw some implications from the debiasing account of formal languages and formalisms for the currently popular dual-process model of human cognition.

I am well aware that the methodological approach adopted here is somewhat unconventional. There is still some opposition to the idea that philosophical analysis should be informed by empirical data (on human cognition or otherwise). Similarly, the relevance of historical investigations for systematic issues remains controversial. Now, if there are reservations towards the philosophical relevance of these two points of view taken in isolation, they are likely to intensify towards a philosophical analysis integrating *both*. I believe, however, not only that a wide range of philosophical problems can benefit from such an integrative approach, but in fact that they cannot be satisfactorily treated unless historical and empirical elements are taken into account. To elaborate on this point, in the concluding chapter, I discuss in more detail the methodological choices made for this investigation, and some of their implications in view of the results obtained.

[8] In the sense of d'Alembert's famous saying, 'Algebra is generous; it often gives us more than what is asked of it.'

I