Utrecht University

Faculty of Humanities

Department of Philosophy and Religious Studies

Research Master Philosophy

Master Thesis

A Wittgensteinian Reflection on Universal Grammar

Author:

Bart Kuijsten

First Supervisor:

Prof. Dr. Albert Visser

August 26, 2015

Contents

Chapter 1

Introduction

In his Managua Lectures, Chomsky talked about the study of linguistics, and the essential framework in which this study must be conducted. One of the questions which must be answered is: "How does this system of knowledge arise in the mind/brain?" (Chomsky, 1988, p. 3), which he takes to be a special case of what he calls 'Plato's Problem'. That is, we need to explain how it is that our knowledge of language encompasses more than what we could obtain through experience, the argument for which is known as the poverty of stimulus argument. His answer is: because we as human beings have a language faculty, the grammatical rules and concepts are innate, and our contact with the world (the data that we need to be exposed to) is only needed in order to program the system, as it were. We only need to attach the language-specific sounds to our innate concepts in order to speak a language. In short, there is no discrepancy between the small amount of experience and the vast amount of knowledge, precisely because we are born with that knowledge. If true, his findings would be remarkable, reason enough indeed to compare his results to the great scientific revolution of the seventeenth century (Chomsky, 1988, p. 92). But alas, his entire idea is based on a conceptually confused notion of rules.

This will be the main aim of our thesis: to show that Chomsky confuses what it is to follow a rule mechanically with a mechanical rule. Precisely because he does not differentiate between the two, he is misled into thinking that any theory about our knowledge of language has to entail all possible applications. That is, that our knowledge of language somehow has to contain all the possible ways in which we use language. Our main arguments in this thesis will be largely based on Wittgenstein's remarks on rule-following, as found in his *Philosophical Investigations*, *Remarks on the Foundations of Mathematics*, and the collected and edited notes from his students as published in *Lectures on the Foundations of Mathematics*. This debate is not new, for example Hacker already commented on Chomsky's theory in "Chomsky's Problems" published in 1990. But although this thesis inevitably has some overlap with Hacker's paper, Hacker neglected the source of Chomsky's confusion, and, as we will argue, leaves the main concern untouched.

In order to reach our conclusion, chapters 2 and 3 contain a reconstruction of Chomsky's theory, and Hacker's argument against said theory. In these chapters we will also briefly reflect

on both. Chapter 4 is devoted to reconstructing some of Wittgenstein's remarks needed to be able to make the conceptual distinction between rule-following as an activity, and rule-following as a mechanical process — which, strictly speaking, is not rule-following at all. In chapter 5 we will return to Chomsky's theory, and disentangle the confusion that lies beneath his theory of the language faculty. But first we will give some attention to the context in which Chomsky's theory arose, as well as some methodological specifics of this thesis.

1.1   Methodological remarks

According to Freidin, the current perspective on the study of linguistics — generative grammar — evolved from a shift in focus "from the external forms of a language to the grammatical system that generates it" (Freidin, 2012, p. 895). The second "major development" Freidin mentions, is Chomsky's interpretation of grammar as a "model of linguistic knowledge in the mind of a speaker, consisting of a finite number of grammatical rules, including a finite lexicon" (Freidin, 2012, p. 900). In this development, language, i.e. ordinary languages, became derivative, derived from grammar, which was characterized as a system of knowledge in the mind of a speaker. This gave rise to a psychological framework, in which four "basic questions" were raised (ibid.). These questions concern *what* this system is, *how* it arises, *how* it is used, and *where* it is located, and are, according to Chomsky, "the essential framework for further inquiry" (Chomsky, 1988, p. 8).

In this thesis we will take the four questions concerning grammar as a system of knowledge as a starting point for our criticism. We will argue that these questions arise precisely because Chomsky is confused about what it is to follow a rule. This results in a slight tension in this thesis. On the one hand, we are focusing on the psychological framework Chomsky introduces, on the other hand we are not interested at all in the technical specifics of the postulated system of knowledge. But when characterizing the framework, some specifics inevitably slip through the cracks. In itself this is not a problem, but the publication we will mainly rely on, viz. *Language and Problems of Knowledge* (also known as the Managua Lectures (ML)), is part of a continuous development in Chomsky's theory. For example, when Chomsky gave his lectures, construction- and language-specific rules — for example in passive transformation — were already replaced by more general rules, a system called the 'Principles and Parameters framework' (P&P) (Freidin, 2012, p. 905). Furthermore, while the earlier (1965) Chomsky saw language learning as something that a child actively does — by testing hypotheses, etc. — when Chomsky gave the lectures, this view was already replaced by what Cowie characterizes as a "less intellectualized conception" where the (in that time new) parameter setting was something that happened to the child (Cowie, 2010, §2.1).

After ML, there have been further developments of Chomsky's theory, where the P&P has been developed into what is called the 'Minimalist Program', which focuses more on the de-

mands of conceptual necessity of grammar (Freidin, 2012, p. 909ff).[1]

Returning to the issue at hand, since we are not interested in the technical details of grammar as a system of knowledge, we could have picked any of Chomsky's works. So, why focus mainly on Chomsky's ML (with the occasional reference to his *Rules and Representations*) when this is not adequate to give a general account of his position? There are two reasons. Firstly, in ML we find a lengthy discussion of the four questions which were characterized as being the essential framework for further inquiry into the study of linguistics. And it is precisely the root of these questions that we want to refute. Secondly, and more pragmatically, since the specifics are irrelevant to our argument focusing on *Language and Problems of Knowledge*, as an accessible introduction, makes our task easier.

Which brings us to our second methodological remark. Our criticism of Chomsky's theory is given from a Wittgensteinian perspective. That is to say, we use Wittgenstein to create a position (give a systematic interpretation), which in turn is used to dispute Chomsky's theory. Therefore we are not concerned with Wittgenstein's own developments — for example the transition from his earlier *Tractatus* to his *Untersuchungen*. Because the main argument will be Wittgensteinian, we will not discuss criticism of Chomsky's theory expect insofar it is relevant to our Wittgensteinian critique. This means that our main sources of literature will be Chomsky and Wittgenstein. Sporadically secondary literature is used, for two reasons. On the one hand, secondary literature is used when a brief overview was needed, for example when we discuss the *Poverty of Stimulus* argument in chapter 3. On the other hand, secondary literature is used if it played a part in the interpretation reached in one of my earlier papers, the conclusions of which are used in chapter 4.

Since this thesis is not the first Wittgensteinian critique of Chomsky, we have to relate (somewhat) to what is already published. Hacker felt like a natural choice because he is well-known to be an expert on Wittgenstein's work, and the article we will use is explicitly linked to what we want to focus on in our criticism. Furthermore, we will argue that Hacker's argument can be improved if his interpretation of Wittgenstein's work would differ slightly, which helps the reader to differentiate between the two interpretations of Wittgenstein and simultaneously serves as an argument in favour of our interpretation.

---

[1]For a more detailed overview of the developments in Chomsky's theory see (Freidin, 1994).

Chapter 2

Ability versus System of Knowledge

In *Language and Problems of Knowledge* Chomsky states: "the idea that knowledge is ability is [...] entirely untenable" (Chomsky, 1988, p. 9). This remark is part of Chomsky's response to what is, according to him, the often held view. Namely, the view that reduces language to a habit system, in order to avoid the fear of 'mentalism', and to "exorcise the Cartesian 'ghost in the machine'" (ibid.). Chomsky gives four separate arguments (or considerations) that are meant to show that this conception of knowledge is mistaken. He argues that knowledge of language cannot be seen as an ability — for example the ability to speak, understand, etc. Responding to Chomsky, Hacker wrote an article, titled "Chomsky's Problems" published in 1990, where he addresses Chomsky's considerations, and argues that speaking a language is an ability, not a system of knowledge.

In the first section of this chapter we will repeat Chomsky's considerations. The second section is devoted to Hacker's counterarguments, and the third and final section is where we reflect on Chomsky's arguments and the confusion that lies beneath them.

2.1   Chomsky's Considerations

In disputing the idea that knowledge is ability, Chomsky gives four considerations. In the first consideration he asks us to consider two people, sharing exactly the same knowledge of a language, in his example, Spanish. "Their pronunciation, knowledge of the meaning of words, grasp of sentence structure, and so on, are identical" (Chomsky, 1988, p. 9). Nevertheless, these people differ greatly in their ability to use the language, one is a poet, the other a 'pedestrian language user'. *Ex hypothesi*, these two people, having exactly the same knowledge of Spanish, will say different things on given occasions. Hence, Chomsky concludes, "it is hard to see how knowledge can be identified with ability, still less with disposition to behavior" (Chomsky, 1988, p. 10).

The second consideration is aimed at exercises that improve one's ability, but do not increase one's knowledge of language. For example taking a course in public speaking. If all goes

well, one's ability to deliver speeches is improved, but there is no new knowledge concerning the language, that is, there is no increase in vocabulary, knowledge of grammatical rules, etc. (ibid.).

The third consideration is about the loss of one's ability to speak, without a loss of knowledge. Chomsky portrays a man, Juan, suffering from aphasia after a severe head wound. Juan, having lost all his ability to speak and understand, might not have lost his knowledge. He can recover, thereby regaining his ability, which would be remarkable if his knowledge was lost: "How did he develop this ability without instruction or experience, something that no child can do?" (Chomsky, 1988, p. 10) The unbelievable alternative, of the miraculous speed at which Juan would learn the language, shows that something must have been retained. What has been retained cannot be the ability — the hypothesis excludes that — but is, according to Chomsky, a system of knowledge: "a *cognitive system* of the mind/brain" (ibid.). Which means that knowing a language cannot be identified with an ability, or a system of dispositions, skills, etc.

Now, for the fourth consideration. A speaker of Spanish knows that 'el libro' refers to a book, not to a table. "This is not a failure of ability on his part. It is not because he is too weak, or lacks some skill, that *el libro* fails to refer to tables" (Chomsky, 1988, p. 12). It is a property of the system of knowledge that 'el libro' refers to a book, and this system of knowledge is precisely what a speaker of Spanish possesses.

In the introduction we briefly stated the shift in focus which arose from Chomsky's work. Where traditional linguistics was focused on the external form of languages, for example English, Chomsky focuses on the grammar that generates such a language. From the psychological framework, characterized by the four questions mentioned earlier, follows that there is a distinction between the knowledge a speaker possesses (his *competence*) and the output a speaker produces (his *performance*). The four arguments are meant to show that equating knowledge of language with the ability to speak a language is misconceived, and that the framework Chomsky suggests should be favoured.

2.2   Hacker's Reply

Hacker's response is that the four arguments are flawed. It is true, Hacker says, that knowing how to do something is not always the same as being able to do something, but the problem is that the criteria for speaking of 'knowing a language' are indeterminate. How much of the vocabulary must one know, and how few grammatical mistakes are permitted before we say that one knows a language?

Hacker gives a response to each of Chomsky's arguments. Regarding Chomsky's first consideration, Hacker argues that given Chomsky's premise, all that really follows is that one can know a language without being able to write poetry. That is, "the phrase 'knowing Spanish' does not subsume the totality of Spanish linguistic skills and arts, but that does not show that it signifies anything other than linguistic abilities" (Hacker, 1990, pp. 130–131).

Hacker's response to Chomsky's second argument follows the same line, in saying that his abilities have improved without any gain in knowledge — although it is altogether unlikely that no knowledge has been gained, but for the sake of argument Hacker accepts that idea — all that really follows is that "linguistic skills are manifold and multi-faceted, incorporating much more than is subsumed under the concept of knowing" (Hacker, 1990, p. 131).

In response to Chomsky's third argument, Hacker points out that there is a conceptual problem, but its resolution does not show that knowing a language is anything other than the ability to speak and understand it (Hacker, 1990, p. 131). In his argument Chomsky uses the notion 'retaining' in saying that something has been retained, enabling the aphasic to recover his speech and understanding without learning anew. Our inclination to say that something has been retained — in cases where someone knew something, forgot it at a certain point, and, still later, remembered it — arises because what was known is not learned *de novo* (Hacker, 1990, p. 132). But, Hacker emphasizes, our notion of retaining is a figure of speech, which we employ in some but not all cases. We employ it in cases of a lapse of memory, in cases where one needs to be reminded, or where a random association will jolt one's memory (ibid.). This 'retaining', is not used to explain why or how one remembers previously forgotten knowledge, our use of 'retaining' merely shows that "Knowing something is compatible with occasional lapses of memory" (ibid.). We would not, for example, say of an amnesiac that he has retained his memory after a long period of not remembering anything, rather we would say that he has regained his memory. It was not a lapse of memory — as in the cases when we speak of retaining — but rather a loss of memory. In case of the aphasic, nothing warrants us to say that he had really retained his knowledge all along.

The tension between having forgotten something and not having to learn it *de novo*, is generated, according to Hacker, "by the picture of knowing as retaining something deposited in the storehouse of the mind or the alleged memory banks of the brain, by the misconception that to know is to be in a persistent *state* [...], and by the idea that if one now knows something because one learnt it or came to know it earlier, then one must have known it *uninterruptedly* throughout the intervening time in as much as it was stored away" (Hacker, 1990, p. 133). Thus, in considering the aphasic, one would like to ask *how* does he know now, if he did not know it during his aphasic period — if it was not stored away? Saying, however, that something was retained, does not answer the question. But even if, Hacker argues, something was retained in the brain during the aphasic period, it could never be knowledge. What one knows is about truths, facts and propositions, and these can be recorded by means of a symbolism, but there is "no such thing as the brain's employing a language" (Hacker, 1990, p. 134). To say that such things are 'filed away' is saying nothing more than that they are known, and not forgotten, and that cannot be said in case of the aphasic.

Regarding Chomsky's fourth consideration, Hacker argues that Chomsky confuses what it is for an expression to refer to or signify an object, with what it is for a speaker to know this relation. The former concerns linguistic rules, while the latter concerns linguistic abilities which

manifest in use. The ability of which one speaks in saying that one knows what "el libro" means, is the second one, not the first. Which is why, according to Hacker, Chomsky's consideration misses the point entirely (Hacker, 1990, p. 134).

2.3   Aftermath

In the previous sections we have seen the four arguments that Chomsky offers in favour of distinguishing knowledge (of a language) from the ability to speak and understand that language, with a reply from Hacker to each of the arguments. In this section we will do roughly the same, but in a different order. First, we will make a brief comment on the first two considerations, after which we will focus on Chomsky's fourth consideration. We will close this section with a comment on the third consideration, because it is indicative of the underlying confusion we want to address.

Chomsky's first two arguments involve the same strategy: stretching what reasonably can be considered to be subsumed under our notion of knowledge, and excluding the abilities in question from that range. Without sounding all too unsympathetic, it must be said that it feels like a cheap trick. Furthermore, it seems flawed. In the first argument there is the distinction between a poet and your average Joe, in the second there is a speaker, learning to be a better public speaker. In both cases there is an ability *a*, writing poetry in consideration one, and public speaking in consideration two, which is *not* part of his knowledge. To see how these cases are problematic, let us introduce a neutral term 'mastery of a language', for now we can keep it primitive, and remain silent on what such mastery consists in. Our difficulty with these arguments is that one would like to think that either *a* is part of mastery of a language, in which case there actually *is* a difference in knowledge of a language — there is, then, something learned in taking a course in public speaking, or poetry — or *a* is not, in which case there is no difference in knowledge of a language, but neither in the mastery of a language. To be more precise, either we include ability *a* in mastering a language, in which case the (new found) skills have to be counted as knowledge of said language as well — articulation, intonation, pace, etc — or, *a* is not part of mastering a language, in which case nothing more is left to be said. Of course, Chomsky can still argue that both abilities are involved in increasing your proficiency in language, but still are not part of one's knowledge of language. But this is just to state that the two are distinct, not to argue for it.

Chomsky's fourth consideration is aimed at the connection between words and concepts. Let us first clarify it, before giving a response. For Chomsky, knowledge of language requires a computational system, as well as a conceptual system (Chomsky, 1980, p. 55).[1] Acquiring knowledge of the word 'el libro', would be establishing its connection with the concept of book, which is part of our conceptual system. This connection would simply be part of a speaker's system of knowledge, his competence. Hacker's response is that Chomsky confuses what it

---

[1] These aspects will be addressed further in the next chapter.

is for an expression to refer to an object, with a speaker's knowledge of this relation. But for Chomsky, these are the same. The established connection between word and concept is the speaker's knowledge of this relation. This argument is difficult to swallow, it presupposes the framework it wants to argue for. The argument fails as soon as Hacker's distinction would be justified. That is to say, if we give primacy to, for example, English as an external language, instead of a system of knowledge in the mind/brain, acquiring a lexicon would then simply be being able to use the words according to rules which have already been fixed.

Chomsky's third consideration has been given most attention by Hacker. Summarizing it crudely, Hacker's point seems to be that simply stating that something must have been retained, does not explain anything — it is merely a way of saying that one did not forget. Furthermore, in the example, the aphasic *did* forget, which implies that saying that something has been retained is not only pointless, it is incorrect. But, even when accepting that there is something like retaining, it could never be the 'system of knowledge' postulated by Chomsky, for knowledge is recorded by means of a symbolism, something which the brain does not employ.

The picture described by Hacker, that of knowing as storing information in the brain, is indicative of what we will call the 'computer metaphor'. That is, we see ourselves participating in everyday activities, whether it be in using language, mathematics, or traffic, and think of ourselves as highly advanced computers capable of storing information and applying rules in order to be able to participate in these activities. We tend to think, 'Surely, something remarkable must be going on in the brain, if a child has to know all these rules to speak a language', not realizing that the sense in which a mechanism can be said to follow a rule is completely different from what we do in our linguistic activities. Chomsky's effort in rejecting the idea that knowing a language is the ability to speak a language, is precisely based on this confusion. In Chomsky's criticism of Skinner's theory of language acquisition, he emphasizes that Skinner's behavioristic view of knowledge of language and language learning cannot account for our choice in responding to other people's utterances, nor explain how it is that we have knowledge extending far beyond the training we have received — see (Cowie, 2010, Chapter 1). Chomsky wants to explain our knowledge and use of language in terms of propositional knowledge, as opposed to know-how, precisely because he wants to give an account of grammar that is capable of generating the linguistic expressions.

In the next chapter we will do some groundwork before we can disentangle this confusion. In the first section, we will reconstruct Chomsky's theory of Universal Grammar. In the second section, we will repeat some of Hacker's remarks on Chomsky's theory. In the third, and final, section, we will say more about the computer metaphor and show how it is present in Chomsky's theory.

Chapter 3

Chomsky's Universal Grammar

3.1   Universal Grammar

In the previous chapter we discussed the four arguments Chomsky gives in favour of rejecting the idea that knowledge of language is an ability or collection of abilities. We argued that he rejected that idea because it cannot account for two of the four questions, that form the "essential framework" for further inquiry into the study of language (Chomsky, 1988, p. 8). The four questions are (Chomsky, 1988, p. 3):

1. What is the system of knowledge? What is in the mind/brain of the speaker of English or Spanish or Japanese?

2. How does this system of knowledge arise in the mind/brain?

3. How is this knowledge put to use in speech (or secondary systems such as writing)?

4. What are the physical mechanisms that serve as the material basis for this system of knowledge and for the use of this knowledge?

Chomsky subdivides the third question into two aspects, namely the perception problem and the production problem (Chomsky, 1988, pp. 4–5). In this section we will focus on the second question, which Chomsky calls 'Plato's Problem', and on the second part of the third question, the production problem, which Chomsky calls 'Descartes's Problem'. Plato's problem is a problem concerning the acquisition of knowledge (of language), that is, how is it possible that human beings know as much as they know based on such limited experience? Descartes's problem is the production problem, the problem of accounting for the creative aspect of language use. That is, how can we account for the free and undetermined use of language if it can be reduced to mechanical processes?

3.1.1 Plato's Problem

Giving a solution to Plato's problem comes down to accounting for the discrepancy between the knowledge obtained and the data we, as human beings, are exposed to. The core of Plato's problem is the *poverty of stimulus* argument. It should be noted that when we gave an indication of the psychological framework in the introduction, Chomsky's theory was not yet committed to innate principles, or concepts.[1] Grammar might ultimately be something in the brain, but at this point this could still have grown from experience. This means that on the psychological interpretation, grammar must be learnable — and actually has to be learned — for any speaker of a language. This is where the poverty of stimulus argument comes in. The data (sometimes called 'primary linguistic data') is insufficient to construct a grammar. It is Plato's problem, and therefore the poverty of stimulus argument which gave rise to Chomsky's solution of innate knowledge.

Chomsky takes a 'modern version' of Plato's solution of anamnesis, namely, the idea that "certain aspects of our knowledge and understanding are innate, part of our biological endowment, genetically determined, on a par with the elements of our common nature that cause us to grow arms and legs rather than wings" to be "essentially correct" (Chomsky, 1988, p. 4). In his lectures, Chomsky considers varying rules of several languages, the interaction between these rules and principles determine "the form and interpretation of the expressions of the language" (Chomsky, 1988, p. 15). These rules and principles are, according to Chomsky, facts about language, thus the problem we face is one of determining how a child comes to master these rules and principles (ibid.). That is, what is the source of this knowledge? Chomsky discerns three factors: "the genetically determined principles of the language faculty, the genetically determined general learning mechanisms, and the linguistic experience of the child growing up in a speech community" (Chomsky, 1988, pp. 15–16). Chomsky's proposal, the "path we must follow", is a combination of the first and the third of the three distinguished factors — remaining silent on the second factor. The path is trying to solve Plato's problem in terms of (Chomsky, 1988, p. 17):

> certain properties of the mind/brain and certain features of the linguistic environ-
> ment. The properties of the mind/brain include several principles of the language
> faculty: the availability of complex constructions with an embedded clausal com-
> plement [...] The linguistic environment must be rich enough to determine the value
> of the parameter associated with the embedded clause property [...] This system of
> knowledge provides the interpretation of linguistic expressions, including new ones
> that the child learning the language has never heard.

Recall the distinction we mentioned in the previous chapter between the computational aspect of the language faculty (grammar) and the conceptual system. Acquiring knowledge of concepts is also susceptible to the poverty of stimulus argument. Since the rate at which a child

---

[1]Systematically, that is. Historically, Chomsky was long committed to the idea that grammar was innate.

learns these complex concepts is too high to be derived from the data the child is exposed to, a grasp of these concepts has to be innate as well. A child already intuitively grasps concepts, such as intention, volition, causation, and even technical concepts! Different languages have different words for these concepts, but the conceptual framework is, according to Chomsky, a "common human property" (Chomsky, 1988, p. 32). Thus, Chomsky concludes (Chomsky, 1988, p. 34):

> It seems that the child approaches the task of acquiring a language with a rich conceptual framework already in place and also with a rich system of assumptions about sound structure and the structure of more complex utterances. These constitute the parts of our knowledge that come 'from the original hand of nature,' in Hume's phrase. They constitute one part of the human biological endowment, to be awakened by experience and to be sharpened and enriched in the course of the child's interaction with the human and material world.

To strengthen this hypothesis, Chomsky asks us to imagine a Martian scientist, studying human language. In studying Spanish, the Martian scientist tries to reconstruct the rule used to transform declarative into interrogative sentences. He finds out, however, that speakers of Spanish do not use a simple rule, but a rule computationally more complex than the rule of his initial hypothesis. Plato's problem, in this case, is the question "How do Spanish speakers know that they are to use the computationally complex rule R-Q, not the simple rule based on linear order?" (Chomsky, 1988, p. 44). The Martian speculates that children are taught to use the complex rule, and children might also mistakenly use the simple rule until corrected. But, Chomsky concludes, the Martian will quickly discover that his speculations are wrong, children *never* make such errors (ibid.). The only plausible conclusion for the Martian will be that there are some innate principles in a human brain. The principles the Martian is trying to discover are already known by the child "intuitively, unconsciously, and beyond the possibility of conscious introspection" (Chomsky, 1988, p. 46).

This, in short, is Chomsky's solution to solve the (supposed) discrepancy between the obtained knowledge of a language, and the data to which we are exposed: learning a language is a combination of innate principles and concepts and experience used to determine which phonetic signs stand for the innate concepts, and which values should be used in setting the parameters in our language faculty.[2] In the last subsection we will give more attention to the relation between a spoken language and the faculty responsible for interpreting it, but first we will take a look at Chomsky's comments on Descartes's Problem.

---

[2]The later developed minimalist program goes even further and suggests that it might be the case that the sole task of language learning is acquiring a lexicon — see (Freidin, 2012, p. 910).

3.1.2   Descartes's Problem

After giving a short discussion of the development of Descartes's problem throughout the years, Chomsky states that it still stands (Chomsky, 1988, p. 147): "We still have no way to come to terms with what appears to be a fact, even an obvious fact: Our actions are free and undetermined, in that we need not do what we are 'incited and inclined' to do, and if we do what we are incited and inclined to do, an element of free choice nevertheless enters." The problem is that, given that the language faculty can be described in terms of physical processes, something needs to account for the element of free choice within this mechanical system. A solution of this problem, however, could simply lie beyond the scope of human capacity (Chomsky, 1988, p. 151): "One possible reason for the lack of success in solving it or even presenting sensible ideas about it is that it is not within the range of human intellectual capacities: It is either 'too difficult,' given the nature of our capacities, or beyond their limits altogether." Accepting this standstill for the time being, let us take a look at Chomsky's remarks on the properties of a language faculty in the next subsection.

3.1.3   Universal Grammar

In the previous subsections we have mentioned the language faculty, posited by Chomsky as a way of solving Plato's problem. But little has been said about what this language faculty is. Let us rectify that now. Roughly, the language faculty is a component in the mind/brain that takes environmental data, and forms a language from it: "a computational system of some kind that provides structured representations of linguistic expressions that determine their sound and meaning" (Chomsky, 1988, p. 60). A theory of the language faculty can be called 'universal grammar' (Chomsky, 1988, p. 61). Thus, universal grammar attempts to give an account, or rather *the* account, of the principles involved in the language faculty in its 'initial state', as Chomsky calls it (ibid.). In line with the parameters and values mentioned earlier, knowledge of a particular language would be a *state* of the language faculty, that is, a particular setting of several parameters, fixed by the environmental data of the particular language. Hence, there needs to be a sufficient amount of experience to set all the parameters (Chomsky, 1988, pp. 62–63). The difference between languages, say Dutch and English, can then be explained by reference to the different values for the parameters. As Chomsky stated: "We may think of the language faculty as a complex and intricate network of some sort of switches that can be in one of two positions. Unless the switches are set one way or another, the system does not function" (Chomsky, 1988, p. 62). In perceiving and understanding what one hears, the only thing required is that the parameters are set correctly, and the mechanisms (in the mind/brain) will operate "virtually instantaneously" (Chomsky, 1988, p. 73).

In analyzing particular languages Chomsky distinguishes several principles and categories. These principles and categories are of no interest to us, but let us briefly reflect on the components that have to be in place. On Chomsky's view the mind is modular, that is, the mind

consists of interacting but separate systems, which have specific properties (Chomsky, 1988, p. 161). The language faculty is one such system, which consists of certain parameters and principles and provides (generates) "paired representations of sound and meaning" (Chomsky, 1980, p. 92). Which interacts with the conceptual system, where the lexical items are connected to (innate) concepts. Note that the conceptual system is not actually part of the language faculty, but more or less our "common sense understanding" (Chomsky, 1980, p. 55). The conceptual system, which provides us with our understanding of concepts, is not a list, but is "like the sounds of language, they enter into systematic structures based on certain elementary recurrent notions and principles of combination" (Chomsky, 1988, p. 31).

The idea is then that Plato's Problem is solved because there is no discrepancy. Granted, the parameters in the child's language faculty need to be set, and this happens through experience, but the background system is already in place and capable of generating and understanding infinitely many sentences. Understood thusly, Descartes's problem is the problem of generating free but appropriate thoughts, which in turn get translated into speech. If this is what Chomsky meant, the problem indeed still stands, but does not threaten his theory of universal grammar as such, rather it is a problem about free will in general.

Chomsky concludes, lacking even the slightest sense of humility, that (Chomsky, 1988, pp. 91–92): "We are beginning to see into the deeper hidden nature of the mind and to understand how it works, really for the first time in history, though the topics have been studied for literally thousands of years, often intensively and productively."

3.2 Hacker's Reply

In the previous chapter we have already discussed Hacker's reply regarding Chomsky's considerations that are meant to show that knowledge cannot be seen as an ability. Apart from this reply, Hacker gives many arguments against Chomsky's theory of universal grammar. In this section, we will repeat his arguments concerning (1) Chomsky's confusion between the concepts of mind and brain, (2) the nature and workings of what can be called a faculty of the mind — specifically directed against the language faculty, (3) Chomsky's comments on Descartes's problem, and (4) Chomsky's solution to Plato's problem.

The Mind/Brain   Hacker's point regarding Chomsky's characterization of the concept mind — as some level of abstraction of unknown physical mechanisms of the brain[3] — is that it is mistaken. It is mistaken because our concept of mind, comparable to concepts such as intending, wanting, etc., is, contrary to our concept of brain, not a theoretical concept. That is, when speaking of one's mind "one is no more 'speaking at some level of abstraction of yet unknown physical mechanisms of the brain' than when one orders eggs and bacon for breakfast one is speaking at some level of abstraction of the chemical structure of proteins" (Hacker, 1990, p.

---

[3]As found on page 7, (Chomsky, 1988).

136). Speaking of mind as some level of abstraction of the processes in the brain is simply a category mistake, full stop.

Faculties   When talking about faculties, Hacker argues, one is talking about powers or abilities. Possessing a faculty means that one has the ability to do things which constitute exercising that faculty (Hacker, 1990, p. 136). Thus, Hacker continues, when Chomsky talks about the language faculty, he must be talking either about the ability to understand a language, or the ability to learn a language — a second-order ability — neither option, however, is part of the brain. Hacker accepts that one can uphold that creatures with a mind have a language faculty, but it is "to traverse the bounds of sense" when one claims that such a faculty contains grammatical principles (Hacker, 1990, p. 137). In support of his rejection, Hacker gives the same arguments as we saw in his response to Chomsky's third consideration, found in the previous chapter, namely that rules are formulated in a symbolism, while there is no such thing as a brain employing a language.

Furthermore, the language faculty, as characterized by Chomsky, is not intelligible. On Chomsky's view, the language faculty selects relevant data for theory construction, but, Hacker argues, "a faculty is not an agent who can make use of data" (Hacker, 1990, p. 138). Besides that, there is "not a shred of evidence" that a child analyzes what it hears and constructs a theory from the obtained data.

Descartes's Problem   In response to Chomsky's remarks concerning Descartes's problem, Hacker states that Chomsky dodges the original problem, that of explaining how, on Chomsky's account, one understands what one says, knows what one means, etc. These are real concerns for Chomsky's account, since the mind/brain should somehow compute what the person says. But how and when would the mind/brain do that (Hacker, 1990, pp. 139–140)? Chomsky's formulation is, according to Hacker, a different question completely. Furthermore, not only does Chomsky not answer the question, he states without reason that it might not be answerable at all. Confronting the original problem, would have given Chomsky cause to stop in his tracks, Hacker concludes (Hacker, 1990, p. 141): "Had he confronted it [the production problem, BK] candidly (and not confused it with philosophical questions about free will), had he disentangled the confusions implicit in it, the rest of his pseudo-mechanical mythology about computational processes of understanding would have sunk without trace."

Plato's Problem   Regarding Chomsky's solution to Plato's Problem, Hacker gives several remarks, three of which will be discussed. The problem Chomsky tries to solve is that of the discrepancy between the (linguistic) knowledge obtained, and the data that has been given, that is, a child learns a language (relatively) quickly, based on inadequate data.

Hacker remarks that the problem Chomsky mentions is puzzling, for it raises the question what it would look like if the child was given adequate data, that is, Chomsky does not make

clear what adequate data looks like. Furthermore, Hacker argues, a child does not learn a language based on data at all. The training and teaching it receives do not constitute evidence which the child employs in theoretical reasoning (Hacker, 1990, p. 142).

The second remark, made by Hacker, concerns the (supposedly) innate knowledge of concepts. According to Hacker, not only is positing innate knowledge not the only conclusion, it is even unintelligible. On Chomsky's account, a child only has to learn the labels for the concepts that are available before experience. But, Hacker argues, there is no such thing as a word being a label for a concept. Words can express concepts, but not label them, since that would require giving criteria for having prelinguistic concepts (Hacker, 1990, p. 142). Next to previously mentioned criteria, Chomsky's position also has to explain what it is to attach a word to a concept — something which cannot be done according to Hacker, since a concept is not a kind of object — and explain how a child can know to which (prelinguistic) concept he has to attach the word in question. And, granting that the child manages to attach words to these concepts, how would we know that he has attached it to the correct concept, for all we know is whether the word has been used correctly or not (Hacker, 1990, pp. 142–143). In short, nothing has been explained about what it means to have these prelinguistic concepts, and no acceptable reason has been given to show that the postulate of innate concepts is required for a child learning a language.

The last remark we will discuss, is concerned with Chomsky's conclusion that a child selects complex grammatical rules immediately, whereas the scientist can only discover it through a taxing process of inquiry. According to Hacker, there are four "interwoven" confusions. First, a linguist trying to formulate such complex grammatical rules already speaks the language, and while formulating such rules requires some effort, it is hardly a cumbersome task. Secondly, a child trying to learn a language is not concerned with formulating a grammar, nor is he translating utterances of an unknown language. If what the child says or asks is well formed, it will tautologically conform to the rules the linguist constructed — after all the rules are formulated to correspond to correct use of language — but that does not show that he has made use of these rules. Thirdly, it does not make much sense to say that a child selects the rule at once, since it is unclear what would count as selecting the rule, nor do we know what the child could do with a rule of which he cannot even understand the formulation. Lastly, Hacker wonders if it is really puzzling that some sentences are correct while others are not. According to Chomsky it ought to be. If that is true, we could just as easily wonder how the child — if he needs a rule to form sentence *a* out of *b* — forms *b* out of *a*, which would require an additional rule, but, Hacker states, "speaking, asserting, asking questions, etc. are not computational processes of *any* kind" (Hacker, 1990, p. 146).

3.3 Computer Metaphor

In the previous sections we gave a partial reconstruction of Chomsky's theory of universal grammar, where a universal grammar is a theory of the language faculty, independently of any specific language. The language faculty, the intricate system of 'switch boxes' in a specific setting, is a mental state, which is knowledge of a language. This faculty is supposed to solve what Chomsky calls 'Plato's problem', i.e. the problem of accounting for the rapid and extensive growth of knowledge (of language in this case), based on inadequate data. In short: explaining the apparent discrepancy between the knowledge obtained and the data available for obtaining this knowledge.

Hacker argued against Chomsky's theory on several points, some of which have been mentioned in the previous section. Here we will briefly reflect on three of Hacker's comments. Firstly, Hacker states that it is not clear what adequate data would look like, and that therefore Chomsky's remark is puzzling, next to being incorrect since the child does not learn a language based on data at all. The first part of Hacker's comment is a remark on the poverty of stimulus argument. This argument has been criticized on many accounts already. To give a few examples, it has been argued that it is unclear what it precisely is that has to be learned, and hence has to be among the linguistic data a child is exposed to. Nor is it clear what the linguistic data contain, and therefore lack. And even if one grants that something has to be innate, it does not follow that it has to take the form of Chomsky's universal grammar — see (Cowie, 2010, §2.2) for an overview of the criticisms against the poverty of stimulus argument. But it is the second part of Hacker's comment that is of interest to our criticism of Chomsky's theory. Granted, Hacker's formulation is somewhat unfortunate, since he talks about theoretical reasoning, while the language faculty as envisaged by Chomsky does not reason at all — it is something that happens to the child. But, no matter in what way it is formulated, ultimately it is something in the brain interpreting the data.

Secondly, Hacker remarks on Chomsky's explication of Descartes's Problem that Chomsky misinterprets the question. Rather the question would be: how does one understand what one says? And Hacker ironically remarks that it raises the question of when the mind/brain of a speaker computes what he says, for example before or after he has spoken. Although we will argue in chapter 5 that Chomsky's confusion arises because he tries to explain language in terms of interpretation, which ultimately lines up with Hacker's conclusion, Hacker's argument regarding Descartes's problem seems to be mistaken. On Chomsky's account, a person would 'know' what he wants to say, because he has a certain configuration of ideas/concepts that he wants to utter. Understanding what someone says is the generation of a mental representation from speech (or written signs), but the other way around the configuration of concepts would be the starting point, which makes Hacker's ironic remark rather off point.

We did say that we will ultimately line up with Hacker's conclusion, but not directly for the same reason. This brings us to our third point. Hacker's argument is that rules cannot be in the

brain, since rules are expressed by means of a symbolism, a language, while the brain does not employ a language. Although we agree that the brain does not employ a language, this is not the strongest argument against the possibility of rules being in the brain. Furthermore, Hacker's point seems to be that rules need to be expressed in a symbolism one can understand, before one can follow a rule. This point seems mistaken, and we will give an argument against it in the next chapter.

For the remainder of this thesis, we will primarily focus on what we have dubbed the 'computer metaphor' and the confusion that lies beneath it. The computer metaphor, in its core, is the tendency to evaluate our activities — such as learning and speaking a language — which (seem to) consist of a system of rules, and imagine our brain as the 'central processing unit' processing these rules and generating correct output.

That Chomsky falls prey to the tendency of looking at the brain as such a processing unit, was apparent throughout our reconstruction: the language faculty is a component of the mind/brain, one that "forms a language, a computational system of some kind that provides structured representations of linguistic expressions" (Chomsky, 1988, p. 60), which can be seen as a "complex and intricate network of some sort associated with a switch box consisting of an array of switches that can be in one of two positions" (Chomsky, 1988, p. 62). It should be noted that this tendency is not exclusive to Chomsky's theory. Copeland points out that both Lashley and Searle already remarked on the historical tendency to model the brain on modern technology — from a hydraulic theory of the brain in Descartes's time to the conception of a brain as a telephone switchboard (Copeland, 1993, p. 182).

Little has been said, however, as to why this way of looking at ourselves is problematic, in the remaining part of this thesis we will take on this task. We will start by reviewing and discussing some of Wittgenstein's remarks on the subject in the following chapter. This will give us the means to show why the computer metaphor is deeply confused.

Chapter 4

Wittgenstein on Rule-Following

4.1 Meaning as use

In the *Investigations* Wittgenstein states, after criticizing the idea that words *only* name objects, that meaning is use. In §43 he says (Wittgenstein, 2009, p. 25): "For a *large* class of cases of the employment of the word 'meaning' [...] the word can be explained in this way: the meaning of a word is its use in language." After several remarks on the role of philosophy, he takes up the question how we are able to *understand* the meaning of a sentence. If the meaning is something we "grasp at a stroke" then it must be something different from use, since use is extended in time (Wittgenstein, 2009, p. 59, §138).

Wittgenstein then continues to question the legitimacy of the interlocutor's proposed solution to the question: what do we precisely grasp, when we grasp the meaning at a stroke? In the paragraphs following 138, he states that it cannot be a picture, since the picture itself could apply to various objects — he discusses the picture of a cube, which could also apply to a triangular prism (§139). Would we still be inclined to say that the picture has the same meaning in both applications? If we want to speak of applying a picture, then *that*, the application, is the criterion of understanding (§146). The question that Wittgenstein's interlocutor then raises is: "When *I* say that I understand the rule of a series, I'm surely not saying so on the basis of the *experience* of having applied the algebraic formula in such-and-such a way" (Wittgenstein, 2009, p. 64, §147). That is, our knowledge of, for example, algebraic formulas cannot be mere past experience — our past application of numbers is not the criterion — since what we developed is finite, but the series is infinite (§147). What follows is a discussion of the nature of knowledge, where Wittgenstein discusses several explanations which do not suffice. To start with knowledge as a state of the mind. Wittgenstein criticizes this idea, since a mental state, like excitement, has a genuine duration. While in the case of knowledge, we do not speak about duration. He says (Wittgenstein, 2009, p. 65): "Compare: 'When did your pains get less?' and 'When did you stop understanding that word?'" Further expanding on our use of the word 'knowing', Wittgenstein considers: "Now I know", an exclamation made by someone

who can now follow a certain mathematical rule. This being able to continue, is then something that "occurs in a moment" (Wittgenstein, 2009, p. 65, §151). But what is it that occurs? Is whatever processes went on in the mind of the pupil — say a feeling of tension — what we call 'understanding'? The problem Wittgenstein is addressing, is that we continue to look for mental processes that are supposed to explain what happens when we understand something. But none of the phenomena satisfy, it is not enough, for example, that a formula should occur to him. Since a formula can occur to someone without him understanding it — just like we can recall a puzzling paragraph of the *Investigations* without understanding it. But even if we were to find a process that occurs in all cases of understanding, we still do not know if that process is the understanding — it might be nothing more than a concomitant process (Wittgenstein, 2009, p. 66. §152). Wittgenstein concludes that it is precisely this search for mental processes in thinking of understanding that confuses us (§154), and, he concludes again, we should look at the particular circumstances under which it is correct to say that you understand it (§155). Wittgenstein extends the discussion to include reading, and here, as in the case of understanding, referring to mental processes, experience of influence, derivation and causation, is not of any use in explaining what we call 'reading'.

Since all these criteria do not suffice, this raises the question: "What is required to carry out an order correctly?", which the interlocutor describes as carrying the order out as it was meant (§186). The idea behind this, the idea that Wittgenstein criticizes is that by meaning something, that which was meant seems to be "in some *unique* way predetermined, anticipated — in the way that only meaning something could anticipate reality" (Wittgenstein, 2009, p. 82, §188). The 'grasping the use at a stroke' (§191), is something for which we have no model — it is a mode of expression which suggests itself to us. In §193 he gives an analogy about a machine as a symbol (Wittgenstein, 2009, p. 83–84, §193): "A machine as a symbol of its mode of operation. The machine, I might say for a start, seems already to contain its own mode of operation." Which Wittgenstein explains to mean: "we are inclined to compare the future movements of the machine in their definiteness to objects which have been lying in a drawer and which we now take out." The idea is that a machine, the steps it will take, is already completely determined, and when we know how the machine works, we know which future movements it will make. When we then continue to reflect on using the machine to symbolize this movement, and realize that it could have moved differently, it seems that the mode of operation must be present in our symbolized machine more determinately than in our actual machine. Being empirically predetermined is not enough, it has to be "really — in a mysterious sense — already *present*" (ibid.). In a mysterious sense because the possibility is something present in the machine movement, but is neither the movement itself nor the physical condition for moving (§194). Just like 'grasping the use at a stroke' we do not have a model for this talk about possibilities, it is just a way of expressing ourselves. The point Wittgenstein is trying to make is made more explicit in §197, where he states that describing what we do as 'grasping the whole use at a stroke' only becomes astonishing or strange when we think that in our grasping

the use of a word the future development must somehow be present, but is not yet present. But, he continues, where is the connection between the meaning of an expression and the future development of the expression? The answer is in our everyday practice (§197).

The underlying problem which has been the focus of this section, is that of interpreting a rule. From understanding as a criterion for grasping the meaning of a word, to mental processes and phenomena satisfying what it is to read. We were looking for prelinguistic criteria that would secure the determinateness of following a rule. The problem Wittgenstein addressed is made explicit in §198 where he states that interpretations do not determine meaning, ultimately leading to the rule-following paradox (Wittgenstein, 2009, p. 87, §201): "no course of action could be determined by a rule, because every course of action can be brought into accord with the rule. The answer was: if every course of action can be brought into accord with the rule, then it can also be brought into conflict with it. And so there would be neither accord nor conflict here." The misunderstanding stems from the idea that every action according to a rule requires interpretation. If we give up the idea that following a rule requires interpretation, and see that following a rule is a practice, that would bring a hold to the infinite regress that is invoked using prelinguistic entities to justify our use.

In the paragraphs following §201, Wittgenstein keeps repeating the point: following a rule is a practice (§202). It is akin to obeying an order. One is trained to react in a particular way (§206). We can only speak of a language if there is enough regularity (§207). Following a rule does not require intuition (§213), there is no choice involved in following a rule — "I follow the rule *blindly*" (Wittgenstein, 2009, p. 92, §219). The regularity in language is agreement, which is not agreement in opinions but in form of life (§241). This last point is made more explicit in §242, and, as we will see shortly, has implications for the objectivity of rule-following. Let us begin by quoting the paragraph in full (Wittgenstein, 2009, pp. 94–95, §242):

> It is not only agreement in definitions, but also (odd as it may sound) agreement in judgements that is required for communication by means of language. This seems to abolish logic, but does not do so. — It is one thing to describe methods of measurement, and another to obtain and state results of measurement. But what we call 'measuring' is in part determined by a certain constancy in results of measurement.

At first glance, this paragraph raises a couple of questions. For example, agreement in judgments is required for us to communicate at all, thus we require agreement to obtain a method of measurement, but simultaneously the result of our measurement in turn determines what we call measuring. It then seems that we 'have to meet up' and agree on what the correct result of a rule should be, but Wittgenstein already said that neither intuition nor agreement in opinions is required (in §213 and §241). It appears, to stress the point, that empirical considerations play a role in determining a practice and further shaping that practice. In order to clarify Wittgenstein's point, we will take a look at his *Remarks on the Foundations of Mathematics* where he

distinguishes between empirical and mathematical propositions.[1] He says (Wittgenstein, 1978, VI-16, pp. 319–320):

> A rule is surely something that we go by, and we produce one numeral out of another. Is it not matter of experience, that this rule takes someone from here to there?
>
> And if the rule +1 carries him one time from 4 to 5, perhaps another time it carries him from 4 to 7. Why is that impossible?
>
> The question arises, what we take as criterion of going according to the rule. Is it for example a feeling of satisfaction that accompanies the act of going according to the rule? Or an intuition (intimation) that tells me I have gone right? Or is it certain practical consequences of proceeding that determine whether I have really followed the rule? — In that case it would be possible that 4+1 sometimes made 5 and sometimes something else. It would be thinkable, that is to say, that an experimental investigation would shew whether 4+1 always makes 5.
>
> If it is not supposed to be an empirical proposition that the rule leads from 4 to 5, then *this*, the result, must be taken as the criterion for one's having gone by the rule.
>
> Thus the truth of the proposition that 4+1 makes 5 is, so to speak, *overdetermined*. Overdetermined by this, that the result of the operation is defined to be the criterion that this operation has been carried out.
>
> The proposition rests on one too many feet to be an empirical proposition. It will be used as a determination of the concept 'applying the operation +1 to 4'. For now we have a new way of judging whether someone has followed the rule.
>
> Hence 4+1=5 is now itself a rule, by which we judge proceedings. This rule is the result of a proceeding that we assume as *decisive* for the judgment of other proceedings. The rule-grounding proceeding is the proof of the rule.

It is interesting to note that what we discussed previously — for example when we wondered what is essential to reading — is also mentioned and rejected here. But, more importantly, there is a point made in distinguishing empirical from mathematical propositions. Namely, that mathematical propositions are overdetermined. That is to say, in having the result (of an application of a rule) as a criterion of carrying out the rule, empirical facts are then excluded from contradicting mathematical propositions. For example, the mathematical proposition $2 + 2 = 4$ does not have to be altered if someone would have 5 as an answer, but the empirical proposition 'If you add 2 and 2 you get 4' can be falsified if someone gets 5 as an answer. It can be falsified precisely because it was an empirical proposition, a prediction in this case. Thus,

---

[1]The following remarks on Wittgenstein on mathematics are based on an earlier paper I wrote for the tutorial Intuitionism. The paper is titled "Dummstein: A Critical Reflection on Dummett's Interpretation of Wittgenstein".

again, the idea is that there is nothing external to mathematics that could falsify mathematical statements.

This, however, still leaves our main concern untouched. Namely, it seems that empirical considerations do play a role in shaping our practice — it appears that we have to agree in what result we ought to get if we add 2 and 2. Thus, despite the mathematical statement being overdetermined, empirical considerations do play a role in determining which mathematical statements we uphold. And Wittgenstein indeed says that empirical regularity lies behind a mathematical law, as found in *Lectures on the Foundations of Mathematics* (LFM) (Diamond, 1976, pp. 291–292):

> Now is it an empirical fact that $\phi(3000) = \psi(3000)$? — We *make* the rule that $\phi(3000) = \psi(3000)$ *because* of the agreement in action — namely, that if we went through these steps we would nearly all get the same results. And this rule then becomes a standard of measurement. The rule doesn't express an empirical connexion but we make it because there is an empirical connexion.

Now, this passage can be understood in two different ways: either empirical considerations play a role in determining our practice, a mathematical system, or empirical considerations play a role within our mathematical system, they are in a position to justify or falsify mathematical statements. Wittgenstein explicitly denies the role of empirical considerations within a mathematical system — there is no room for interpreting the rule. The role of empirical considerations is to point out a certain *practice*, namely our mathematical language-game. It is in virtue of our language-game that we are able to practice mathematics at all. A lack of an empirical regularity would then not falsify a mathematical statement, rather there would be no mathematics. This is the difference between a negation within a system, falsifying a certain statement or result, and negation of a system, denying the system itself. As Gerrard puts it: "It is one thing to say that we would not have our mathematics if some empirical considerations had not obtained (true to Wittgenstein), and quite another to say that the propositions of mathematics *state* these empirical considerations or can be falsified by empirical considerations (false to Wittgenstein)" (Gerrard, 1996, p. 185).

This answers our main concern regarding §242. Our concern was whether this agreement in judgment was based on our opinions, or intuitions regarding which result we should count as correct. If so, choice would govern our practice, and meaning would become contingent. Returning to Wittgenstein's remarks on mathematics, it is true that mathematics as a whole is contingent and indeed not responsible to anything. But it is not the case that, given our mathematical practice, we can choose to hold as necessary whatever statement we like, it is not the case that we are not responsible to anything within our mathematical practice. As Wittgenstein explains in LFM, we have two kinds of responsibility: mathematical responsibility where a proposition has a responsibility towards other propositions, and the responsibility of the system as a whole (Diamond, 1976, pp. 240–244). The second kind of responsibility only refers

to the presence of a practice. It is then very clear that within a mathematical system, there would be no room for deciding which statements are necessary. For Wittgenstein the right answer is already secured by the rule, the answer is, as we saw, overdetermined. But still, one might say, we do choose the right answer through experiments — namely in determining the practice. Wittgenstein responds, however, that we do not choose anything: "Mathematical truth isn't established by their all agreeing that it's true — as if they were witnesses of it. *Because* they all agree in what they do, we lay it down as a rule, and put it in the archives. Not until we do that have we got to mathematics. One of the main reasons for adopting this as a standard, is that it's the natural way to do it, the natural way to go — for all these people" (Diamond, 1976, p. 107). We have a certain practice but not by choice, we simply do it in such-and-such a way.[2]

Returning to the more general case of meaning and rule-following. If neither meaning nor understanding can be understood as interpretation, but are evaluated in use, the same goes for rule-following. Rules can only be followed insofar as there is a practice. The remarks discussed in the previous paragraph are remarks about the origin of such a practice, or perhaps the requirements for there being a practice. Within a practice we can speak of correct or incorrect use, and there it is not the case that anything goes. The practice itself however, lacks such normativity. The presence of a practice is not a matter of choice, but something that comes naturally. Two things can be distinguished within such a practice. First of all, it requires a certain constancy, a practice would lose its point if confusion prevailed. Furthermore, a practice has to have a purpose. That is to say, the language game must have a certain function, we have to do something with the technique. This becomes clear when Wittgenstein asks: "Does a calculating machine *calculate*?" (Wittgenstein, 1978, p. 257, V-2). To which Wittgenstein answers that it is essential that "its signs are also employed in *mufti*" (ibid.). The calculator does not calculate, it is us calculating with the calculator.

Recap In this section we have reconstructed Wittgenstein's notion of rule-following, starting from the idea that meaning is use. We started with his criticism of the search for prelinguistic entities which secure meaning, whether it be a picture, a mental state or a mental process. We have seen that neither one of these suggestions is of any use in explaining our linguistic activities. The suggestion, that meaning is use, has been subject to the same line of inquiry: what is required to carry out an order correctly? What is it we understand when we grasp the meaning at a stroke? The problem the interlocutor raised is that of interpreting the rule, and his search for criteria of interpreting the rule, is what ultimately led to the rule-following paradox:

---

[2]Wittgenstein's remarks on mathematics raise further questions. We are in the process of reconstructing Wittgenstein's remarks on rule-following, and the characterization in his remarks on mathematics is applicable for the case of natural language, on the Wittgensteinian view. But his remarks as a position in the philosophy of mathematics is not uncontroversial. For example, platonists in the philosophy of mathematics would disagree with much that has been said in the previous paragraphs. One can even agree with much of what Wittgenstein has to say about natural language, while still upholding such a strong metaphysical view of mathematics. This need not concern us too much. The point is that we are reconstructing his notion of rule-following, not giving a defense of his view on mathematics.

namely, that there is no rule-following at all, since all actions can be brought into accord and likewise into conflict with the rule. Wittgenstein's response is that the problem lies in thinking that following a rule requires interpretation. Instead of searching for criteria for interpreting rules, we should realize that following a rule is a practice. It is a system of shared human behavior, where we act in a certain way — a natural way to go. This is where language starts, and because we all at large agree, we can adopt it as a standard, lay it down as a rule.

4.2 Interpretation

We have given an interpretation of Wittgenstein's account of meaning and rule-following, and saw how rule-following is only possible within a practice. What we have reconstructed involved a transition in the status of the rule. On our interpretation the transition is (coarsely put) from that of having a rule first and foremost that enables us to communicate thereafter, to having shared human behavior — already enabling us to communicate in a broad sense — which enables us to adopt standards.[3] There remains, however, a certain tension between a practice and a rule. It has been clear so far that there can only be rules, language, etc., within a practice/custom. Taking for granted that we have a certain shared technique, including rules. The contingency lies with the practice as a whole, but within the practice there is responsibility towards the rules. When it comes to applying our technique, let us say applying our technique of multiplying numbers to numbers so large that the proof will be unsurveyable — where the technique cannot be applied, i.e. where there is no internal relation between the terms — then a new decision is required, it has ceased to be multiplying. This decision is then an experiment of expanding the technique to include these numbers, after which it can be adopted as a calculation — we stipulate the internal relation. This stipulation is, however, not by choice. We do not invent the correct result but *find* that this is the correct result (Diamond, 1976, p. 101).

Now, if we find out by experiment what we all at large do, why do we need rules at all? After all, we have our answer already, and if it was lacking — there was no agreement at large — we could not arrive at rules anyway. The answer that we gave in the previous section, concerned the status of a rule. Empirical propositions had to be revised if there was an empirical fact contradicting it, mathematical propositions — and on our interpretation, by extension all rules — have an internal relation. The idea then is that by 'putting it in the archives', we create the standard of judgment, a normative standard that our agreement did not possess.

There is some merit to this interpretation. By taking agreement in judgment as a starting point, we have our natural way to act. That is, this gives an explanation of Wittgenstein's remarks that following a rule does not require interpretation or choice, but that we follow it blindly. We have put what came natural in the archives, and this is now the way we must do it, we are trained to apply a technique in dealing with situations. That is why we ourselves do not know more than the examples we can give (Wittgenstein, 2009, §209, p. 89). We can now train

---

[3]To clarify: a standard would then be a norm, a rule in a broad sense, which can only exist in a practice.

someone to follow a rule, namely to act in such and such a way. A short point of clarification is in order. If taken this way, the same problem of interpretation seems to appear. How are we going to apply whatever we put in the archives? In Wittgenstein's words (Diamond, 1976, p. 105): "We have the metre rod in the archives. Do we also have an account of how the metre rod is to be compared with other rods? [...] Couldn't there be in the archives rules for using these rules one used? Couldn't this go on forever?" The point is, according to Wittgenstein, that we recognize what is put in the archives, that is, that we are already acquainted with the technique. This putting in the archives only makes sense if the technique is already present, and can now be fixed to be used for further instruction, for example the teaching of pupils.

Our interpretation also offers a clear explanation of the machine analogy mentioned in the previous section. If we forget that our rules presuppose our customs, and look at how our rules could be interpreted differently — reflect on the possibility of different usage — we are deceived in thinking that our rules must contain our future use in a stronger sense than merely having shaped our way of comprehending things. We start looking for additional information that tells us how we are supposed to use the rule. It is no longer enough that the machine as symbol shows us a technique, it now has to show us all its applications as well.

On our interpretation meaning as use could still be objective, since it is within our practice that rules are interrelated, and responsible towards one another.[4] But also normative, because the practice determines how we have to approach certain situations — the technique shows how we have to judge. The point is that questions about interpretation can only be asked within our language. It is only because we have customs that we can talk about these things, and precisely because we can communicate, we can also invent new games and techniques.

Before we take up the task of pointing out Chomsky's confusion regarding the notion of rule-following, we will first, briefly, reflect on Hacker's criticism of Chomsky in the next section.

4.3   Hacker's Notion of Rule-Following

As mentioned, Hacker criticizes several points of Chomsky's theory. And parts of Hacker's criticism line up with Wittgenstein's remarks on language. For example, Hacker lines up with Wittgenstein when he states that the criteria of a person knowing certain things is exhibited in his behavior (Hacker, 1990, p. 128). As well as the language acquisition of a child — that it is something learned through training. Furthermore, Hacker's argument against knowledge being a mental state — that it lacks genuine duration (Hacker, 1990, p. 133) — seems to be directly derived from Wittgenstein, when he asks (Wittgenstein, 2009, §148, p. 64): "Let me ask: *When* do you know that application? Always? Day and night?" Hacker's notion of rules, however, deviates from Wittgenstein's, and it is precisely on this point where Hacker could have strengthened his criticism. That is, by focusing on the expression of the rule — which, according to Hacker, needs to be formulated in a symbolism — he misses the point that there is

---

[4]That is to say, it is not something to which opinions are relevant.

simply no room for interpreting the rule, and that that is precisely the problem for Chomsky.[5] Furthermore, Hacker seems to uphold a rather strange notion of what is needed for one to be able to follow a rule (Hacker, 1990, p. 137):

> However the supposition that it makes sense to talk of a creature's knowing the rules of a language before it has learnt a language is misguided. For rules, like propositions, must be formulated in a symbolism. One can only be said to know a rule if one can understand its articulate expression, and one can only be said to understand a rule if one knows what *counts* as acting in accord with it. Hence a rule that cannot be formulated in a symbolism which the follower of a rule can understand is not a rule which such a creature can follow. A creature can only be said to follow a rule in so far as it uses the rule as a guide to, standard of, and justification of its conduct.

These comments are part of the section where Hacker aims to clarify Chomsky's confusion by explaining our concept of faculties. The point Hacker is trying to make is that it is senseless and incorrect to speak as if the child knows any grammatical rules while it would not even recognize any formulation of these rules. The argument, however, seems plain wrong. Of course if the training is complete then a speaker should be able to recognize expressions of the rules he learned. But that is nothing more than being able to speak the language — if you have mastered a language, then you will recognize well formed sentences by definition. It is in no way necessary for a rule to be formulated in a symbolism before one can act upon it. If examples are needed, take shaking your head to mean the same as saying: "No." Of course being able to translate it to the symbolic representation you have just read, means that we can give descriptions of the rule, but such a description is in no way necessary for the rule to be effective. A child learns many rules before it is able to read or write — or even speak for that matter. Maybe Hacker is suggesting that the technique must be 'put in the archives', but even so, why should that be a symbolic representation? It would be interesting to see Hacker give a symbolic expression of the rule for correctly using the word 'blue' — after all, reacting in a certain way to that colour is following a rule.

The more problematic issue is that Hacker seems to miss the most important part, namely that following a rule does not require interpretation. By stipulating that we need to understand the expression of a rule, and pointing out that such an expression cannot be found in the brain, he seems to leave the door open for some future neo-Chomskyan to say that it is not a symbolic representation that is found in the brain, but a certain pattern in a neural network.[6] While the important point is that rules simply cannot be reduced to mechanical processes no matter in which way these processes are disguised.

---

[5]Point of clarification: We can interpret rules, namely by substituting different expressions of a rule, for example for one who is not familiar with the concepts used in the first expression. The point is that we already need to know what it is to follow a rule before it makes sense to talk about interpreting the meaning of something.

[6]Comparable to switching from Von Neumann architecture to Connectionism.

Chapter 5

Computer Metaphor

From the previous chapters a few things become clear. The Chomskyan mental representation of sentences was explicitly rejected by Wittgenstein. The meaning of a word, the acquiring of a lexicon in Chomsky's theory, was disputed by Wittgenstein on the account that it invokes a vicious regress of interpretation of pictures — if knowing the meaning of a word is having the correct picture, then we require a further rule of applying that picture, etc. Rather, on the Wittgensteinian account, meaning is use. Furthermore, for the Chomskyan, external language, as infinitely many sentences, can be known because the underlying principles (the grammar and the conceptual system) are finite, and can generate infinitely many sentences. This, we take it, is the principle of compositionality, and explains how we are able to understand and produce infinitely many sentences. Namely, because the meaning of sentences is determined by the meaning of the words, and these words in turn are linked to the conceptual system, the concepts of which "enter into systematic structures based on certain elementary recurrent notions and principles of combination" (Chomsky, 1988, p. 31). On the Wittgensteinian account this picture is incorrect. For one thing, it neglects the multiplicity of ways in which words are used. We can use language to express thoughts, but that is not the only way in which we use it — that is why Wittgenstein compares the function of words with tools in a toolbox (Wittgenstein, 2009, §11). Furthermore, on Chomsky's account the origin of productivity is found in the individual speaker, he (or more precise his language faculty) is capable of generating infinitely many sentences. But on the Wittgensteinian account, language is only possible if there is enough regularity, if there is a practice. Hence it cannot be the case that a speaker of a language, being very intelligent or something of the like, can produce sentences which are not recognized as such by the community. That becomes clear when Wittgenstein says (Diamond, 1976, p. 104): "There is nothing there fore a higher intelligence to know—except what future generations will do. We know as much as God does in mathematics." The point is, that there is not something beyond the practice, as if we were only hindered by our cognitive capacities. The first comment is part of Wittgenstein's criticism of the Augustinian picture of language, and part of his argument that meaning is use. This has implications for what it is to know the rules of language, it is of no use, as we saw, to think of knowledge of rules as some sort of mental interpretation of them.

In chapter 3 we have mentioned how Chomsky's notion of the language faculty, and universal grammar as a theory of that faculty, is confused about what it is to follow a rule. We have said that Chomsky falls prey to the computer metaphor. That is, the tendency to think of the brain as some sort of processor processing rules to generate correct output. Now, the confusion is conflating the rules of a certain game with processes in the brain. While whatever goes on in the brain is merely an uninteresting byproduct — it makes no difference in our practice. Of course this is not to deny that something goes on in the brain, it is merely to say that whatever goes on in the brain does not tell us anything about rules in language, or any other game. The rules of language simply cannot be found in the brain. In the first section we will try to show how this confusion arises, that is, why we are inclined to explain rule-following by reference to brain processes. In the second section we will argue why such an enterprise is mistaken. After which we will show how this confusion undermines Chomsky's theory in the third section.

5.1 Mechanical Rules and Following a Rule Mechanically

We have already stated that it is confused to reduce rule-following to brain processes. Keep in mind that this tendency, to look at our brain as a mechanism in order to explain our behavior, is not exclusive to Chomsky's theory. In the philosophy of Artificial Intelligence, there are, for example, proponents of what is called 'the strong symbol system hypothesis', the hypothesis that only universal symbol systems can think, that is, every being that thinks is in essence nothing more than a universal symbol manipulator containing an expressive language to represent objects, a large knowledge base of the world, and input devices that obtain symbolic representations from stimuli. Thinking would then be nothing more than applying operations with the stimuli and knowledge base as input to create output (Copeland, 1993, p. 80, 182).

We will give further arguments to show that it is indeed confused in the next section, here we are interested in a possible explanation for our tendency to evaluate rule-following thusly. The confusion arises by conflating mechanical rule-following with following a rule mechanically.[1,2] As Wittgenstein says: "One follows the rule *mechanically*. Hence once compares it with a mechanism. 'Mechanical'—that means: without thinking. But *entirely* without thinking? Without *reflecting*." The problem is that we, trained to act in a certain way, on reflection try to explain how it is what we did. But since there was nothing special of which we were conscious, we are lured into thinking that it must be something that happens in our brain. This is the comparison with a mechanism. We are thinking about ourselves as both human beings in our everyday activities and simultaneously as natural systems. This might seem intentionally

---

[1]The following remarks on the distinction between mechanical rule-following and following a rule mechanically are based on an earlier paper I wrote for the course Philosophy of AI. The paper is titled "Wittgenstein on Turing's Notion of Computability".

[2]I owe a great deal of my interpretation of Wittgenstein's remarks on the distinction between the two notions to the first chapter of Shanker's *Wittgenstein's Remarks on the Foundations of AI* where he comments on precisely this confusion (Shanker, 2006).

obscure, but it is not — at least not intentionally. We can communicate because we at large do the same, we agree in actions and judgment, the techniques we use, in which we are trained, are ways of dealing with situations. Ways of comprehending them. The idea is that the entire notion of rule-following only has meaning within our form of life, and it is within our customs that we speak of following rules. The problem is that we take these notions and try to explain them in terms of ourselves as natural systems. This is how the idea of us following the rule mechanically gets translated to mechanical rules — surely that is what must happen in the brain. We expect to find corresponding processes, not realizing that it is only within our practice that the words 'knowing', 'understanding', 'meaning' have meaning at all.

5.2   Rules and Processes

Let us start with examining what it is that we do when we follow a rule. In our previous chapter, we saw that a rule is part of a technique, and a technique is a way of approaching a situation — a manner of conduct. The reason we can communicate at all, is because we have a shared custom, and following a rule is only possible within such a framework. It would become senseless to speak of a rule if there is no shared practice. Senseless because it would lose its function, it is no longer a guide for conduct for there is no criterion of judging. That is why regularity in action is not enough in order to speak of following a rule, it has to be part of a practice, but rules as brain processes can only be described in terms of regularity. The point is that the rule serves as a justification for your action, there is a certain *must* in following the rule — the internal relation between the terms as discussed earlier. With processes in the brain, however, it makes no sense to speak of justification — it does not make sense to ask why certain neurons fired — for there is no must, it can only be described in terms of what happens. The proposition is empirical, while the notion of a rule cannot be empirical, by definition — an empirical proposition can never serve as a guide for conduct.[3]

This is why even complete knowledge of the brain could never tell us what, for example, reading is. For reading, understood as a technique, can only be explained through reference to our practice, and to say someone is reading can only be understood by examining his actions being part of the practice. Reading can only be attributed to a person in the light of his behavior, his reaction to written signs. Reading is then the practice of reacting to certain signs in *that* way, and it would not be reading if he did not react in that way. Mental processes simply play no part in our use of the word 'reading'. But here, again, we should be careful not to confuse the

---

[3]A similar point is made in Stein's *The Fiber and the Fabric*, where he argues that Chomsky's notion of knowledge of a language, as a particular brain state, is empty as soon as one discards external languages as the object of such knowledge. And without any such object, the language faculty cannot even be classified, for we threw away our possibility of referring to the intended function — see (Stein, 1997, chapter 3.VI). Which means that we no longer are in a position to judge whether the output is as intended, or whether the mental state is 'properly configured'. This, I take it, comes down to roughly the same point we made, namely that on Chomsky's account we can only describe whatever goes on in the brain in terms of regularity, there is no longer a criterion for correctness.

description 'reacting in such-and-such a way' with the prescriptive nature of a rule. Because it is not the case that the sum of all descriptions can give us necessary and sufficient conditions for following a rule.

All of this is not to say that nothing happens in our brain, or that our constitution is completely irrelevant for our actions and forms of life. Our practices would not be the same if we were shaped differently. Imagine human beings being more limited than they are, and as being able only to comprehend sentences no longer than four words, or human beings lacking the ability to see. These limitations would have tremendous effects on our way of life. The point is that although our constitution is a prerequisite, ultimately the meaning of words, our notion of understanding, and so on, are found in what we do — in our practice. What this section is meant to point out, is that even though there might be certain brain processes concomitant to certain actions, following a rule could never be reduced to brain processes since they lack the normative aspect of rule-following. In the next section we will see the problems Chomsky's theory faces, precisely because it does reduce rule-following to brain processes.

5.3   Chomsky's Rules

As remembered Chomsky's aim was to solve what he calls 'Plato's Problem', that is the discrepancy between the linguistic knowledge obtained and the environmental data that is available. His solution, which is according to him the only path that can be taken to solve this problem, is a combination of (genetically determined) principles in the mind/brain and linguistic experience. We have a language faculty, a mechanical system where grammar is in place, combine this with our conceptual system and we only need linguistic experience to set the switches of this system. The discrepancy is dissolved, because, as it turns out, all the knowledge is already there. It just needs to be labeled correctly.

In our chapter on Wittgenstein we have interpreted him as making a switch in priority of shared human behavior, where we can communicate, to rules as codification thereafter. The idea was that by focusing on rules first and foremost, we were stuck with the question of how to interpret a rule, and creating an objective standard then seemed to require much more than the rule. We needed to determine all the possible applications of it. Various notions were discussed, all of them concerned prelinguistic entities. This is precisely what Chomsky's solution involves. By asking how people can know as much as they know, he is asking how it is possible that we know an infinite amount of applications of a (linguistic) rule, while it is impossible that one should learn them all (Chomsky, 1988, p. 17): "This system of knowledge provides the interpretation of linguistic expressions, including new ones that the child learning the language has never heard." We can now see that this question is the same one Wittgenstein criticizes (Wittgenstein, 2009, §147. p. 64): "So you mean that you know the application of the rule of a series quite apart from remembering actual applications of numbers? And you'll perhaps say: 'Of course! For this series is infinite, and the bit of it that I could develop is finite.'"

For Chomsky, this is where the language faculty is supposed to provide an answer: with this innate system, the steps are completely determined. That is, this innate system of knowledge would then provide us with the criteria of correctly carrying out the rule. The problem is, how do we know how to apply these prelinguistic concepts? If they are supposed to determine the meaning, we need further criteria to know how to apply a concept to a given application — the method of projection Wittgenstein discusses in §141. But then we would also need criteria for applying the method of projection correctly, and so on. This is why Wittgenstein said that every "interpretation hangs in the air together with what it interprets" and does not determine meaning (Wittgenstein, 2009, p. 86, §198). Far from being a solution, the entire idea that we have a language faculty is the source of the rule-following paradox, because it tries to explain our rule governed activities in terms of interpretation.

Chomsky's solution of positing innate mechanisms does not only give rise to the rule-following paradox, but according to Hacker it is also the cause of what Chomsky calls 'Descartes's problem'. If speaking and understanding a language would be nothing more than the language faculty processing algorithms, then our actions should be completely predictable — but they are not. Far from being a genuine problem for our understanding of language, it is a problem generated by his own theory. As we saw in chapter 3, Hacker says that Chomsky confused the original question (how we understand what we say and hear) with the question of free will, and confronting the problem candidly would have given Chomsky cause to 'stop in his tracks'. Hacker's point is — besides his criticism of Chomsky's idea that problems of free will can be solved by investigating empirical linguistics — that his problem is created by his own mechanical theory. We interpreted Chomsky's point a bit more charitable. We deemed it to mean that it is not clear how we are able to combine concepts, creating free but appropriate thoughts. On Hacker's interpretation, there is not much more to be said about the production problem: if you look at rules as causal mechanisms, then indeed you cannot account for creative language use. But, luckily, as we saw, rules are not causal at all. On our interpretation, the problem and its solution are basically the same. It is not a matter of computing what one says, but a matter of how thoughts can be combined in a free way. The conceptual system, however, is as much a physical mechanism of the brain as the language faculty is. If the goal ultimately is to describe the conceptual system in terms of combinatorial rules as well, the same argument goes: these are not rules.

Part of Hacker's criticism on Chomsky is that his questions are "a curious melange of philosophical questions and questions in psychological learning theory" (Hacker, 1990, p. 127). That is, Hacker criticizes Chomsky for rejecting the distinction between philosophy and science without argument. The problem Hacker addresses is that Chomsky's theory conflates empirical questions with conceptual questions, and the problem is that the empirical questions do not help to understand normative sciences such as grammar or logic. His theory does not help us in answering conceptual questions about "the nature of the mind", nor does it clarify our concepts of language, meaning, etc. (Hacker, 1990, pp. 146–147). And indeed, this conceptual confusion

is what created the problems in Chomsky's theory. By reducing rule-following to empirical propositions about brain processes, he is not only mistaken about the nature of rules, but is the source of the problems he wants to solve. Plato's problem is not a problem at all unless we ask how our limited experience could provide interpretation for all possible applications — how our understanding/our system of knowledge could determine all the steps in advance.

This was Chomsky's confusion: in thinking that interpretation is required to explain how we follow linguistic rules — after all the application of a rule extends far beyond any training (Plato's Problem) — he tries to determine the correct interpretation through positing mechanical rules in the mind/brain. Not realizing that his solution is not only mistaken about the nature of rule-following — that there can only be rules insofar as there is a custom, that is, that the normative aspect of these rules only exist in a shared practice — but actually is the source of his problem. If prelinguistic entities are needed to provide interpretation as to how to follow the rule, then they would have to be interpreted as well, leading to a vicious regress. While the real solution to his problem is noticing that following a rule does not require interpretation at all, we can only follow rules insofar as there is a shared practice, and this shared practice is nothing more than being trained to look at situations in this or that way, you do not choose between different ways of looking at things — you do not interpret which way the rule dictates — you are trained to see only one way of looking at things, and this is how the steps are already determined.

Now where does this leave the Chomskyan? We have argued that the theory of universal grammar fails because it neglects the normative aspects of rule-following. If we want to explain what language is, we have to look at what we do. Simultaneously we said that what we do is partly determined by our nature and our environment. This seems to leave room for a neo-Chomskyan to say that these innate mechanisms simply are what is meant by 'our nature', and determine our practice. One could, of course, set up numerous experiments, perhaps in a laboratory, to see which parts of the brain 'light up' when one is speaking, for example. In a sense this seems alright, in our reconstruction we did not try to deny that anything goes on in the brain, nor that it is completely irrelevant for what we do. As long as one keeps in mind that these underlying processes are not a language, and are not an essential part of our speaking a language.[4]

---

[4]Let us even say that our knowledge of the brain would be complete, all we would be able to produce are still no more than empirical propositions. We might even imagine that we are always right in predicting what will happen on the basis of a certain brain state, but this still would not tell us what language is. For example, what would constitute a mistake in language? Having the incorrect brain process? How would we determine what would be the correct one?

Chapter 6

Conclusion

In this thesis we have argued that Chomsky's theory of Universal Grammar is based on a confused notion of rule-following. Chomsky's confusion lies in what we have called the 'computer metaphor', which comes down to trying to evaluate our activities as human beings in terms of ourselves as natural systems. The problem was that this tendency is based on failing to distinguish following a rule mechanically, from mechanical rules. Mechanical rules are nothing more than descriptions of the predetermined movements of a machine, while following a rule mechanically simply means following a rule without reflecting upon your actions. Rule-following, however, is part of a practice, and rules themselves are parts of techniques, which is to say that they serve as a guide for conduct. They are prescriptive, not descriptive.

This confusion is the source of Plato's problem. The attempt of trying to reduce the normative status to descriptions — the idea that a rule cannot but be followed if all these descriptions are true — is based on the idea that one needs to interpret the rule in order to be able to follow it — the meaning must be more than the limited amount of applications that we encounter. After all, the possible applications are infinite. But precisely this attempt means that we would never be able to follow rules at all, after all we can then wonder whether *this* is the process that means that one is following a rule, we need further criteria.

The real answer to Plato's Problem is not some mechanical account of interpretation, but realizing that we cannot understand rule-following in terms of prelinguistic entities at all. There is no room for interpretation, rather we are trained in a certain way of dealing with out everyday situations, and it is precisely because we are so trained in a technique that we see only one way of applying it.

Bibliography

Chomsky, N. (1980). *Rules and Representations*. New York: Columbia University Press.

Chomsky, N. (1988). *Language and Problems of Knowledge: The Managua Lectures*. MIT Press.

Copeland, J. (1993). *Artificial Intelligence: A Philosophical Introduction*. Blackwell Publishing.

Cowie, F. (2010). "Innateness and Language". In Zalta, E. N., editor, *The Stanford Encyclopedia of Philosophy*. Summer 2010 edition.

Diamond, C., editor (1976). *Wittgenstein's Lectures on the Foundations of Mathematics, Cambridge 1939*. Chicago: The University of Chicago Press.

Freidin, R. (1994). "Conceptual Shifts in the Science of Grammar: 1951–1992". In Otero, C., editor, *Noam Chomsky: Critical Assessments*, pages 653–690. London: Routledge.

Freidin, R. (2012). "A Brief History of Generative Grammar". In *The Routledge Companion to the Philosophy of Language*, pages 895–916. London: Routledge.

Gerrard, S. (1996). "A Philosophy of Mathematics Between Two Camps". In Sluga, H. and Stern, D., editors, *The Cambridge Companion to Wittgenstein*, chapter 5, pages 171–197. Cambridge: Cambridge University Press.

Hacker, P. (1990). "Chomsky's Problems". *Language & Communication*, 10(2):127–148.

Shanker, S. (2006). *Wittgenstein's Remarks on the Foundations of AI*. New York: Routledge.

Stein, H. (1997). *The Fiber and the Fabric: An Inquiry into Wittgenstein's Views on Rule-Following and Linguistic Normativity*. PhD thesis, Amsterdam: Institute for Logic, Language and Computation.

Wittgenstein, L. (1978). *Remarks on the Foundations of Mathematics*. Oxford: Basil Blackwell.

Wittgenstein, L. (2009). *Philosophical Investigations*. Blackwell Publishing Ltd.