



Terms and Conditions of Use of Digitised Theses from Trinity College Library Dublin

Copyright statement

All material supplied by Trinity College Library is protected by copyright (under the Copyright and Related Rights Act, 2000 as amended) and other relevant Intellectual Property Rights. By accessing and using a Digitised Thesis from Trinity College Library you acknowledge that all Intellectual Property Rights in any Works supplied are the sole and exclusive property of the copyright and/or other IPR holder. Specific copyright holders may not be explicitly identified. Use of materials from other sources within a thesis should not be construed as a claim over them.

A non-exclusive, non-transferable licence is hereby granted to those using or reproducing, in whole or in part, the material for valid purposes, providing the copyright owners are acknowledged using the normal conventions. Where specific permission to use material is required, this is identified and such permission must be sought from the copyright holder or agency cited.

Liability statement

By using a Digitised Thesis, I accept that Trinity College Dublin bears no legal responsibility for the accuracy, legality or comprehensiveness of materials contained within the thesis, and that Trinity College Dublin accepts no liability for indirect, consequential, or incidental, damages or losses arising from use of the thesis for whatever reason. Information located in a thesis may be subject to specific use constraints, details of which may not be explicitly described. It is the responsibility of potential and actual users to be aware of such constraints and to abide by them. By making use of material from a digitised thesis, you accept these copyright and disclaimer provisions. Where it is brought to the attention of Trinity College Library that there may be a breach of copyright or other restraint, it is the policy to withdraw or take down access to a thesis while the issue is being resolved.

Access Agreement

By using a Digitised Thesis from Trinity College Library you are bound by the following Terms & Conditions. Please read them carefully.

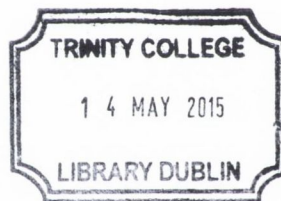
I have read and I understand the following statement: All material supplied via a Digitised Thesis from Trinity College Library is protected by copyright and other intellectual property rights, and duplication or sale of all or part of any of a thesis is not permitted, except that material may be duplicated by you for your research use or for educational purposes in electronic or print form providing the copyright owners are acknowledged using the normal conventions. You must obtain permission for any other use. Electronic or print copies may not be offered, whether for sale or otherwise to anyone. This copy has been supplied on the understanding that it is copyright material and that no quotation from the thesis may be published without proper acknowledgement.

**CHOMSKY QUINE AND
NATURALISTIC
PHILOSOPHY**

PHD IN PHILOSOPHY

2015

DAVID KING



Thesis 10503

DECLARATION

(1) This thesis has not been submitted as an exercise for a degree at this or any other university.

(2) This thesis is entirely the work of David King.

(3) I agree that the thesis may be lent.

This permission covers only single copies made for study purposes, subject to normal conditions of acknowledgement.

David King

David King

PhD Acknowledgments

This thesis would not have been possible without the help of many people. I would like to thank my supervisor Dr James Levine for all his help with the writing this thesis. Without his guidance it would not have been possible to complete this thesis. I would also like to thank the department of philosophy Trinity College Dublin for supporting the writing of the thesis. The philosophy department in Trinity provides the perfect environment for academic research. Professor Noam Chomsky graciously answered any questions I asked him about his work and its relation to Quine and I would like to note my sincere appreciation for this help. Finally I would like to record thanks to my wife Mito King who was a constant support to me over the last four years.

Yours greatfully,

David King

CHOMSKY, QUINE AND NATURALISTIC PHILOSOPHY

CHAPTER 1: Naturalism and the Analytic/Synthetic Distinction

Part 1: Rationalism, Empiricism, Naturalism

Section 1: Types of Naturalism 8

Section 2: Empiricism and Rationalism 13

Section 3: Methodological and Ontological Dualism 21

Part 2: The Analytic and the Synthetic

Section 1: The Object of Attack 29

Section 2: Two Dogmas of Empiricism: A Summary 30

Section 3: Chomsky's Defence of the Distinction 34

Part 3: Analyticity Naturalised?

Section 1: Types of Analyticity 50

Section 2: Faculty Psychology and Analyticity 57

Section 3: Chomsky and Quine on Analyticity 60

CHAPTER 2: Substantive Disagreements

Part 1: Underdetermination and Language Acquisition 66

Part 2: Underdetermination and Science 77

Part 3: The Philosophy of Linguistics

Section 1: Quine's Criticisms of Chomsky 97

Section 2: Chomsky's Criticisms of Quine 110

CHAPTER 3: Quine contra Chomsky: The Question of Indeterminacy

Part 1: The Indeterminacy of Translation: A Synopsis 120

Part 2: The Indeterminacy of Translation and Underdetermination 130

Part 3: The Indeterminacy of Translation and Empirical Research

Section 1: Evidence for Innate Concepts of Body and Mind 132

Section 2: Ontology and Language Learning: An Empirical Test	137
Section 3: Theory of Mind and Triangulation	147
Section 4: Markman and Word Learning	149
Part 4: Evaluating the Relevance of the Experiments	153
Part 5: The Aptness of the Analogy	159
Chapter 4: Poverty of Stimulus Arguments: The Evidence	
Part 1: Challenges to the APS	
Section 1: Implications for Chomsky and Quine	163
Section 2: Some Recent Criticisms of the APS	172
Section 3: Negative Evidence	183
Section 4: Evidence of Negative Evidence	188
Part 2: Conceptions of the APS	
Section 1: The Real APS?	196
Section 2: The Structure of Collins' APS	203
Section 3: Which APS Does Chomsky Work With?	207
Part 3: Chomsky's Latest Defence of the APS	219
Chapter 5: Where Do We Go From Here?	
Part 1: Nativism and Anti-Nativism	233
Part 2: Some Tests for the Future: Quine	238
Part 3: Some Tests for the Future: Chomsky	243
Part 4: Naturalised Epistemology	245

List of Abbreviations used in Thesis

PLD= Primary Linguistic Data

IDT = Indeterminacy of Translation

UD = Underdetermination

NHSLM = *New Horizons in the Study of Language and Mind*

WO = *Word and Object*

WObjs = *Word and Objections*

APS= Argument from Poverty of Stimulus

When speaking about a concept I will use italics and caps, for example the concept of mind= *MIND*.

When mentioning a word it will be enclosed in quotation marks, so for example the word which denotes Car when mentioned = 'Car'.

'Gavagai' when beginning with a capital letter = a sentence.

'gavagai' when beginning with a small letter = a term.

INTRODUCTION

...I remarked that Quine's commitment to behaviourism may well prove to be his Achilles' heel. Nevertheless, I think we may firmly conclude that while Quine's naturalistic-behaviouristic theory of language, especially that part pertaining to language acquisition, is in a sense incomplete, it has not been refuted. (Gibson 1982, 205)

It has been thirty years since Roger Gibson wrote his book on the philosophy of Quine. In his book he discussed Chomsky and Quine's disagreement on the nature of language and language acquisition. He concluded his book by noting that it was still an open question as to whether Quine or Chomsky had the correct theory of language acquisition. He further commented that despite what many people believed, Chomsky had not refuted Quine's conception of language. In the thirty years since Gibson wrote his book, mountains of data have been made available which can be used to try and decide which conception of the nature of language (if any of them) is the correct one. When evaluating this data it is argued that even now we still have not obtained enough evidence to conclusively decide between the two competing pictures. However the new data does indicate that, if anything, there is evidence implying that Chomsky's theory is less convincing than it was thought to be thirty years ago. Quine's theory, on the other hand, has not been tested sufficiently; this is because Chomsky's rhetoric has given the mistaken impression that behaviourism is a dead research programme. So, while recent research has certainly clarified the nature of the debate between Quine and Chomsky on language acquisition, the question of which theorist's view is correct still remains open.

Contemporary philosophical scholarship typically behaves as though behaviourism were a dead research programme. It is assumed that Chomsky has

shown that the behaviouristic approach adopted by Quine is a failure. Chomsky's poverty of stimulus arguments¹ are supposed to have shown that, contrary to what Quine thinks, induction, and positive and negative reinforcement play little role in language acquisition. The primary aim of this thesis is to evaluate this claim. However, prior to doing this, it is first necessary to clear away the ground of mutual misunderstanding between Quine and Chomsky. To this end, the first two chapters will be devoted to explicating where exactly both thinkers agree and disagree with each other. It will be shown that they both adopt the same type of methodological naturalism as each other. It will also be demonstrated that they disagree with each other in the relative importance of competence and performance in our linguistic theory. A further disagreement between them is in how they deal with Underdetermination²; Chomsky typically responds to UD by postulating a faculty to overcome UD, whereas Quine typically argues that UD is a fact of life which we just have to live with.

Chomsky views the Indeterminacy of Translation³ as a form of UD which can be overcome by postulating innate concepts. Furthermore his APS is an example of UD applied to syntax acquisition. In Chapters 3 and 4 the evidence which Chomsky uses to support his postulations of innate concepts and innate syntax to overcome UD is evaluated. I will argue that in the case of the UD of our concepts, he has not demonstrated in a non-question begging manner why we should postulate innate concepts, while in the case of UD in syntax acquisition he has not studied people's primary linguistic data, or actual behaviour in enough detail to justify his claims that UD affects syntax acquisition. This evaluation will show that Quine's alternative picture of language acquisition is still a viable option. It will be further shown that

¹ Henceforth Poverty of Stimulus = APS

² Henceforth Underdetermination = UD

³ Henceforth The Indeterminacy of Translation = IDT

there is some experimental evidence which supports Quine's view that reinforcement and induction play a larger role in language learning than Chomsky admits.

I will argue that while Chomsky's claims about language acquisition have been tested over the last fifty years and have not all stood up to critical scrutiny, Quine's views have not been tested in enough detail because most people have the erroneous belief that his views have already been refuted. Overall there is not enough evidence in to decide whether Chomsky or Quine have the correct theory of language acquisition. To this end, I will propose experiments which I argue will help solve the debate between both thinkers on this issue. Ultimately I argue that getting clear on whose theory is the correct one is vitally important because doing so will help us understand whose naturalistic theory of the nature of mind is correct. Obviously, discovering this will be a huge step in the right direction of constructing a naturalistic epistemology. The thesis ends with the claim that if naturalised epistemology is to be something more than a buzz-word philosophers need to resolve whether Chomsky's or Quine's theory of naturalistic epistemology is correct. Naturalised epistemology, if it is to be a valid research programme, needs to be constantly subject to tests in light of new experiments; furthermore, it needs to be proposing new tests. A truly scientific epistemology will not stand still with Quine's conception but will evolve as new scientific challenges are noted. The debate between Quine and Chomsky has greatly improved our understanding of how people go from stimulus to science. The experimental tests which I recommend in Chapter 5 will go even further in clarifying the issues at hand. Once we have resolved the debate between Quine and Chomsky, we will be left with a much more accurate picture of naturalised epistemology than we began with.

CHAPTER ONE

NATURALISM AND THE ANALYTIC/SYNTHETIC DISTINCTION

The central aim of this chapter is to compare the views of Chomsky and Quine on the nature of naturalism. In Part 1 I consider their views on Methodological Naturalism and Methodological Dualism as well as their views on Ontological Naturalism and Ontological Dualism, and I argue that neither thinker endorses any kind of dualism and both of them are methodological naturalists.

Non-naturalistic philosophers typically pointed to our grasp of a priori truths and necessary truths to argue that human epistemic capacities could not be explained naturalistically. Mere induction and trial and error reasoning they argued would not be sufficient to explain how we grasped a-priori truths and necessary truths. To explain our grasp of such truths they typically appealed to a mysterious faculty of intuition which could not be explained naturalistically. Logical positivists who wanted to naturalise our knowledge did so by arguing that all our supposed grasp of a priori truths and necessary truths could be explained in terms of analytic truths. Quine, of course, famously argued against the existence of analytic connections in natural language. He argued that a truly empirical philosophy will reveal upon close inspection that we have no justification for presupposing analytic connections in our theory of the world. Chomsky has argued that Quine's views on analyticity are incorrect. He argues that the empirical evidence clearly shows that analytic connections exist in natural language.

The primary aim of discussing analyticity is to clarify how both thinkers view analyticity, and to evaluate how their views on analyticity are connected to their naturalism. To do this, I will explicate what Chomsky's positive picture of analyticity is, showing that his conception in no way helps explicate any supposed a priori

knowledge nor our purported grasp of necessary truths. It will be demonstrated that Chomsky's actual conception of analytic truth is that of a contingent feature of natural language which is encoded into the brain. Chomsky's conceptual connections in natural language are entirely naturalistic posits which he uses to explain people's linguistic intuitions. So the debate between Chomsky and Quine on this point is an entirely empirical and naturalistic one.

I argue that Chomsky has not provided any substantial evidence to support his positive conception of analyticity. I conclude that Chomsky has not dealt with Quine's criticisms of analyticity and his positive conception of analyticity is better explained in the Quinean web of belief manner. The main aim of this chapter is to situate both thinkers in relation to their philosophical ancestry, and emphasise their shared methodological naturalism. Having shown that both thinkers are methodological naturalists I emphasise their different views on analyticity and show how their dispute on this topic is entirely naturalistic.

PART 1: RATIONALISM, EMPIRICISM AND NATURALISM

SECTION 1: TYPES OF NATURALISM

Chomsky and Quine both adopt a thoroughly naturalistic approach to the study of the mind and the world. However, obviously there are different types of naturalism: ontological naturalism, epistemological naturalism and metaphysical naturalism. So it is important to be clear from the outset which type of naturalism they endorse.

Chomsky distinguishes between two kinds of naturalism which are prominent in contemporary philosophy: metaphysical naturalism and epistemological naturalism. He endorses a version of naturalism which he considers different from either of the above types of naturalism; he calls it methodological naturalism. According to

Chomsky, epistemic naturalism is simply a branch of psychology concerned with how we come to hold the beliefs we do. Metaphysical naturalism is the view that any philosophical accounts of the mind, language and knowledge must be consistent with the hypothesis advanced by the natural sciences. So for example such accounts should not include hypothetical abstract entities such as Platonic entities etc. (2000b, 81).

From the point of view of the methodological naturalist the enterprise of epistemological naturalism is an unproblematic branch of psychology. However, Chomsky argues that the tenets of metaphysical naturalism are much more controversial. The difficulty with such an approach is that its claim that our philosophical accounts of the mind, language, and knowledge should be consistent with the hypothesis advanced by science is ambiguous. The ambiguity resides in saying what 'continuous with science' means. Chomsky asks rhetorically if this means continuous with current physics? (ibid, 82). If so, then metaphysical naturalism would be absurd because we have no idea whether current physics will continue on in its present form. Future physics may deviate radically from current physics. So, for example while Platonic Senses may not be consistent with current physics there is nothing to say that they will not be continuous with the physics of the future. And if the metaphysical naturalist is insisting that our current studies of the mind, language and nature must be consistent with some kind of idealised physics which will be constructed in the future, such a claim can be dismissed out of hand. Clearly it would be an absurd criterion to ask for current theories of language, mind and knowledge to be consistent with some future undiscovered physics. Based on such considerations, Chomsky concludes that metaphysical naturalism entails that our philosophical theories of the mind, language and knowledge must be consistent with our current physics. However, one has reason to doubt that it is sensible to ask that the entities

postulated to explain the mind and language be consistent with the entities postulated by current physics. To drill this point home Chomsky discusses a chapter in the history of science:

Still more instructive is what was taking place just before the unification of chemistry and physics. Prior to unification, it was commonly argued by leading scientists that chemistry is just a calculating device, a way to organise the results about chemical reactions, sometimes to predict them. In the early years of the last century, molecules were regarded in the same way. Poincaré ridiculed the belief that the molecular theory of gases is more than a mode of calculation; people fall into error because they are familiar with the game of billiards, he said. Chemistry is not about anything real, it was argued: the reason is that no one knew how to reduce it to physics. In 1929, Bertrand Russell—who knew the sciences well—pointed out that chemical laws ‘cannot at present be reduced to physical laws’; not false but misleading in an important way. It turned out that the phrase ‘at present’ was out of place. Reduction was impossible, as was soon discovered, until the conception of physical nature and law was (radically) revised. (1980a, 72)

The main point here is methodological; when we are constructing a scientific theory in a particular domain what we seek is the simplest possible theory which predicts and explains the datum we are interested in. If this theory needs to posit entities such as propositions or other abstract entities which do not seem to be consistent with current physics then so be it. The current physics may be in error and may undergo radical revisions which make the posits of psychology etc. consistent with the new physics. Chomsky’s discussion of chemistry and physics illustrates an instance in history where it was the physics which was revised to make reduction possible, not the less fundamental science of chemistry. This historical episode should warn metaphysical naturalists from preaching that certain entities should be banned from our ontology because they are not consistent with current physics.

A typical move of metaphysical naturalists is to ban abstract entities from our ontology, and to insist that the only things which exist in the world are the physical objects discovered by physics. Chomsky has argued against this approach because he believes that no sense can be given to the notion of physical. Our concept of what

makes something physical is constantly evolving as we learn more and more about the world in areas like quantum physics etc. So to say that everything is physical, is either to say that everything conforms to our current theory of the world, or it is to say that all future science will conform to our current conception of the physical. The latter option is absurd and completely lacking in evidence, while the former claim could be stipulated to be true but it would be nothing more than a stipulation. Much contemporary science invokes entities such as propositions, numbers and classes which are patently not physical. And we have seen that such entities cannot reasonably be ejected from our ontology merely because they are inconsistent with our current physics. So from a Chomskian perspective the notion of physicalism is an arbitrary doctrine espoused by the metaphysical naturalist.

Quine's views on this topic are largely consistent with Chomsky's. Quine espouses methodological naturalism and he is not a supporter of ontological naturalism. For Quine, ontological questions are to be decided based on the utility of the particular entity which is being posited. Quine denies the existence of some abstract entities. For example, he argues against admitting propositions into our ontological system. However his argument against propositions does not centre on the fact that propositions are spooky abstract entities which are not consistent with physicalism. Rather his arguments centre on the fact that he can explain things such as verbal behaviour and scientific knowledge without positing such entities. He also argues that such entities have no clear identity conditions. So, based on these arguments he thinks that we should not admit such entities into our ontology. Quine is quite explicit that he does not reject such entities because they are abstract objects but rather because they have no clear identity conditions, and because the facts they are introduced to explain can be explained in a simpler manner.

He argues that we should admit some abstract objects into our ontology. He argues, for instance, that we must admit classes into our ontology. His reasons for admitting them are because they provide us with massive explanatory power which is otherwise not available to us. Quine's approach is that of a methodological naturalist. Like Chomsky he has no arbitrary commitment to physicalism and he is prepared to admit any entity into his ontology as long as it has explanatory value and clear identity conditions.

Quine excludes mental entities from his ontology. His reasons for doing this are again methodological rather than metaphysical. He argues that we can do without mental events in explaining human behaviour. So since such entities are explanatorily barren he sees no reasons to admit them into our ontology.

So we can see that both Quine and Chomsky are methodological naturalists, and neither is an ontological naturalist. Both theorists also argue for a naturalised conception of epistemology. However Chomsky is very critical of Quine's version of epistemology naturalised:

Contemporary epistemic naturalism derives from Willard Quine's "epistemology naturalized," which stipulated that the study of knowledge and belief must be incorporated within a narrow branch of behaviourist psychology of no known scientific significance, a strange move in itself, which has evoked surprisingly little challenge (2000b, 80)

Chomsky is here criticising Quine's Naturalised Epistemology because of its behaviourist orientation. He is not offering a non-naturalistic alternative to Naturalised Epistemology; rather he is saying that Quine's version of Naturalised Epistemology is false. So in this sense the debate between Chomsky and Quine on this point is an entirely empirical one. To decide between them we merely have to evaluate which version best fits the empirical facts. However while the debate between both thinkers is indeed partially empirical, some aspects of the debate result

from mutual misreading of each other's projects. So in the following section I will diagnose where both thinkers are merely misreading each other's project, and where there is actual disagreement between them. A major contributor to the mutual misreading between both thinkers stems from Chomsky's interpretation of the history of philosophy and his partial equating of Quine's views with those of traditional empiricists.

SECTION 2: EMPIRICISM AND RATIONALISM

Chomsky, as well as being the father of modern linguistics and a respected contributor to contemporary political theory, has written extensively and deeply about philosophy and its contribution (good and bad) to the field of linguistics. An important point to note is that Chomsky sees himself as an heir to the philosophical tradition of the seventeenth century. Philosophers such as Hume, Descartes and Leibniz are the types of figures who Chomsky has the closest affinity to. A central thread running through Chomsky's writings is a belief in philosophical naturalism, along with a deep respect for the rationalist philosophers of the seventeenth century. Furthermore, Chomsky believes that contemporary philosophy, if it is to learn anything from linguistics, should learn from the success of adopting a rationalist approach to the study of science as opposed to the emptiness which he believes ensues when one studies science from the point of view of an empiricist. (2000a, 145).

A perceptive reader will have noticed a tension that emerged in the above paragraph. Chomsky is, on the one hand, criticising empiricism as a philosophical position. On the other hand, he is proclaiming an affinity to Hume, one of the greatest empiricists of all time. The seeming contradiction can be resolved by considering what Chomsky has to say about his universal grammar and its relation to Hume

An enquiry into universal grammar... falls within what Hume called 'moral philosophy' that is 'the secret springs and principles, by which the human mind is actuated in its operations', and most importantly 'those parts of (our) knowledge' that are derived from 'the original hand of nature' (1980a, 30).

In other words, he views himself and Hume as engaging in the same project. This project is a scientific and empirical research program. Their affinity lies in the fact that both of them are psychologists trying to explain different aspects of the mind. Hume's views on the mind are empirical claims. They are subject to empirical tests and refutations, just like Chomsky's views on universal grammar are. However, according to Chomsky, Hume's model of the mind was incorrect. Hume's mistake lay in certain empirical assumptions which he made about the structure of the mind. These are the assumptions which marked Hume out as an empiricist as opposed to a rationalist.

Chomsky's characterisation of the difference between Rationalism and Empiricism is as follows:

Each postulates innate dispositions, inclinations, and natural potentialities... The crucial question is not whether there are innate potentialities or innate structures. No rational person denies this, nor has the question been at issue. The crucial question is, whether this structure is of the character of Empiricism or Rationalism; whether it is the character of 'powers' or 'dispositions'; whether it is a 'passive' system of incremental data processing, habit formation or induction, or an 'active' system which is the 'source of linguistic competence' as well as other systems of knowledge and belief (1975b, 216).

Obviously, Hume's description of how the mind works follows the empiricist model. Chomsky believes that empiricist models such as the one above cannot explain the complexity and creativity of the mind and hence are false. However he does admire the scientific manner in which Hume framed the question.

Chomsky has claimed that Quine's theory of the mind falls within the category of empiricism. So he attributes to Quine the view that when we are studying the mind

we must begin with the assumption that the mind begins as a passive system of incremental data processing, which learns using induction together with habit formation (1975b, 199). He opposes Quine's empiricism to the rationalist approach which Chomsky claims to follow. However, clearly Chomsky is attributing to Quine views which he does not hold; this can be seen when we consider Quine's views on the nature of rationalism and empiricism.

Quine would describe himself as both a naturalist and an empiricist. He describes naturalism as the claim that it is within science itself, and not in some prior philosophy, that reality is to be identified and described (1981, 21). His naturalism is similar to the methodological naturalism Chomsky espouses in that Quine does not limit the type of entities that can be posited to explain phenomena as long as they have explanatory value. Quine makes the following claim about science:

Descartes' dualism between mind and body is called metaphysics, but it could well be called science, however false. He even had a causal theory of the interaction of the mind and body through the pineal gland. If I saw indirect explanatory benefit in positing sensibilia, possibilia, spirits, a Creator, I would joyfully accord them scientific status too, on a par with such avowedly scientific posits as quarks and black holes. (1995, 462)

Furthermore, Quine even claims that as far as scientific explanation of the phenomenon is concerned, Descartes' scientific explanation is superior to empiricists such as Locke's:

Descartes thought we had innate knowledge and innate ideas. Locke thought not. I despair of sharpening the issue by defining the term 'idea'. Definition even of 'knowledge' is in trouble since Gettier's challenge of the definition of knowledge as true and warranted belief. However, we need no such sharpening of the issue to see that the evidence favours Descartes over Locke (2008a, 176).

So here we see Quine defending rationalism against traditional empiricism. This shows that his views are distinct from those of his predecessors. Furthermore, Quine and Chomsky are in agreement on a number of points. Firstly, both thinkers

agree that the theories of the seventeenth century philosophers are best viewed as scientific theories. Secondly, Quine, like Chomsky, believes that the evidence favours Descartes' rationalist approach over the empiricist approach of Hume. However, they do disagree in how they interpret the nature of empiricism. The nature of this disagreement can be clarified by further explicating Quine's views on empiricism.

We have seen Chomsky's views on the nature of traditional empiricism. In his (1969a), Chomsky places Quine firmly within this tradition by claiming that *Word and Object* is a worked-out Humean conception of the philosophy of language. When Quine replied to this by claiming that he allowed any innate mechanism whatsoever as long as they could be cashed out experimentally, Chomsky construed this as a rejection of his empiricism:

Quine's later views do not fall within the class of systems E, though the earlier ones do in so far as it is clear. Thus, if we are prepared to welcome any innate mechanism, however elaborate, we are not bound by procedures of the character of E, but can explore richer and, I believe, more adequate theories. (1975b, 200)

Chomsky distinguishes between Quine 1960, who worked with an empiricist model of the mind and Quine 1969, who worked with a model which is more consistent with rationalism. This interpretation is incorrect. Quine worked with the same empiricist model throughout his career; however, it was not the narrow empiricism which Chomsky attributes to Hume.

Quine differentiates himself from his empiricist predecessors by claiming to be an externalised empiricist (1966, 58). By calling himself an externalised empiricist Quine is distancing himself from a philosophical tradition which begins with Descartes and runs right through to Kant. Rene Descartes had inaugurated a tradition in philosophy which began with the subjective certainty of the Cogito and tried to build the world up from that certain starting point. Empiricist philosophers followed

in this Cartesian tradition, though they used their subjective experiences as a foundation to try to construct the world rather than the certainty of the Cogito. Kant followed in this tradition as well and took it in an even more devastating direction by claiming that we could never know the world as it is in itself but only in so far as it conforms to our mode of cognition. What these thinkers have in common is a belief that the philosopher can provide a foundation which will ensure that our scientific claims are justified and stand on a firm ground; also that this foundation can be provided through the use of pure reason, or by analysing how much of our beliefs can be justified on the basis of our subjective empirical experiences. Quine is totally against this picture. He claims that the idea that philosophy can provide a foundation for science which is more secure than science is an illusion:

Naturalism is the recognition that it is within science itself, and not in some prior philosophy, that reality is to be identified and described; again that it is the abandonment of the goal of a first philosophy prior to natural science. (1995, 462)

So, on Quine's picture, the philosopher has two jobs. The first is to explain how it is that scientists manage to construct our picture of the world; this is naturalised epistemology and is, he claims, a chapter of empirical psychology. The second is to make explicit the ontological commitments of our best science by translating it into the language of quantificational logic; this is metaphysics.

An important point to note is that Quine has no problem with the traditional empiricist or rationalists when they operated as scientists offering differing models of the mind. Rather he had a difficulty with the foundationalist epistemology espoused by both groups. Quine's externalised empiricism stands against the traditional world picture on two fronts. Firstly, Quine does not begin internally and deduce the world from that vantage point, rather; he begins in the external world and tries to discover

how the organism manages to construct a total theory of the world from the impacts that impinge on its body. Quine then extrapolates whatever innate apparatus is necessary to account for how human beings are able to construct a theory from the impacts that impinge on their senses. Secondly, while Quine does try to explain how our best scientific theories are constructed, he does so from within science and not from some a priori vantage point a la Descartes.

The above two points are connected and can be understood in relation to traditional philosophical problems such as scepticism about the external world, and the problem of induction. Firstly the problem of induction as traditionally formulated by Hume seemed to put the validity of scientific laws such as Newton's in doubt. Traditionally philosophers felt that they had to solve this problem in order to make secure the truths of science. Quine however argued that we have no more certain knowledge than the knowledge of science; he summed this up with his famous quote '*The Humean condition is the Human condition* (1969, 56)'. Quine further argued from natural selection that the cause of humans making such accurate inductions is selection pressure acting over our evolutionary history. This evolutionary argument obviously does not justify induction since it relies on knowledge gained through induction; rather it shows why our inductions tend to be correct according to our best scientific theories.

Creatures who are invariably wrong in their inductions have the pathetic but praiseworthy characteristic of dying of before reproducing their own kind (1969, 56)

In other words, natural selection would cull those creatures that are unable to make accurate predictions about their environment. Here Quine is explaining induction internal to our overall scientific picture rather than trying to do what he believed was impossible: to justify induction by an a priori argument.

Quine's answer to the sceptical problem about our ability to justify our belief in the existence of the external world has a similar structure to his above discussion of induction. In *The Roots of Reference* Quine makes the following point:

The crucial logical point is that the epistemologist is confronting a challenge to natural science that arises from within natural science (1974, 2)

He claimed that philosophers such as George Berkeley were concerned with deriving depth from two-dimensional data because the surface of the eye was two dimensional. So since these sceptical challenges to science typically sprung from within science, scientists are free to use science to refute them. On Quine's view, all supposed philosophical questions, if legitimate, are scientific questions and all scientific answers are by their very nature fallible.

However, for Quine, the simple picture of falsification which Popper outlined in his *The Logic of Scientific Discovery* does not capture how science actually proceeds. According to Quine, our scientific theories are a massive interconnected web of sentences. This web of sentences is connected to the world via observation sentences. Quine's famous holistic view is exemplified by his slogan 'The Maxim of Minimum Mutilation'. This maxim states that when faced with a perceived falsification of a theory we try to accommodate it by rescinding the sentence within the theory which implied the false prediction. Which sentence we rescind is decided by pragmatic considerations, sentences which infiltrate our total web of beliefs are less likely to be rescinded than those which will have less impact on our overall theories. Hence we are less likely to rescind logical laws than we are to rescind other aspects of our theories.

The central point of Quine's is that no aspect of our theory of the world is immune from revision; any belief can be falsified. While Quine qua naturalist can offer reasons which explain that we are in touch with the external world. The reasons he

offers are fallible, hence do not speak to the type of certainty that the sceptic is looking for. Furthermore Quine begins from a different vantage point than the sceptic. Quine accepts the scientific picture of man as a creature in the world. He sketches a picture of how this creature manages to construct his scientific picture of the world from the impacts of the physical world that impinge on him. While Quine can acknowledge the possibility that his beliefs about the world can be wrong, he will only be worried when his actual theory of the world starts to make persistent false predictions which cannot be eradicated by modification of the web of belief.

Obviously, then, Quine's externalised empiricism is not of the character of the empiricism that Chomsky attributed to Hume, it is not a passive system of incremental data processing. In "Linguistics and Philosophy" Quine argued as follows:

'Externalised empiricism or behaviourism sees nothing uncongenial in the appeal to innate dispositions to overt behaviour, innate readiness for language learning. The as yet unknown innate structures, additional to mere quality space, that are needed in language learning, are needed specifically to get the child over this great hump that lies beyond ostension, and induction. If Chomsky's anti-empiricism or antibehaviourism says merely that conditioning is insufficient to explain language-learning, then the doctrine is of a piece with my doctrine of indeterminacy of translation'. (1966, 58)

Quine's empiricism aims to begin with the external facts of behaviour and to *discover* the mechanisms that make this behaviour possible. Unlike Chomsky's Hume, he does not begin by assuming a cognitive model of incremental data processing.

Quine's externalised empiricism, unlike the models of Hume etc., does not presuppose a particular model of the mind at the outset. Quine makes no assumption that the mind must have the structure of incremental data processing etc. He simply says we should begin with external behaviour and postulate whatever innate mechanisms are necessary to explain the behaviour.

Thus far we have seen that Quine agrees with Chomsky that their rationalist and empiricist predecessors were offering scientific models of the mind. Furthermore,

Quine argues that the evidence favours the rationalists on the innate ideas issue. However, despite this fact Quine still calls himself an empiricist. His empiricism differs from that of his empiricist predecessors, because his externalised empiricism rejects their foundationalist epistemology. For this reason, Quine is not guilty of beginning his theorising about the mind by presupposing an empiricist model of incremental data processing, induction etc. Rather, like Chomsky he is prepared to posit any innate mechanisms which are necessary to explain people's linguistic abilities.

SECTION 3: METHODOLOGICAL AND ONTOLOGICAL DUALISM

Aside from arguing (incorrectly) that Quine and Hume adopt the same kind of empiricism, Chomsky also argues that there is a big difference between Hume and Quine. He claims that the difference is that Hume is a thoroughgoing naturalist who studies the mind and the world using the same scientific methodology, whereas Quine, on the contrary, adopts an arbitrary methodological dualism. In order to clarify this point I will need to discuss Chomsky's views on Naturalism.

As I have already said, Chomsky considers himself a methodological naturalist. He opposes this naturalism to what he calls methodological dualism:

Methodological Dualism: The doctrine that in the quest for theoretical understanding, language and mind are to be studied in some manner other than the ways we investigate natural objects, as a matter of principle (2000b, 135).

He calls his version of naturalism 'methodological naturalism'; and claims that it amounts to applying the Galilean method of inquiry to all areas of discourse. The idea is that; 'when one is studying any aspect of the world one should construct intelligible explanatory theories that one then tries to reduce to the core sciences' (1980a, 5). This 'methodological naturalism' is opposed to 'methodological dualism' which Chomsky accuses the vast majority of contemporary philosophers of engaging in.

Chomsky considers Quine a ‘methodological dualist’, despite the fact that Quine describes his project as naturalistic. According to Chomsky, Quine is a methodological dualist because he treats the mind as an area which should be studied by different methods than all other aspects of the natural world. This methodological dualism occurs for Chomsky as a result of arbitrary stipulations on what is deemed good evidence for psychology and linguists, while such methods are not adhered to in the study of any other aspect of the natural world. Chomsky accuses Donald Davidson, Richard Rorty, and Hillary Putnam, amongst others, of engaging in this methodological dualism.

The reader could be forgiven for at this point being a bit confused. Such a reader may be wondering why Chomsky, who aligns himself to Descartes (The King of Dualists), is accusing Quine, a stern critic of Cartesian Dualism, of being a methodological dualist. Chomsky’s reasons for arguing so are that even though Descartes was indeed a dualist, his dualism was a scientific position as opposed to the arbitrary methodological dualism to which people like Quine supposedly adhere. A brief digression discussing Chomsky’s take on Descartes’ mind/body problem will help clarify exactly what Chomsky’s position is.

According to Chomsky, the mind/body problem arose as a result of Descartes’ failure to explain the creativity of language in the mechanical terms of his time. What Descartes meant by the creativity of language can be explained in three important points: (1) Language is potentially infinite (there seems to be no limit to the number of sentences we can construct); (2) language is free from stimulus control⁴; (3) language use is coherent and has appropriateness to the situation (1966, 59). For

⁴ The term “free from stimulus control” is ambiguous. A natural reading is that here Chomsky is referring to external stimulus control, that for example nothing in the environment determines how the child will speak. In fact Chomsky means something more radical. He means free from both external and internal stimulus control. I will discuss this point in detail in Chapter 4.

Descartes it was the second and third aspects of language use that he could not reconcile with his mechanical philosophy. There seemed to be no room within the deterministic mechanical philosophy for a creature who could utter words which were free from stimulus control, but which were coherent and appropriate. For example while a machine could be created which could constantly mouth novel sentences in a manner free from external stimulus such a machine would make no sense to others. Likewise it is possible an organism could be created which can be trained to mouth words in certain determinate situations through conditioning but such a creature would be for Descartes a mere parrot. It is for this reason, according to Chomsky, that Descartes postulated a substance separate from matter which he called a mind (a thinking thing). However, in Chomsky's view, after Newton, this distinction became pointless as Newton postulated such occult forces as gravity which he could not explain in materialistic terms but which helped him to make incredibly accurate predictions. Newton's gravity was no less a mystery than the creativity of language. The substance of the universe was obviously more complex than Descartes had realised. His reasons for distinguishing the mind from the body were no longer valid. Descartes' machine was destroyed so there was no need for the ghost.

The important aspect of this discussion is that Descartes' dualism resulted from a scientific problem and was dissolved as a result of further advances in the science of his time. According to Chomsky, the approach in analytic philosophy is completely different. He cites Quine's IDT argument as a typical case of contemporary philosophers holding the study of the mind to higher methodological standards than the study of any other aspect of the empirical world. I will not at this time discuss the IDT argument because I will be discussing it in detail in Chapter 3. Chomsky's essential point is that the IDT is no different than ordinary empirical UD.

He argues that the fact that Quine treats UD as harmless in physics but as fatal in linguistics shows that Quine is a methodological dualist (Chomsky: 1980a, 14-16).

So for Chomsky, when Descartes postulated a theoretical entity (thinking substance) to explain the behaviour of a body (in this case a man), he was just using the standard methodology of science. He was behaving in an analogous manner to the physicist who postulates gravity to explain why a body drops. According to Chomsky, rather than engaging in scientific study, Quine was arbitrarily stipulating that we should use different techniques in linguistics and psychology than in the study of all other sciences.

We have seen in our above discussion that Quine shares Chomsky's views on the nature of the traditional rationalist/empiricist debate. Furthermore, Quine explicitly endorses methodological naturalism, which is the same as the methodological naturalism which Chomsky endorses. Chomsky accuses Quine of being a methodological dualist because of the way he treats UD in physics, and UD in linguistics. In Chapter 3, I will argue that Chomsky's accusation against Quine of inconsistency in his treatment of UD in linguistics and physics is incorrect. What our discussion so far has shown is that both Chomsky and Quine are methodological naturalists, who view traditional philosophical questions as questions which should be answered in a scientific manner.

In his book *The Philosophy of Quine: An Expository Essay* Roger Gibson characterised the disagreement between Quine and Chomsky as follows:

Quine rejects what he takes to be Chomsky's irreducibly mentalistic posits, and Chomsky rejects what he takes to be Quine's unreasonable, behaviouristic restrictions (1982, 180).

What Quine means by an irreducibly mentalistic posit would be for example explaining the meaning of a word in terms of correspondence with an idea in the mind. Quine's problem with explanation in terms of ideas is as follows

The ideas in between are as may be and may vary as they please, so long as the external stimulus in question stays paired up with the word in question for all concerned (2008, 315)

In other words, our internal idiosyncratic ideas do not determine the meaning of the word. Two people understand each other as long as the sounds they mouth share the same stimulus meaning.

Quine defines language as a system of dispositions to verbal behaviour. Dispositions are for him physical states that we do not yet know how to explain physically so we must explain them behaviouristically. Imagine, for example, a dog that barks every time a human enters his garden. A dispositional account would say that dog x is disposed to engage in behaviour y in circumstance z, and would try to represent this probabilistically. However, a person engaging in this type of account would not be explaining the barking at a physical level, though that would ultimately be the aim. As Quine put it

The deepest explanation, the physiological, would analyse these dispositions in explicit terms of nerve impulses and other anatomically and chemically identified organic processes. (ibid, 312).

So for Quine it is only when we are clear what the data is we are trying to explain that we can begin to give it a physical explanation. And the data we want to explain is a disposition to emit verbal behaviour in certain circumstances. Quine has no problem with mental explanation per se as long we are clear precisely what it is we are talking about:

Mental entities are unobjectionable if conceived as hypothetical physical mechanisms and posited with a view strictly to the systemisation of physical phenomenon...Any vagueness or complexities that might obstruct that objective should be minimised. (1974, 32)

So if a person wants to explain something in terms of mentalistic idioms then this is fine as long as it is a stop gap to reducing such talk to physiological explanation; however, familiar talk about things such as ideas and meanings can and in fact does obscure what the data is we are trying to explain.

Chomsky explicitly argues in various different sections of his writings that he is not endorsing irreducibly mentalistic posits, as can be seen by the following claim:

When I use such terms as ‘mind’, ‘mental representation’, ‘mental computation’, and the like, I am keeping to the level of abstract characterization of the properties of certain physical mechanisms, as yet almost entirely unknown. (1980a, 5)

Furthermore he recommends that when studying ‘the mind’ we use the Galilean method; in other words we ‘construct abstract models that are accorded more significance than the ordinary world of sensation...’ (ibid p.9). So he would claim that far from endorsing irreducibly mentalistic posits he is rather explaining ‘mental events’ in terms of abstract models, in the exact same manner he would explain the motion of the earth in terms of abstract models. Gibson claims that Quine’s difficulty with Chomsky is that he relies uncritically on mentalistic semantics. Quine endorses Gibson’s interpretation by saying Gibson ‘Understands my position so fully, presents it so clearly, and defends it so cogently’; however it should be clear by now that Chomsky would deny endorsing uncritical mentalistic semantics.

So here Quine is partially misreading Chomsky’s intentions when he speaks of mentalistic entities. A possible source of Quine’s misinterpretation of Chomsky as a dualist can be traced to a common reading (rather misreading) of Chomsky’s famous ‘‘Review of Skinner’s Verbal Behaviour’’⁸. Conventional wisdom has it that when Chomsky critiqued Skinner’s attempt to explain language behaviouristically, he was vindicating ordinary folk psychology. Chomsky’s constant talk of language being a

faculty of the *mind* seems to further strengthen this view. However Chomsky was not making this point at all; rather his point was that neither Skinner's account nor ordinary folk psychology adequately explains the complexity of human behaviour. His arguments against Skinner centre on the fact that all his supposed explanations of things such as beliefs, hopes etc. in terms of conditioned response actually take us no further along the line in terms of explanation than the original folk terms. He claims that all Skinner does in effect is re-label folk terms in pseudo-scientific language (RSVB, 13) He agrees with Quine that the mental (qua ordinary folk psychology) scarcely deserves the name of an explanation. However, when he uses the term 'mental' he just means the brain at a different level of abstraction. In his criticism of Skinner Chomsky makes the following claims about Skinner's terminology which make explicit his views on ordinary mentalistic discourse:

It is not unfair, I believe, to conclude from Skinner's discussion of response strength, the 'basic datum' in functional analysis, that his 'extrapolation' of the notion of probability can best be interpreted as, in effect, nothing more than a decision to use the word 'probability', with its favourable connotations of objectivity, as a cover term to paraphrase such *low status* words as 'interest', 'belief', and the like'...Naturally the terminological revision adds no objectivity to the familiar 'mentalistic' modes of description...A mere terminological revision, in which a term borrowed from the laboratory is used with *the full vagueness of ordinary vocabulary*, is of no conceivable interest(1959 p.22). (Italics added by me).

Chomsky clearly is not asking people to explain human behaviour in terms of *low status vague* mentalistic discourse; rather he is claiming that Skinner's behaviourism should not be adopted because it remains at this paltry level of explanation. Chomsky claims that generative syntax takes us beyond this crude level of mentalistic explanation. So he aims to describe the rules that govern linguistic competence, and presumes these rules are instantiated in the brain. However, like Quine, he agrees that at present we do not know enough about the brain to track this process in any detail.

Obviously, analysing whether Chomsky's views on Skinner are correct is beyond the remit of this thesis. However, his criticisms of Skinner reveal certain facts about Chomsky's view of the mind. First, Chomsky is an eliminativist about the mental in the sense of ordinary folk psychology. His criticism of Skinner is not that he tries to do away with mentalistic idioms; it is that Skinner tries to do away with the mentalistic idiom by replacing it with something just as vague. Quine, like Skinner, tries to do away with the ordinary folk conception of meaning, mind etc. and replace with it with a more scientific conception of human behaviour. Chomsky also rejects Quine's behaviouristic analysis of language. However, his rejection stems not from the fact that Quine is guilty of being an eliminativist about the mind, but rather because Quine focuses on performance issues and believes that we can explain the behaviour of humans in terms of the probability of an utterance in particular circumstances. This is a debate on the correct methodology to be adhered to when studying the mind; it is not a debate about ontology. Both theorists correctly believe that our ordinary intuitive folk psychology is not a reliable tool for the study of any aspect of the world. So in terms of ontology, neither Quine nor Chomsky are ontological dualists. Furthermore, neither thinker is an ontological naturalist. Despite their mutual misreading, the position adopted by both thinkers is methodological naturalism.

PART 2: THE ANALYTIC AND THE SYNTHETIC

SECTION 1: THE OBJECT OF ATTACK

Despite the fact that Chomsky and Quine are both methodological naturalists, they nonetheless disagree on a number of issues. In this section I will consider one such disagreement: their differing views on the status of the analytic/synthetic distinction. I will show that this dispute between them is an entirely naturalistic dispute. Traditionally, philosophers who used analyticity in their theorising did so in the service of a non-naturalistic philosophy. Such thinkers used analyticity as an explanation of non-naturalistic notions such as a-priori knowledge and our grasp of metaphysical necessity. One of my main reasons for discussing the dispute between Chomsky and Quine on analyticity is to illustrate a disagreement between both thinkers which, despite superficial appearances is entirely naturalistic. I will evaluate the evidence put forth by both thinkers and show that Chomsky has not offered sufficient evidence to justify his claim that analytic connections exist in natural language.

Quine's attack on the analytic/synthetic distinction in his *Two Dogmas of Empiricism* is probably one of the most famous papers in the history of analytic philosophy. However, while the paper has been extremely influential, it has long been noted that the object of Quine's attack was unclear even to Quine himself. In *Two Dogmas* Quine seemed to be attacking three different sources (1) Truth in virtue of meaning; (2) Necessary Truths; and (3) A Priori Knowledge. This fact was understandable because at the time Quine wrote *Two Dogmas* most philosophers believed that an analytic truth was a truth which was both an a priori truth and a necessary truth. Logical Positivists had argued that the only sensible way to account for supposedly a priori truths or necessary truths without postulating some mysterious

faculty of intuition was to do so in terms of analyticity. So it is understandable that Quine would have believed that by attacking analyticity he was also attacking the a-priori and necessity.

Post Saul Kripke few philosophers would conflate analyticity, necessity, and the a-priori. It is worth bearing in mind the above distinctions when evaluating Quine's arguments because doing so will help us to key in on what his arguments are directed towards, and on whether his arguments are effective. The a-priori/a-posteriori distinction is an epistemological distinction. The necessary/contingent distinction is a metaphysical distinction. The analytic/synthetic distinction is between purely semantic truths and truths which rely on semantic facts and extra-linguistic information. In the 1950's, Quine, like most philosophers, believed that these distinctions lined up. He furthermore believed that, since analyticity was the only way of explicating the a-priori and necessity, then by attacking analyticity he was attacking the other distinctions. Quine was not merely concerned with attacking analyticity in so far as it was used as an explanation of the a-priori and necessity, he was also concerned with attacking the very notion of a purely semantic truth. He believed that truth by virtue of the meaning of the terms was a senseless notion. I will now discuss his actual arguments against analyticity in *Two Dogma's* focusing firstly on his arguments against the semantical notion of analyticity.

SECTION 2: TWO DOGMAS OF EMPIRICISM: A SUMMARY

When Logical Positivists spoke of analytic truths they included truths which we would group as necessary truths, a priori truths, and purely semantic truths. Under the banner of analytic truths one would find philosophically important notions such as truths of logic, truths of geometry and arithmetic, as well as trivial semantic truths.

So, for example, one would find complex mathematical statements grouped under the same category as a trivial statement such as ‘All bachelors are unmarried men’. Quine who was working within this logical positivist tradition grouped trivial analytic statements together with more philosophically important analytic truths.

In the first four sections of “Two Dogmas” Quine was critiquing the very notion of an analytic truth. Quine argued that we could make sense of what he called Logical Analyticity, an example of which is (1) ‘All bachelors are bachelors’. However he stated that he could make no sense of analytic truths which results from substituting synonyms for synonyms, for example (2) ‘All bachelors are unmarried men’. We will call analytic truths derived by substituting synonyms; Meaning Analyticity. Quine argued against Meaning Analyticity because he could find no cogent way of explicating what it amounted to. Any attempt to define meaning analyticity relied on using notions which themselves could only be made sense of if we helped ourselves to the prior notion of meaning analyticity. However, since analyticity was the very notion which we were trying to define, if the other terms which we use to define it themselves rely on the notion of analyticity to be made sense of, then our explanation will be circular. Quine’s objection to the notion of analyticity is usually called his ‘circle of terms’ argument. We cannot define analyticity without terms such as “synonymy”, “necessarily”, “meaning” etc. and these terms can only themselves be defined in terms of the ill-understood notion of analyticity. So Quine argued that, since we cannot define analyticity in a non-circular manner and the notion makes no sense behaviourally, then we have no justification to argue that analytic sentences, as traditionally conceived, exist.

The primary aim of the first four sections was to criticise the notion of meaning analyticity. In these sections Quine was concerned with attacking the notion

of purely semantic truth. In the final two sections Quine is attacking what he takes to be the dogma of reductionism. He begins by attacking the verification theory of meaning. He notes that, according to the verification theory of meaning, the meaning of a statement is the method of confirming or infirming it (1951,37). He then notes that an analytic statement is that limiting case which is confirmed no matter what (ibid., 37). Quine claims that it is possible to argue for a conception of analyticity using the verification theory of meaning. One could say that two statements are analytic if they have the same conditions of verification.

Quine rejects the definition analyticity in terms of conditions of verification because of deficiencies which he sees in the verification theory of meaning. He famously argued that the unit of empirical significance is the whole of science (ibid., 42). He pointed out that all empiricist attempts at reduction had failed. Furthermore, he noted that such failures were failures in principle; all such attempts relied on undefined terms to construct their reduction. It was because of this failure that Quine argued for his famous web of belief. When cashing out the nature of the web of belief Quine claimed that: 'our statements about the external world face the tribunal of experience not individually but as a corporate body' (ibid., 46).

An important consequence of this holism is its effect on the nature of falsification of our scientific theories. Since our total theory of the world is an interconnected web of belief, and this web faces the tribunal of experience as a corporate body then a falsification of a prediction made by the theory will reverberate throughout the theory. However, Quine notes that if we are faced with a supposed falsification of a prediction, how we modify the theory which made the false prediction is to be dictated by pragmatic considerations. So, for example, since mathematical and logical principles are so deeply connected to every area of our

scientific theories they will be the last thing that we will modify if our theory makes a false prediction. Nonetheless there is no reason in principle that we cannot revise mathematical or logical laws if it seems the best way to accommodate recalcitrant experience.

Laws of mathematics have traditionally been viewed as being special in that they could not be explained empirically⁵. Kant claimed that mathematical laws were a priori synthetic truths, because, while they were a priori true, they were not analytic truths. An analytic truth, was for Kant, a statement where there was nothing in the predicate concept that was not already contained in the subject concept. However, Kant noted that in the case of mathematical truths we could not explicate them in terms of analyticity. So, for example, if we consider '2+2=?', according to Kant, no analysis of preceding concepts will give us the answer '4'. Kant claimed that if we want to discover an answer in maths we have to construct it in intuition, merely analysing concepts will get us nowhere.

Logical Positivists argued, contrary to Kant, that mathematical truths could be reduced to logical truths. They recognised two types of truth, analytic truth and synthetic truth. Mathematical truths, like truths of logic, were viewed by them as analytic truths.

Quine's web of belief argument, with its emphasis on the fact that mathematical and logical truths are revisable in principle but in practice are rarely ever revised, is pertinent to the considerations of Kant and the Logical Positivists. To explain the special status of mathematics Kant had to postulate a mysterious kind of knowledge (a priori synthetic knowledge). Logical Positivists tried to explain mathematics in terms of logic which was in turn explained by linguistic convention.

⁵ John Stuart Mill was an exception to this rule.

Quine demonstrated that the doctrine of truth by convention was dubious because in order to *consistently* apply any logic we would need to presuppose a logic prior to such an application (1934). While Kant's postulation of a priori synthetic knowledge was never made sense of in a non-mysterious way, Quine's web of belief picture explained why mathematical and logical truths seemed to have a special status in our overall theory of the world. Furthermore, it managed to avoid the difficulties which plagued the pictures of the Positivists and Kant.

Quine's web of belief argument shows that we do not need to appeal to analytic truths in the sense of truth in virtue of meaning to explain the apparent specialness of mathematics and logic. It furthermore shows that we do not need to appeal to a-priori synthetic knowledge to explain what is apparently special regarding mathematics. Quine's picture of mathematics and logic being deeply embedded into our total theory of the world explains their apparent specialness adequately. Furthermore, by admitting that basic laws of mathematics and logic can be revised, he is making sense of the history of science in a way that people like Kant do not.

However, as Putnam pointed out in his paper "The Analytic and the Synthetic", while Quine's argument for the web of belief is persuasive as an explanation of non-trivial analytic sentences, it is far from persuasive when it comes to trivial analytic sentences. While it is sensible to explain away truths of mathematics conceived as analytic truths by appeal to them being deeply embedded in our web of belief, it is not sensible to claim that a trivial statement such as 'All bachelors are unmarried men' is deeply embedded into our web of belief. I will explore this weakness with Quine's explanation of analyticity later in Section 4. Firstly I will explicate Chomsky's difficulty with Quine's claim that there is no sensible distinction to be drawn between analytic and synthetic sentences.

SECTION 3: CHOMSKY'S DEFENCE OF THE DISTINCTION

It appears, then, that one of the central conclusions of modern philosophy is rather dubious: namely, the contention- often held to have been established by the work of Quine and others- that one can make no principled distinction between questions of fact and questions of meaning, that it is a matter of more or less deeply held belief. (2000b, 63)

As a linguist, Chomsky was primarily with whether an analytic/synthetic distinction was, as a matter of empirical fact, a distinction that ordinary speakers of natural language recognised. He argued contra Quine that analytic connections do exist in natural language, and that it is the job of the cognitive scientist to study such connections. As in the case of syntax, we begin with people's intuitions; in the case of analytic connections, we need to test whether people can reliably distinguish sentences which are analytic from sentences which are synthetic. So the question arises: would an ordinary speaker of English be able to distinguish between an analytic sentence like 'All Bachelors are Unmarried men' from a synthetic sentence like 'All Bachelors like Seinfeld'? For Chomsky's claim that analytic connections are a fact of natural language to be shown to be correct, experiments need to be conducted. However, experimental research on the analytic/synthetic distinction has not exactly been overwhelming. There have been some experiments on the distinction, for example: Apostel, L., W. Mays, A. Morf, and J. Piaget. *Les liaisons analytiques et synthétiques dans les comportements du sujet*. Arne Naess. *Interpretation and Preciseness*, as well as Katz and Fodor's paper, "The Structure of a Semantic Theory". These studies do imply that people have a gradualism of statements that they find intuitively synonymous. And this view seems to support Chomsky's claim that people do have intuitive conceptions of what sentences in natural language are analytic and what ones are synthetic.

Chomsky argues that since the experiments show that people intuitively distinguish between analytic and synthetic sentences, then we need to explain this fact. His explanation centres on poverty of stimulus considerations. He argues from the speed that children acquire words (about 12 words a day at peak periods), and the incredible complexity of the words acquired, to the claim that the children must be merely labelling concepts they are born with. From his argument for innate concepts, he draws the following conclusion:

This would appear to indicate that the concepts are already available, with much or all of their intricacy and structure predetermined, and that the child's task is to assign labels to concepts, as might be done with limited evidence given sufficiently rich innate structure. And that these conceptual structures appear to yield semantic connections of a kind that will, in particular, induce an analytic-synthetic distinction, as a matter of empirical fact. (2000b, 62)

On the face of it, this argument simply does not work. Chomsky's poverty of stimulus argument for concepts relies on the assumption that when children acquire words this ability indicates a grasp of a complex concept which the word means. So, for example, it implies that when a child acquires a word such as 'Mama', the child understands the same complex concept which adults do when they explicate what the concept *MAMA* means. Chomsky has never provided any evidence to support this wildly implausible claim. If we take away the unproven claim that the words children are learning at peak periods of language acquisition express full-blown adult concepts, then the acquiring of 12 words a day is less impressive than Chomsky claims. If a child using a word such as 'mama' had only a simple understanding of what it meant, then the fact that the child was learning such words so fast would not seem so miraculous. And there would therefore be no need to postulate innate concepts to explain the speed of word acquisition.

While there is some evidence that children from as young as four months of age demonstrate some conceptual understanding of concepts such as *OBJECTS*,

CAUSATION and *AGENT*, this evidence does not support Chomsky's more radical claims about concepts. I will discuss the evidence for children's conceptual abilities in chapter three; here I will merely note that it is at best suggestive. The evidence does show that children have *some* concepts prior to learning language. However it does not conclusively demonstrate that these concepts are innate, nor that children possess *most* of their concepts prior to learning language. Chomsky is making much more radical claims than that some abstract concepts such as object, or causality are innate. In a personal communication with me he made the following claims:

Also, there is good evidence that innateness of concepts goes far beyond the philosophically interesting examples that you mention. It's hard to imagine how else people could acquire the meanings of the simplest words in the language -- "river," "tree," "person,".... They have rich properties for which there is no empirical evidence for the child. (Chomsky: Personal Communication)

Furthermore in his *New Horizons in the Study of Language and Mind*, Chomsky made even more radical claims about the innate conceptual abilities of Children. Chomsky considered an argument from Putnam (1988) that evolution could not have equipped us with an innate set of concepts including ones like *Carburettor* and *Bureaucrat*. Putnam argued that evolution could not have equipped us with these concepts because in order to do so evolution would need to be able to anticipate all future contingencies. Chomsky replied to this by noting that Putnam's argument is incorrect because we do not need to assume that evolution anticipated *all* possible contingencies, just the particular ones under question.

Chomsky claimed that a similar argument to Putnam's was used in immunology and that recent work by Niels Kaj Jerne challenges this idea:

Notice that a very similar argument had long been accepted in immunology: namely, the number of antigens is so immense, including every artificially synthesized substances that had never existed in the world, that it was considered absurd to suppose that evolution had provided an 'innate stock of antibodies', rather formation of antibodies must be a kind of 'learning process'.

But this assumption might well be false. Niels Kaj Jerne won the Nobel Prize for his work challenging this idea... (ibid, 65)

This analogy proves nothing, it is not in doubt that it is *possible* that people are born with an innate stock of concepts like *CARBURETTOR* and *TREE*. What is in doubt is whether there is any evidence that supports the claim that such concepts are innate. Chomsky has provided none whatsoever here. Chomsky vaguely gestures in the direction of poverty of stimulus considerations to support his conjecture about innate concepts:

Furthermore, there is good reason to suppose that the argument is at least in substantial measure correct even for such words as *Carburettor* and *Bureaucrat*, which, in fact, pose the familiar problem of poverty of stimulus if we attend carefully to the enormous gap between what we know and the evidence on the basis of which it is known. (2000b, 65)

Here Chomsky is using a poverty of stimulus argument similar to the argument he uses in the area of syntax. I will discuss the poverty of stimulus argument in the area of syntax in chapters two and four. The poverty of stimulus argument in syntax has the merit of being precise and testable. Chomsky has provided no such explicit model of a poverty of stimulus argument for innate concepts. So we have no reason to believe Chomsky's radical claims about innate concepts until more evidence is provided.

It is of course possible that Chomsky is correct that children are born with innate concepts such as: *CARBURETTOR*, *TREE*, *BUREAUCRAT*, *RIVER*, etc.; however an incredible amount of evidence is needed to support such an incredible claim. Chomsky vaguely points to poverty of stimulus considerations. However, he provides no details of where the supposed gap between the knowledge children display of the concept *CARBURETTOR* and the data the child receives when learning the concept lies. So until he spells out his poverty of stimulus argument in more detail his claims on these matters should be viewed with extreme scepticism.

More importantly, even if Chomsky did provide evidence that such concepts are innate, it would not follow that they will automatically yield analytic connections. Jerry Fodor has long agreed with Chomsky that most of our concepts are innate; however he does not think that analytic connections exist in natural language. It is worth briefly considering Fodor's views on innateness and analyticity because doing so will help us better understand Chomsky's views.

Fodor agrees with Chomsky that most of our concepts are innate. However he does not use the same arguments as Chomsky to reach this conclusion. While Chomsky's poverty of stimulus argument points to a gap between what we know and the inability of the data of experience to account for our knowledge, Fodor argues that because of certain facts about the structure of concepts, all known theories about how concepts are *learned* are incorrect in principle. He claims that there are two different ways of accounting for our knowledge of concepts. The first one is to say that most of our concepts are definitions which are defined in terms of primitive concepts. The primitive concepts are either defined as sensory primitives such as *RED*, *SQUARE* etc., or as abstract concepts such as *CAUSATION*, *AGENCY*, and *EVENT* etc. What the primitive concepts are will depend on the nature of the theory being expounded. So for example, the empiricist philosopher David Hume argues that the primitive concepts are sensory experiences. While the linguist Stephen Pinker argues that the primitive concepts are abstract concepts such as *CAUSATION*, *AGENCY*, and *OBJECT* which we are born with knowledge of. Fodor thinks that Pinker, Hume and all definitional theorists are wrong because concepts cannot be explicated in terms of definitions.

The other theory of concept acquisition thinks that our concepts are basically prototypes which are learned statistically. Fodor argues against this view because he

thinks that the prototype theory cannot account for the fact of compositional concepts. I will not here consider Fodor's arguments against prototype theory because they are not relevant to the concerns of this thesis. Furthermore I will not consider the fact that Fodor has lately retracted his views on radical nativism because again this is not relevant to the overall argument of the thesis. I am only concerned with explicating why Fodor does not follow Chomsky in arguing that analytic connections follow from the fact of innate concepts. Fodor's reason for not believing that analyticity follows from innate concepts is that he does not think that innate concepts have internal structure. His argument that concepts do not have internal structure is derived from his belief that concepts are not definitions. Fodor essentially uses three different arguments against the claim that concepts are definitions:

(1) **Lack of definitions:** In general we have not found any definitions for concepts. There are millions of concepts and very few definitions (maybe fifty or so). (1998, 46)

(2) **Developmental argument:** All concepts cannot be definitions. Some of the concepts must be primitive. The concepts which are definitions must be defined partly in terms of the primitives. This being so we must assume that there is a developmental stage when children only know primitive concepts. However we have no evidence to support this claim. Therefore, concepts cannot be definitions. (ibid, 47)

(3) **Production argument:** If concepts were definitions, then sentences which contained complex concepts would be harder to produce than those which contained simpler concepts. The reason is that a complex concept would take longer to call to mind the definition of when speaking the sentence. However the experimental evidence indicates that the production of sentences involving complex concepts takes no longer than the production of sentences which contain less complex concepts. (ibid, 49).

Fodor believes that his arguments against definitions are decisive. He claims that without definitions, concepts have no internal structure, and furthermore if concepts have no internal structure, then there are no analytic sentences. Kant famously defined an analytic sentence as a sentence in which the predicate is entirely contained within the subject. However obviously if a concept has no internal structure, then nothing can be contained within it. For this reason, Fodor thinks that there are no analytic sentences. He notes that one of the main arguments for definitions is that they explained people's felt intuitions of analyticity. So this gives us two options: (1) accept that people's intuitions of analyticity are a reason to believe in conceptual connectedness; (2) explain away the intuition. Since Fodor has offered reasons to deny that there are complex concepts, he opts for number (2), to explain away the intuitions of analyticity.

EXPLAINING AWAY ANALYTICITY INTUITIONS:

As we saw above Quine explained away our intuitions of analyticity in terms of his web of belief picture. This worked for supposed non-trivial analytic sentences like $F=MA$. However it was less plausible as an explanation for apparently more trivial analytic sentences like "All bachelors are Unmarried men". We cannot plausibly explain away cases like "All Bachelors are Unmarried men" by saying that they are deeply embedded into our total web of belief.

As we have already seen, Putnam claimed that Quine's argument worked as a criticism of the a priori but not as a criticism of apparently trivial analyticity. He claimed that a sentence such as "All Bachelors are Unmarried men" is an analytic sentence, and its analyticity derives from the fact that it is a one criterion concept. The concept *BACHELOR* is not according to Putnam connected to any other concepts in

our web of belief other than to the concepts *UNMARRIED* and *MEN*. Putnam argues that sentences which people intuitively find analytic but which are not deeply connected to our web of belief, can be explained by his notion of one-criterion concepts. Fodor disagrees with this explanation because he feels that it is a circular argument. In order to define which concepts are learned by one criterion, you need to know that it is analytic already. So your explanation of analyticity presupposes the existence of the very thing we are doubting. So we are back to Quine's circle of terms again.

Fodor offers a different account of one-criterion concepts which he thinks is less problematic than Putnam's is. Fodor agrees with Putnam that one-criterion concepts exist. He furthermore thinks that he can use these one-criterion concepts to explain people's intuitions of analyticity. However, Fodor argues that since he is not using one-criterion concepts to explicate analyticity, then his story is not circular in the way that Putnam's is. Here is Fodor's story in a nutshell:

Suppose you think the only epistemic route from the concept *C* to the property it expresses depends on drawing inferences that involve the Concept *C**. Then you will find it intuitively plausible that the relation between *C* and *C** is conceptual; specifically, that you can't have *C* unless you have *C**. And the more you think that it is counterfactual supporting that the only epistemic route from *C* to the property it expresses depends on drawing inferences that involve the concept *C**, the stronger your intuition that *C* and *C** are conceptually connected will be. (Ibid, 83)

Fodor argues that people's intuitions of analyticity are really just epistemic intuitions which are confused for semantic connections. So for example I can only know that someone is a bachelor by knowing that he is an unmarried man. According to Fodor, I know that all bachelors are unmarried men because of epistemic access to the facts in the world, not because of the internal structure of concepts. He holds that people's intuitions of analyticity are illusions which result from conflating epistemic properties with semantic properties.

So we can see from Fodor's argument that analyticity does not follow from innateness. If the innate concepts are unstructured, then they will not automatically yield analytic connections. Chomsky assumes that the children's innate concepts are structured and as a result will automatically yield analytic connections. Furthermore unlike Fodor and Quine, he takes it for granted that people's intuitions of analytic connections are good evidence that such connections exist.

In a personal communication I asked Chomsky why he believed that analyticity followed from innateness:

You're quite right that it isn't obvious, but I think it is correct. It's an empirical issue, in principle, but a hard one to investigate -- and like other empirical questions, certainty is unattainable, just a high degree of plausibility. Personally, I think that serious inquiry would reveal that "tomorrow is two days after yesterday" or "my uncle is male" or "nightmares are dreams" and much else would withstand the harshest tests, and I don't think that Quine or Fodor have suggested any reason to doubt these expectations. (Chomsky: Personal Communication)

Chomsky's reply is instructive because he simply claims that certain constructions are obviously analytic and that any future tests would confirm this belief. He does not answer the question about why analyticity follows from innateness. He merely claims that it is in principle an empirical question. Presumably his reasoning is that people do recognise a distinction between analytic and synthetic sentences, and he believes that people are born with a massive amount of innate concepts. From this he believes it follows that the structure of these innate concepts are our best explanation for our intuitions of a distinction between analytic and synthetic sentences. However this vague argument relies on an unproven claim about innate concepts, and a further unproven claim about the structure of these purported innate concepts. So Chomsky thus far provides no real explanation for people's felt intuitions of a distinction between analytic and synthetic sentences.

Chomsky appeals to the fact that certain sentences are obviously analytic, presumably he is here claiming that such sentences are intuitively felt to be analytic. However obviously the fact that people have intuitions of analyticity does not tell us whether Fodor's story is correct or if Chomsky's is.

Since Chomsky has offered no compelling evidence to support his claim that humans are born with innate concepts, a debate between him and Fodor on whether innate concepts have internal structure is a nonstarter. Chomsky has offered no evidence for innate concepts, so he cannot support his views on analyticity by appealing to innate internally structured concepts. Fodor has offered some arguments for innate concepts⁶, however they offer little support to Chomsky's position.

It is of course open to Chomsky to use Fodor's arguments for innate concepts; however a vital part of Fodor's argument, is that concepts are not definitions, so they have no internal structure. So if Chomsky did use Fodor's arguments for innate concepts, he would be left in a situation where he could claim that concepts are innate but have no internal structure, so do not yield analytic connections. Either approach leads to the conclusion that analytic sentences do not exist. So Chomsky has offered us no real reason to believe that we can derive analyticity from innate concepts.

We have shown that analyticity does not follow from innate concepts. It is worth noting that the converse is also true. If one is denying that innate concepts exist one is not thereby denying that analytic connections exist in natural language. Empiricists such as Ayer, Carnap and Hume who are not committed to the view that concepts are innate, nonetheless argue that analytic connections exist in natural language. Ultimately, then, Chomsky's arguments for innate concepts are irrelevant to facts about analyticity unless he can provide evidence that such innate concepts are

⁶ I will not here evaluate whether Fodor's arguments for innate concepts work because it has no relevance to the argument of the thesis.

structured in such a way as to yield analytic connections. Thus far Chomsky has provided no evidence about the structure of these supposedly innate concepts, so we therefore have no reason to believe that innate concepts lead to analytic connections in natural language.

Independent of his unproven claims about innate concepts yielding analytic connections, Chomsky's arguments for analytic connections in natural language focus on three primary facts. One is that ordinary language speakers can readily tell whether a sentence of natural language is true by virtue of the meaning of the words involved and extra linguistic fact, or whether it is made true entirely by facts of meaning. His second reason is that he has never seen a convincing counter example to clear cases of analytic sentences, such as 'Tomorrow is two days after Yesterday'. And his third reason is that he thinks that particular sentences of a relational structure such as the sentence "If John killed Bill then Bill is Dead" are more difficult to find counter examples to than the simple examples focused on in the philosophical tradition such as "All Bachelors are Unmarried men".

Chomsky's first claim that people intuitively recognise a distinction between statements which are synonymous and those which are not has been tested experimentally through the using of questionnaires. The small literature does seem to indicate that people have an intuitive sense of a distinction which loosely corresponds to what is traditionally called an analytic/synthetic distinction. However such studies involve showing people different sentences which are categorised into two different groups. The fact that people categorise in a way consistent with an analytic/synthetic distinction proves nothing. Such categorisation may reflect an epistemic categorisation rather than a semantic one. The various different studies do not distinguish between the Quine/Fodor story and the Chomsky story. The Quinean

model would predict that people would have an intuitive sense of a distinction between truths like “ $2+2=4$ ” and “Most dogs are less than six feet tall”. The difference between Quine and Chomsky on the question of Analyticity is whether; a person’s intuitive sense of Analyticity can be explained best in terms of an innate system of concepts, or in terms of deeply embedded beliefs in a person’s holistic theory of the world. The weak point of Quine’s views on analyticity is that his web of belief story works well for epistemologically significant concepts such as “ $2+2=4$ ” but not for “All Bachelors are Unmarried men” which cannot be reasonably claimed to be deeply embedded into our total theory of the world. However Putnam’s one criterion concepts as modified by Fodor works perfectly to capture trivial cases of analyticity. Fodor has shown that such cases can be viewed as epistemic knowledge confused with conceptual knowledge. The fact that people have intuitions of analytic connections in natural language is not that significant, such intuitions could be epistemic knowledge confused with conceptual truths. They can be explained away as intuitions of truths deeply embedded in to our web of belief, or one-criterion concepts.

Chomsky’s second reason is not very convincing, as philosophers have shown over the last fifty years that some statements which seem immune to revision are in principle revisable; see for example Putnam on the sentence “All Cats are Animals”. Now Chomsky could reply to this that while some sentences which are analytic have been shown to be in principle revisable, the paradigm statements have not been shown to be revisable. If he were to argue so, in order to make his claim a testable one, he would need a list of paradigm cases of analytic sentences. He would then need to show how such sentences could withstand the probing of the most strenuous of tests. Chomsky has never done anything like this. His remarks on analytic sentences, remain just that, remarks; and remarks which are unproven.

Chomsky has also claimed that philosophers erroneously focus on analytic sentences with to simple a structure; he claims that sentences of a relational structure provide much more evidence for analytic sentences. To test Chomsky's claims we will need to consider some examples of sentences of a relational structure and test whether they support his claims over Quine's. He has used two examples over and over again, they are: (1) If John killed Bill, then Bill is Dead. (2) If John persuaded Bill to go to college, then John caused Bill to intend to go to college. Here Chomsky discusses these sentences:

Notice again that we appear to have connections of meaning in such cases as these; we have a rather clear distinction between truths of meaning and truths of fact. Thus, if John persuaded Bill to go to college, then Bill at some point decided or intended to go to college and did so without duress; otherwise, John did not persuade Bill to go to College. Similarly if John killed Bill, then Bill is dead (though John may or may not be depending on the facts). These are truths of meaning not of fact. (2000b, 62)

Sentences 1 and 2 are examples of sentences with a relational structure. The use of the predicates 'killed' and 'persuade' in the antecedents will have direct bearing on what words are allowed in the consequent. Take, for example, Sentence 1 above- Let us consider the sentence as an argument schema: If X killed Y then Y is....?. What Chomsky seems to want to say here is that the only way to make the argument schema true is to fill 'dead' into the blank or something synonymous with 'dead'. Of course such an interpretation is patently false as can be seen by the following; we can insert 'is not going to play football next week', 'is not going to run a marathon next week' etc. and the truth value will remain invariant. So it could be claimed that the relational structure does not in any obvious sense indicate that, from the fact of John killing Bill, we have to infer that Bill is dead. However, upon closer inspection it becomes clear that anything we fill into the blank must be consistent with the fact that Y is dead. So, for example, if one were to offer the following to fill in the blank 'is

going to sing his favourite song tomorrow’, one would be constructing a false argument. And the reason is that singing a song is not consistent with being dead. The two attempts to fill in the blank above which result in true arguments work because obviously *not* doing something, is largely consistent with being dead. From this we can conclude that any blank we fill into the consequent must at least be consistent with being dead, however obviously it does not have to be synonymous with being dead. So if we consider ‘X killed Y so Y is -’, only something consistent with being dead can be slotted into the consequent and the truth value remain invariant. So any conditional which has an antecedent such as ‘Y being killed, so Y is...’ must have a consequent which is consistent with Y being dead or the conditional will be false.

If one puts ‘is dead’ into the consequent one gets an analytically true statement. Likewise, if one puts ‘is not playing tennis’ one also gets an analytic sentence. This construal of the argument leaves us in a situation where we cannot equate the analyticity with synonymy. Obviously the statement that ‘X killed Y’ is not synonymous with the statement that ‘Y is not playing tennis’, despite the fact that Sentence 1 ‘X killed Y so Y is not playing tennis’ is an analytic statement. Chomsky could argue that Sentence 1 is analytic because of the internal structure of the complex concept *KILLED*. He could argue that the complex structure of such concepts which we are born with will automatically tell us what inferences can be derived from it purely based on the meaning of the concept⁷. So our language faculty determines what arguments can be filled into the schema ‘X killed Y so Y is-’ in order to yield analytic connections. Here we have an argument which seems to show how we can go from an innate concept such as *KILLED* to construct analytic sentences

⁷ For a detailed discussion of issues connected to the above discussion see Fodor 1970 ‘Three Reasons for Not Deriving Kill from Cause to die’, and Pietroski’s ‘Small Verbs, Complex Events, Analyticity without Synonymy’. The debate between Pietroski and Fodor would take us too far a field from the concerns of this thesis but IS worth reading in its own right.

based on the internal structure of the concept. However, such an argument again relies on assumptions about innate concepts which have not been proven correct. So if Chomsky wants to construct such an argument he needs to find evidence to support his belief in innate concepts such as *KILLED*. Independent of the argument from innate concepts all that is left is the claim is that some analytic statements are difficult to refute, but of course this claim is largely consistent with Quine's claims about analytic statements being deeply embedded into our overall theory of the world.

So Chomsky's claim that sentences of a relational nature provide stronger kinds of evidence for analytic connections in natural language than the simpler cases that philosophers typically talk about has been shown to be erroneous. His assumption that this is so relies heavily on his belief in innate concepts. The claim that sentences of a relational structure seem to offer cases of analytic connections can be explained in terms of how embedded notions like *KILL* and *INTENTION* etc. are to our overall theory of the world.

The above discussion shows that Chomsky has offered no compelling arguments which demonstrate that Quine's critique of an analytic/synthetic distinction is incorrect. It is possible in principle that Chomsky's positive views on analyticity are correct; however, we do not as of yet have any empirical evidence to support Chomsky's claims on this matter. In the next section I will try to explicate what type of analyticity Chomsky is committed to if we assume he is correct that analytic sentences exist in natural language as a result of certain constraints which are imposed by the concepts which we are born with. In the next section I will assume that Chomsky is correct on this point, and consider how much analytic connections being encoded in the brain of a subject meets the criterion of analyticity as traditionally conceived.

PART 3: ANALYTICITY NATURALISED?

SECTION 1: Types of Analyticity

In Part 2 I discussed Quine's arguments against the notion of analyticity. It was noted that Quine was attacking three notions, (1) Analyticity as a defence of the a priori (2) Analyticity as a form of necessity, (3) Analyticity as a purely semantic truth independent of collateral information. Chomsky's claim that an analytic/synthetic distinction was a fact of natural language was then discussed. Throughout the discussion I treated Chomsky's arguments as pertaining to semantic analyticity. By semantic analyticity I mean analytic truth as being true in virtue of the meaning of the terms and independent of collateral information. In this Part I will consider whether Chomsky's positive conception of analyticity can support a priori knowledge, and our grasp of necessary truths.

As we have seen above, when Chomsky is discussing language and mind he is doing so in a thoroughly naturalistic manner. Our language results from an innate module of the mind which he attempts to explain in a purely naturalistic manner. Chomsky tells an evolutionary story about the origins of language (See for example Chomsky, Fitch, and Hauser 2002). He argues that any analytic connections in natural language result from the internal structure of the innate language faculty. He also claims that human concepts have idiosyncratic features which were of no particular use to our ancestors so it is unlikely that they were formed through natural selection. He assumes that if our innate conceptual abilities were not formed through natural selection then they must have resulted from a random mutation. Obviously,

conceptual abilities which resulted from a random mutation will in no way provide us with any kind of a priori justification.

When discussing the rules which govern our language, Chomsky has been quite explicit on the fact that we are not justified in applying the rules of natural language. He argues that we just automatically implement them in the way a computer implements a programme:

In short, if I follow R, I do so without reasons. I am just so constituted. So far, these conclusions offer no challenge to the account discussed earlier. I follow R because S' maps the data presented into S*, which incorporates R; then I apply R blindly. There is no answer to the Wittgenstein Sceptic and there need be none. My knowledge, in this instance, is ungrounded....I have no grounds for my knowledge in any useful sense of the term and no reasons for following the rules: I just do it. If I had been differently constituted with a different structure of my brain, I would have followed a different rule. (1986b, 225)

In the above passage, Chomsky is responding to Kripke's rule following paradox. He is claiming that he does not have a solution to the paradox. He cannot say why calculating according to the plus function is justified whereas calculating according to the quus function is not justified. His answer is to point to the fact that we do just calculate one way as opposed to the other way, and to explain this fact by invoking innate constraints on the type of mathematical rules we are capable of following. The innate structure of our brain means that we interpret '+' in terms of 'plus', however if our brain was structured differently we may well interpret '+' in terms of 'quus'. I will not here discuss Chomsky's solution to the rule following paradox; my main aim is to show that his answer to Kripke shows the folly of reading Chomsky's conception of analyticity as an epistemological conception.

Chomsky has explicitly claimed that people are born with innate concepts some of which will as a matter of empirical fact yield analytic connections. He has furthermore claimed in a number of places, that concepts of natural language do not refer to entities in the mind-independent world in the manner in which people like

Quine, Frege, and Tarski believe they do. This being so, it is clear that on the Chomskian picture concepts like *BACHELOR* and *UNMARRIED MEN* do not mean the same thing because they pick out the same mind-independent entity, rather, they both mean the same thing because they both pick out the same concept.

This view of course puts Chomsky in a very weak position. It means that we cannot *justify* our sentences as being true by virtue of the meaning of the terms involved. Instead we just automatically categorise them so, and if our brain was constituted differently we would categorise them differently. A number of points follow from this immediately. The types of analytic connections which Chomsky defends have no epistemological significance. The analytic connections he is concerned with are not connections which justify us in holding certain beliefs true. So it could be argued that the analytic sentences which Chomsky is concerned with are of a philosophically uninteresting type.

However such an interpretation is complicated by the following quote from Chomsky.

These are truths of meaning, not of fact. The a priori framework of human thought, within which language is acquired, provides necessary connections among concepts, reflected in connections of meaning among words and, more broadly, among expressions involving these words... (2000b, 63)

When Chomsky talks about an *a-priori* framework of human thoughts and about *necessary* connections amongst concepts he seems to be making claims which his explicit theory would not allow. No contingent fact about our brains would justify us holding a belief true, while it is likewise highly unlikely that a contingent feature of our brain structure would somehow put us in touch with a necessary truth about the universe. So it seems that Chomsky's claim about humans being born with analytic connections amongst concepts is inconsistent with his claim that analytic truths provide a priori justification or knowledge of metaphysical necessity. Throughout the

remainder of this section I will endeavour to reconcile these seemingly inconsistent views.

Given that Chomsky has claimed elsewhere that our syntactic rules are not justified but are merely followed blindly, it is obvious that we should not interpret Chomsky's claims about an a priori framework of human thought epistemically. Rather we should instead interpret an a priori framework of thought to mean an innate framework of thought. But this of course leads to the question of what Chomsky means by *necessary* connections amongst concepts? Consider the concepts: 2, 4, +, = these can be combined to create the necessary truth $2+2=4$. Now philosophers such as Hume and Ayer have claimed that the truths of arithmetic are analytic truths. Given Chomsky's various discussions of analyticity, and his views on concepts expounded in his reply to Kripke, it seems that Chomsky also considers the truths of arithmetic as analytic, in the sense of being true based on the meaning of the terms involved. But it seems that this type of analyticity is not strong enough to support metaphysical necessity.

In order for our analytic truths to yield metaphysical necessity, there would have to be some sort of correspondence between the truths of meaning which are encoded in the brain and necessary truths which exist in an unchanging mind-independent world. If the connection were a mere contingent one, then it would have been just the result of time and chance. To grasp the distinction between contingent and necessary truths using the framework of possible worlds is sometimes helpful. According to Chomsky, it is a fact of the human brain that some of the concepts we are born with will automatically yield analytic connections. Surely, though, this being so, it is also possible that the human brain could have been constructed differently. It is obviously possible that humans could have been born with innate concepts but no

analytic connections between their concepts: call such creatures Fodorsapiens. If we think of the analogy between male and females who are both members of the species Homosapien, it is also possible that there were two different types of Homosapiens, who were distinct from each other but could reproduce with each other; call them Chomskysapiens and Fodorsapeins. Both strands of Homosapiens are descended from the same ape ancestors and they just diverged from each other because of a random mutation which did not affect their ability to reproduce with each other. We can imagine in our story that Fodorsapeins eventually died out because their slightly different DNA made them unable to fight off the common cold. The point is that there are many possible worlds which can be constructed in which it was Chomskysapiens who died out, or in which a random mutation changed which sentences the Chomskysapien finds analytic. The truths of meaning which Chomsky claims are component of the mind brain seem to be nothing more than contingent results of our evolutionary history, and his claim that they provide necessary connections remains unargued.

So it would seem that since the brain is the contingent product of its evolutionary history, then any concepts it is innately born with will correspond with a mind-independent reality, if at all, only coincidentally. Here, of course, the type of analytic truths will be of paramount importance. If we are to count the truths of arithmetic as analytic, then we can sketch a story of how an innate number sense would be vital to survival. We could cite how a creature who could not calculate whether the number of foes was equal to the number of friends would not survive very long. We could point to creatures such as monkeys and dogs who indicate a grasp of a number sense to show how this sense is probably a universal feature of any creature who wants to survive in a hostile environment.

When one tries to explain other analytic truths such as ‘All bachelors are unmarried men’, or ‘All dogs are animals’, sketching a story in terms of survival value seems more difficult. The idea of ‘All bachelors are unmarried men’ being a necessary truth or a truth which is justified independent of the facts seems unintelligible. However the idea that the sentence is analytic because ‘bachelor’ and ‘unmarried’ correspond to the same mentalese concept is obviously intelligible. So the question we are faced with is the following: since analytic connections are not limited to the language faculty, what is their place in the overall taxonomy of the mind which Chomsky has constructed so far?

In a personal communication Chomsky tried to clarify some of these questions for me. He claimed that:

In this context, “a priori” means what Konrad Lorenz called “biological a priori,” that is, part of the cognitive equipment with which an organism organizes and approaches external data, mapping it into experience, and then interpreting it. The issue of justification doesn’t arise, any more than it does in the case of “book” rhyming with “took” (in my speech). This doesn’t conflict with Boghossian. He is talking about something different. (Chomsky: personal communication).

So Chomsky here agrees with my analysis of his views on analyticity. His views offer nothing towards helping us clarify such notions as a priori knowledge and necessary truths.

Even if Chomsky is correct that people are born with an innate system of concepts which yield analytic connections, this fact will still leave him and Quine in agreement on some central points. Quine was criticising how philosophers had used analyticity as a philosophically neutral explanation of a priori knowledge and necessary connections. He wanted to show that the only type of analyticity he could make sense of, stimulus analyticity, could not explain the philosophically meatier conceptions of the a priori and necessity. Chomsky’s alternative conception of

analyticity, like Quine's, does not purport to explain our grasp of necessary truths, nor any supposed a priori knowledge. So both thinkers are in agreement on this point at least. However, as we have already seen, Quine was not only arguing against analyticity in the sense of an explication of the a priori and necessity. He was also arguing against the semantical notion of analyticity. So one clear area of disagreement between both thinkers centres on whether there are analytic connections in the sense of sentences whose truth is fixed purely in terms of meaning independent of extra-linguistic fact? Quine gave arguments against this in the first four sections of "Two Dogmas", and these are arguments which Chomsky never managed to answer. Furthermore, Chomsky has offered little evidence in support of his positive conception of analyticity. So it is clear that Chomsky has not really answered Quine's objections against the idea of an analytic/synthetic distinction.

Since Chomsky's conception of innate analytic connections in natural language and Quine's conception of stimulus analytic connections rest on such radically different pictures of how language is learned it would be foolish to say that they are in agreement on the nature of analyticity. However, it is significant that both thinkers are in general agreement as to the fact that analytic sentences do not explain a-priori knowledge, or necessary truths. So the debate between them on analytic sentences is an illusory one when it comes to metaphysics and epistemology. There is, however, a substantive debate between the two thinkers as to the nature of analytic sentences as a part of natural language and I will return to this point later in the essay.

SECTION 2: FACULTY PSYCHOLOGY AND ANALYTICITY

Chomsky has at various different times spoken about the language faculty, a science forming faculty, a folk physics faculty, and a moral faculty. To say that we have a faculty is to say that we have domain specific knowledge of something, and that this knowledge has not resulted from a general inductive procedure. A clear indication of a faculty operating in the brain is when as result of a disease or an accident, a person loses a certain ability (e.g. linguistic ability) but their general intelligence remains unimpaired. So, for example, people such as Marc Hauser have argued for the existence of a moral faculty based on the fact that selective brain damage results in people losing their moral sense but retaining their overall general intelligence. Likewise, Chomsky has discussed cases of people losing their linguistic abilities without it having any effect on their overall general intelligence.

Another key reason to believe that some abilities are the result of a faculty is that the ability arises at set times in the person's development. Furthermore the ability arises in people at more or less the same time, independently of their general intelligence. Again Hauser argues for the existence of a moral faculty based on the fact that children uniformly answer tests in a similar way when at the same age of their development.

In Chapter 2 I will introduce Chomsky's arguments for a science-forming faculty and for a language faculty. I will argue that Chomsky has not provided enough evidence to support his claim that humans have a science-forming faculty. However I argue that he has provided much more robust evidence to justify his claim that humans are born with a language faculty. In Chapter 3 I will provide what I believe is strong evidence that children come equipped to the world with intuitive folk physics and folk psychology. The fact that these two different domains of knowledge are not

commensurate indicates that they may belong to different faculties of the mind. And there is independent evidence that folk physics and folk psychology are faculties of the brain. A person's folk psychology does develop at set times in the child's development independently of general intelligence. So, for example by the age of four children can typically pass the false belief test with the notable exception of autistic children, most of whom fail the test at this age. The fact that autistic children fail this test is instructive because they typically show no deficit in general intelligence. Furthermore, children who do in fact have a lower IQ, such as children with Down syndrome, manage to pass the test by the age of four. This shows that we *may* be justified in claiming that folk psychology is a distinctive faculty of the human brain.

Overall I would claim that there is some reasonable evidence (though far from overwhelming) to support the postulation of a faculty for folk physics, folk psychology, morality, language, amongst others. When Chomsky claims that humans are born with innate concepts which will automatically yield analytic connections he fails to detail the nature of these concepts, and whether these concepts are unique to particular faculties. Consider the following purported analytic statements: (1) All ducks are animals. (2) If John persuaded Bill to go to school then Bill decided/intended to go to school and did so without being under duress. (3) Three times five is half of thirty. It is a fair bet that most people would consider these statements to be analytic, though it must be admitted that some people may answer 3 by saying it is a priori synthetic. Let us assume for a minute that most people would consider all three sentences analytic. The question is: where, for a Chomskian, does their analyticity derive?

The subject matter of 1 is biology, 2 folk psychology, 3 is mathematics. All three statements are, to use the traditional definition, true by virtue of the meaning of

the terms and not because of extra-linguistic facts. For Chomsky, as we have seen, the *meaning* which creates these analytic connections is something we are born with, not something which is true by convention. The three different statements contain concepts of agency, number, and biological entities. What we do not know is if those concepts grow in the brain separately from the different faculties, or whether they grow internal to the different faculties. So, for example, if a mathematical truth such as $2+2=4$ grows internal to a faculty of mathematics, then damage to this faculty would not have any significance for our ability to understand analytic truths which grew internal to other faculties. This is a question which Chomsky has not in general been very clear on. He correctly notes that more empirical research is required before we can decide on what the correct answer to the question is, and that it is empirical research that is necessary, not a priori philosophical speculation.

Whatever the exact structure of the various faculties are, Chomsky thinks that if there are analytic connections in different domains they will emerge internal to the faculties. Chomsky is explicit on this point in a personal correspondence with me:

Given my language faculty, it's a necessary truth that "book" rhymes with "took" and that "chase" entails "follow." Horwich is quite right. This has no bearing on, say, whether arithmetic truths are synthetic a priori... On the Hume questions, we have to distinguish what our innate sense of number determines, and what is the truth about numbers (if we regard them as entities that exist independently of our cognitive capacities, as Platonic entities of some sort). The two presumably coincide, but we can imagine an organism in which they would not (and maybe that's us, as a matter of fact). (Chomsky: personal communication)

One important point to note about this communication is that he here admits a kind of scepticism about the truths of mathematics. He claims that while it is possible that our innate number sense does correspond to mind-independent Platonic entities, we cannot say for certain that they do. Now, given that our best scientific theories about

the world are couched in mathematical formulae, this is a startling admission. Chomsky is claiming that we have no justification for acting in accord with our innate number sense, and that we have no way of knowing whether our innate number sense corresponds with any facts in the objective world.

Furthermore, Chomsky is arguing that different types of analytic connections will emerge internal to our various different faculties. So we will have analytic truths derived internal to our innate number sense, and analytic truths derived internal to our language faculty etc. This means that Chomsky's view on the nature of analyticity is parasitic on his views on the various different faculties of the mind. It also shows that truths which are typically grouped together by philosophers as analytic derive their analyticity internal to particular faculties. So Chomsky's justification for arguing that analytic truths exist depends on accepting his controversial views on (1) innate concepts, (2) faculty psychology, (3) potential radical scepticism about the truths of mathematics. Chomsky's conception of analyticity seems to ask more questions than it answers.

SECTION 3: CHOMSKY AND QUINE ON ANALYTICITY

In the previous sections we established that neither Chomsky nor Quine have a conception of analyticity which is rich enough to provide a priori justification and evidence of necessity. However, I also noted that, despite having this much in common, their views on the nature of analyticity diverge radically. In this section, I compare the two views and consider which view deals with the available data better.

Having already explained Quine's negative critique of analyticity at the beginning of this chapter, I will not repeat the material here. In this section I will focus on how Quine explains away people's intuitions of analyticity.

As is well known, Quine criticised the notion of analyticity as it was used by philosophers from Descartes to Ayer. Quine, of course, did not rest with such negative criticisms; he also explained why people held such intuitions of analyticity in the first place. Quine was well aware of experimental work which showed that people had an intuitive conception of analyticity. However, for Quine, such intuitions do not provide evidence that analytic connections actually exist in natural language. He notes that analytic intuitions typically set in where people have difficulty figuring out what the person who denies the truth of such a sentence may mean by its denial. This reaction is perfectly consistent with Quine's web of belief story. A person who denied something which is deeply embedded into our web of belief would indeed result in an interpreter having difficulty figuring out what the person meant by the denial.

We saw above that Putnam criticised Quine's blanket criticism of the notion of analyticity. Putnam correctly argued that the web of belief story does not account for trivial cases of analyticity. In *Word and Object* Quine told a story about analyticity which, he argued, was largely consistent with the story told by Putnam.

Quine noted that terms like "Indian nickel" and "Buffalo nickel" are not stimulus synonymous while "bachelor" and "unmarried men" are. Quine explains this disparity by noting how the different terms are learned. He argues that terms like "Indian nickel" and "Buffalo nickel" are learned by association with non-verbal stimuli. Terms like "bachelor" and "unmarried men" are learned by association with verbal stimuli. So, for example, one learns the meaning of "bachelor" by learning to associate it with "unmarried man". Obviously from a social point of view the stimulus meaning of the term "bachelor" will vary from person to person. A child can be told the meaning of the word "bachelor" by ostensive definition; however, obviously, the particular person who is being pointed to will not be representative of

the class of people in the world who are bachelors. So a child who learned the meaning of “bachelor” this way would be unable to generalise the term using this technique. Even if they pointed to a vast number of different bachelors and named them the child would not be able to generalise the term because there are no particular characteristics shared by bachelors alone other than being unmarried men. If an adult wants to teach a child what a bachelor is he can do no better than tell the child that bachelors are unmarried men. In the case of the terms “Indian nickel” and “Buffalo nickel” things are different; we learn these terms through ostension. If we learn that “Indian nickel” and “Buffalo nickel” pick out the same object, we again do this inductively. Quine argues that some technical terms of science do not have a socially fixed stimulus meaning to govern how they are used. In this sense, Quine argues that they are like terms such as “bachelors” and “unmarried men”. There is, however, one difference between the technical terms of science and terms like “bachelor”; “bachelor” gets its meaning by being connected to “unmarried man”, while in the case of technical terms their meaning is governed by their relation to a whole scientific theory.

Quine’s discussion of analyticity above is largely consistent with the story told by Putnam. Quine argues that a sentence like “All bachelors are unmarried men”, appears to be analytic because we learn the meaning of “bachelor” through associating it with “unmarried men”. It is not actually analytic in the sense of being immune to revision, nor does it provide us with any kind of a priori justification. The sentence merely has the appearance of being analytic because of its mode of being learned. As language evolves the sentence ‘All bachelors are unmarried men’ may no longer be considered analytic though in terms of present usage it certainly gives the appearance of being analytic. Likewise Quine’s web of belief story explains our grasp

of less trivial sentences which have the appearance of being analytic, though as in the 'bachelor' case, he again accepts that such sentences are subject to revision as our language and theories evolve.

Like Fodor, and unlike Putnam, Quine does not beg any questions because he is not trying to explain the existence of analytic truths by using one-criterion concepts. Quine is merely arguing that since our manner of learning some concepts will be through one-criterion concepts this explains why some people have analyticity intuitions. Ultimately, Quine argues that there is no sensible way of explicating an analytic/synthetic distinction. For this reason he argues that we have no reason to believe that such a distinction exists in language. He acknowledges that people have an intuition of analyticity but he explains away this intuition through his notion of one-criterion concepts.

Overall then Chomsky and Quine are agreed on certain points. Neither thinker argues that any sense can be made of the notion of analyticity as a defence of a priori justification, or of our supposed grasp of necessary truths. Both thinkers, however, do manage to offer at least some explanation of why people have an intuition that a distinction between analytic and synthetic truths exists. Chomsky's explanation of analyticity is sketchy and vague and has little evidential support. He claims that people are born with innate concepts which as a matter of fact yield analytic connections; however, he does not offer empirical evidence to support this claim. Furthermore, his claims about analyticity commit him to the view that analytic connections exist in different forms in different faculties. So his purported explanation of analyticity relies on prior commitments to faculty psychology, and while this commitment may turn out to be justified, it is very much an open question. More importantly, his views on the analyticity of mathematics yield a type of scepticism

which cannot be taken seriously. Quine's arguments against the notion of an analytic/synthetic distinction were not refuted by Chomsky. His explanation of why people have an intuition of analyticity relies on a conception of language acquisition which he has not shown to be correct. The question whether Quine's conception of people's intuitions of analyticity is correct or not is parasitic on whether his conception of language is correct or not. At this moment in time the evidence does not indicate whether Quine is correct on this point or not. What is clear is that Chomsky's positive arguments for the existence of an analytic/synthetic distinction have little evidence in support of them.

SUMMARY

In this chapter we have considered two areas of debate between Chomsky and Quine. The central point of the chapter was to clear up supposed areas of debate between both thinkers which result from mutual misreading. So, for example, it was demonstrated that both thinkers are methodological naturalists. The supposed debate between them on the status of naturalism was dissolved by demonstrating through textual analysis that they both actually endorse the same type of methodological naturalism.

The debate between them on the status of the analytic/synthetic distinction was also analysed. It was shown that both thinkers are in agreement that analyticity cannot be used as an explanation of any supposed grasp of a priori knowledge or grasp of metaphysical necessity. Because both thinkers are naturalists they obviously reject any appeal to a priori knowledge or to a grasp of metaphysical necessity. However, while both of them take an entirely naturalistic approach to analyticity, Chomsky, unlike Quine, argues that analytic truths in the sense of truth in virtue of meaning exist. This is the key difference between Chomsky and Quine on analyticity;

Quine denies that we can make sense of analyticity in the sense of truth in virtue of meaning.

It was demonstrated that the supposed debate between them on the nature of naturalism was merely a result of mutual misreading. It was shown that once read aright it is clear that both Chomsky and Quine are methodological naturalists. In the debate between them on the analytic/synthetic distinction it was demonstrated that again the debate between them was an entirely naturalistic one. I discussed their debate on whether semantic analyticity existed. It was shown that Chomsky's positive conception of analyticity has virtually no evidence to support it. Furthermore, Quine's efforts to explain away our intuitions of analyticity, while cogent have not been proven to be correct yet either.

CHAPTER 2: SUBSTANTIVE DISAGREEMENTS

Chapter 1 centred on clearing away pseudo-debates between Chomsky and Quine. It was demonstrated that both thinkers are naturalists, and while Chomsky believes that an analytic/synthetic distinction exists, his conception is an entirely naturalistic one. This chapter will consider more substantive debates between the two thinkers. The primary aim of this chapter is to consider how both thinkers deal with Underdetermination in various different areas. Understanding the nature of UD will occupy the rest of the thesis. Both thinkers take different approaches to the phenomenon of UD. Their different approaches result in them disagreeing with each other substantively on how language and concepts are acquired. This chapter will introduce how both thinkers deal with UD in a variety of different areas. In Part 1 I will deal with UD in the area of syntax acquisition. I will compare how Chomsky and Quine's different conceptions of how language is acquired deal with the supposed UD facing the child as they begin learning the syntax of their language. In Part 2 I will discuss the global UD of our total scientific theory of the world. I will compare the different ways that Chomsky and Quine deal with global UD. Chomsky postulates a Science Forming Faculty to overcome the problem whereas Quine argues that global UD is just a fact of life that we simply have to live with. In Part 3 I compare Chomsky's and Quine's views on the correct methodology to be used when studying language. Chomsky thinks that language should be understood in terms of competence, while Quine focuses almost entirely on performance issues. The primary aim of this chapter is to get clear on how both thinkers deal with UD, and what their philosophy of language is. The rest of the thesis will compare both of their views on language acquisition and with the methodology both thinkers use to solve the difficulties they encounter.

PART 1: UNDERDETERMINATION AND LANGUAGE ACQUISITION

We saw above that part of the disagreement between Chomsky and Quine was a result of mutual misreading. So, for example, Chomsky misunderstood Quine's views because he wrongly attributed to Quine a narrow empiricist model of the mind. In this section I want to key in on a very real difference between the two thinkers, namely, their views on language acquisition. Their disagreement results from Chomsky's emphasis on the poverty of stimulus argument. Quine's model ignores the supposed poverty of stimulus and emphasises trial and error learning from the outset. The primary purpose of this section is to introduce a debate between both thinkers which will be explored in detail throughout the rest of the thesis. So, for example, in Chapters 3 and 4 I will be dealing with UD in both syntax acquisition, and concept acquisition.

In Chapter 1, I mentioned how Chomsky situated himself in the rationalist tradition, how he understood rationalism as a position which emphasised the innate mechanisms of the mind which went beyond abilities such as incremental data processing, habit formation and induction. I further emphasised that, for Chomsky, rationalism was not just an abstract philosophical position; it was rather, an empirical claim about the structure of the mind. For Chomsky, one important reason to make explicit his philosophical position was so that it could be subject to empirical tests.

In the years preceding the ground-breaking publication of Chomsky's *Syntactic Structures*, Behaviourism was the dominant position in psychology. The central figure of the then dominant Behaviourism was the psychologist B.F. Skinner. Skinner had claimed that he could explain all aspects of human psychology in terms

of conditioning. In his book *Verbal Behaviour*, Skinner claimed to have explained human linguistic abilities in terms of conditioning. Chomsky's "Review of Skinner's *Verbal Behaviour*" criticised Skinner on the grounds that his theory of conditioning could not explain certain aspects of language, such as recursion, compositionality etc. Chomsky claimed that Skinner was in principle incapable of explaining language acquisition because he was implicitly working with an empiricist theory of the mind.

As I showed in Chapter 1, Chomsky believed empiricism and rationalism were models one could adopt to explain human psychological abilities, and that these models were to be tested empirically. According to Chomsky, Skinner's empiricist model failed to explain language acquisition because it could not explain certain features of language, such as our ability to understand and speak a potentially infinite number of sentences. So Chomsky reasonably conjectured that since the empiricist model was not sufficient to the task, a more sophisticated model would be needed.

Chomsky's central task in explaining language acquisition is similar to the 'naturalised epistemologist' in that he has to explain how the child goes from meagre input (language exposure, conditioning), to torrential output (potentially infinite amount of sentences). Chomsky's reasoning is simple; whatever data is not provided by the environment must be brought to the occasion by the child himself. His argument was partly logical and it is called Argument from the Poverty of Stimulus (APS). Essentially he claimed: (1) The actual data that a person is exposed to is fragmentary and limited; (2) children learn to mouth complex constructions which go beyond the data they have been exposed to; (3) what cannot be explained in terms of experience must be explained in terms of innate constitution (2000a, 7)

Assuming that one accepts the validity of the above argument, the next question to ask is whether premises (1) and (2) are in fact correct? This is of course no

longer a conceptual question but is rather an empirical one. In order to test whether the two premises are true we simply need to discover grammatical rules that ordinary English speakers all accept (at least implicitly) and then see whether children are exposed to these rules. If people are not exposed to enough data to discover these rules, then the APS is not just valid but also sound.

Consider the following sentence⁸ (A) Mary is at home. In the interrogative form (A) is as follows (B) Is Mary at home? Now consider sentence (C); Bill can sing. We can turn (C) into an interrogative as follows: (D) Can Bill sing? If a person wanted to explain what was going on with A-D they might propose the following rule. *Rule 1*: To turn a statement into a question find the auxiliary and move it to the front of the sentence. However this does not work in certain other situations. Consider (E) Mary *will* believe that Frank *is* here. Here there are two auxiliaries. So a child following the above rule could derive (F) *Is* Mary *will* believe that Frank here. To an ordinary English speaker (F) is obviously ungrammatical. Or the child could try (G) *Will* Mary believe that Frank *is* here. The previous sentence is obviously a grammatical construction. Now a behaviourist would claim that using ordinary inductive intelligence the child could arrive at *Rule 1*, and then produce sentences (F) and (G) when trying to turn (E) into a question. Furthermore, the behaviourist would argue that through positive and negative reinforcement the child would learn that sentence (G) is correct. And from this fact derive *Rule 2*: When turning a sentence into an interrogative move the first auxiliary to the front of the sentence. Of course, even *Rule 2* will not work. Consider (H) The man who *is* tall *will* leave now. If I followed *Rule 2* I would arrive at (I) *Is* the man who tall *will* leave now; which again is ungrammatical. Now presumably a conditioned response theorist would claim that

⁸ Here I am following Howard Lasnik's Syntactic Structures Revisited.

when the child tries to use *Rule 2* for sentence (H), he would receive negative reinforcement and would try to find a more general rule. The correct answer is *Rule 3*: Move the auxiliary beside the main clause to the front of the sentence⁹. Now in sentence (E) the sentence is about what Mary believes, so the auxiliary beside Mary will get moved to the front.

According to Chomsky, the important point to note from above is that while it is possible to tell a vague story about the child arriving at *Rule 3* through conditioning, it is just that: a story. Some empirical research e.g., Crain and Nakayama 1987 indicates that children virtually never offer the deviant sentences above, nor hear them spoken. They automatically construct sentences according to *Rule 3* (See also Crain Nakayama 1987 and Legate Yang 2002).¹⁰ There is also other empirical research (Brown and Hanlon 1970), which indicates that correction plays a very small role in language acquisition. So the poverty of stimulus argument offers a plausible account of language acquisition, and it makes plausible the claim that the data we are exposed to are not enough to account for the language we speak. So it claims that we need to postulate some innate apparatus in the brain (more importantly, domain-specific innate apparatus) which will determine how we speak and understand language.

Believing that he has shown that it is necessary to postulate an innate language faculty, Chomsky then sets out to demonstrate its actual construction. He proceeds as follows: he abstracts from the concrete situation and works from an idealised model, ignoring certain limitations which are inherent in humans, and sets out to study what the nature of the language faculty must be in order to account for the empirical facts. An important point to note is that Chomsky is concerned with linguistic competence,

⁹ I will discuss this particular Poverty of Stimulus argument in greater detail in Chapter 4 and 5.

¹⁰ Research into the child's linguistic data casts serious doubts on Chomsky's claims about the nature of the linguistic data which the child is exposed to. I will discuss this in greater detail in Chapter 4 and 5.

rather than performance. He claims that any theory of performance which is studied should be worked out after we have sketched a theory of competence¹¹. Chomsky defines the study of language as follows:

Linguistic theory is concerned with an ideal speaker-listener, in a completely homogeneous speech community, who knows its language perfectly and is unaffected by memory limitations, distractions, etc. (1980a, 25-26).

The APS purports to show that our language faculty is wired into the subject, but it leaves an obvious question: why, if all humans are genetically programmed at birth to speak a language, do people from different countries speak different languages? Chomsky explains this fact through his principles and parameters approach. He claims that language is governed by certain abstract principles, and these principles can be realised within certain narrow parameters. The scope of the parameters determines the degree of variation which we see in the various world languages.

When explaining how a child learns his first language, Quine does not consider the Chomskian APS. According to Quine, when a human is learning a language, he is learning a theory of the world. Hence he draws a parallel between a child learning a language and a scientist learning a theory of the world. Quine considers language as a social intersubjective art which involves mouthing words in intersubjective scenarios and having our words reinforced either positively or negatively. This description of language learning as being based on conditioning immediately commits Quine to a conception of language learning which relies on general intelligence. A child born in France, Spain or England will hear sounds spoken to him in various different circumstances and will mimic them. Furthermore, the child will babble words in various different circumstances, and will be positively and negatively reinforced. As a

¹¹ Chomsky holds out virtually no hope that a theory of performance can be sketched; however he claims that those who try to develop such a theory, should at a minimum already have an accurate theory of competence.

result of this conditioning, the child will learn the right situations to use various observation sentences in. And the child will eventually learn to break down observation sentences into words, and learn how to insert these words into other sentences by means such as analogical substitution. In this sense, each child is using his general intelligence to figure out what language his elders are speaking. Quine has stressed time and time again that the radical translator is in a similar situation to the child learning his first language¹². This, of course, emphasizes that the child is like a scientist in trying to construct a theory of the data which surrounds him, using his general intelligence.

On the Quinean picture, various different cultures speak different languages which have different structures, and the child born within a particular culture must *learn* the language of that culture. For Chomsky, on the contrary, all humans speak the same language, and they just sound different because of parametric variations between the languages. In their introduction to Chomsky's book *On Nature and Language*, Adriana Belletti and Luigi Rizza describe the nature of a parameter in the following manner:

The child interprets the incoming linguistic data through the analytic devices provided by Universal Grammar, and fixes the parameters of the system on the basis of the analysed data, his linguistic experience. Acquiring a language thus means selecting, among the options generated by the mind, those which match experience, and discarding other options (2002, 17).

A simplified example of a parameter is what is known as 'The movement parameter'. Consider question formation. When forming questions, human languages have two different options to choose from: the first option is to move the interrogative phrase (who etc.) to the front, to a position in the left periphery of the clause; the second is to leave the interrogative phrase in the clause-internal argument position in

¹² A similar situation but not an identical one. Some differences between the child learning his first language and radical translation will be discussed in Chapter 4.

which it is interpreted (ibid., 17). English takes the first route: ‘Whom did you meet?’, Chinese the second ‘You love whom?’; while French uses both rules at the same time. The point is that the language faculty is a machine in the brain which implements abstract principles when exposed to certain data. The parametric variations which the principles allow mean that people exposed to different sets of data will develop what seem to be different languages. The parametric variations will obviously affect the structure of the sentences; however, they leave the propositional content of the sentences invariant.

According to Chomsky, people speak the language they speak because the data they are exposed to switches certain parametrical variations on underlying principles in their language faculty. They do not *learn* French; rather, they are programmed to interpret the data of experience according to innate principles subject to parametric variation. For Quine, on the contrary, we *learn* French or English by keying our sounds and the sounds of our community to states of affairs in the world, and by deducing how observation sentences are constructed. In other words, Quine thinks that we learn a language using domain-general procedures such as induction, analogy and reinforcement, whereas Chomsky thinks that we have a distinct language faculty encoded in our brains which automatically constructs sentences when exposed to certain data.

In describing language in terms of learning, Quine is directly contradicting Chomsky’s views. This can be seen by considering the subject-auxiliary inversion that I discussed above. Consider sentences such as (E) Mary *will* believe that Frank *is* here (G) *Will* Mary believe that Frank *is* here? (F)* *Is* Mary *will* believe that Frank here. Now, as we illustrated earlier, Chomsky contends that positive or negative reinforcement does not play any role here because children virtually never mouth the

deviant sentences. So, by emphasizing reinforcement as a key factor in language learning, Quine's conception of language acquisition is moving in the opposite direction to Chomsky's.

Quine argues that a necessary condition of any language learning is that we have innate similarity responses which guarantee that we see certain things as more similar to some things than to others. He claims, for example, that a particular quality space could determine that we would see a red ball as more similar to a white ball than to a red table. He correctly notes that the exact dimensions of the said quality space are to be determined experimentally. Quine is quite clear on the point that an innate similarity quality space is a *necessary* condition of any language learning because without it no conditioning would be possible¹³. In *Word and Object* Quine mentions two modes of learning sentences: (1) association of sentences with non-verbal stimuli, and (2) analogical synthesis which is the association of sentences with sentences. Quine offers one example of analogical synthesis, which is called *analogical substitution*, which he explains as follows:

Having been directly conditioned to the appropriate use of 'foot' (or 'This is my foot') as a sentence, and 'hand' likewise, and 'My foot hurts' as a whole, the child might conceivably offer 'My hand hurts' on an appropriate occasion, though unaided by previous experience with that actual sentence (Quine: 1960, 9)

Now obviously analogical substitution is only one of many different types of analogical synthesis, and Quine emphasises this point by claiming:

Further inter verbal associations are required which provide for the use of the new sentences without tying them, even derivatively, to any fixed range of non-verbal stimuli (*ibid.*, 10).

Quine's point is that we do not yet know what all of the different types of analogical synthesis are, and that we need to do experimental research to discover them. He has

¹³ Note here that Quine cannot conceive of language acquisition being possible without conditioning

continually emphasised that he sees nothing in principle uncongenial to innate readiness to language learning. Quine's openness to any innate apparatus may suggest that he and Chomsky could arrive at the same innate apparatus to explain language learning though approaching it from different angles. However, since Quine begins his research by assuming that language is *learned* he is obviously not prepared to postulate the same amount of innate mechanisms as Chomsky is.

According to Quine, at the beginning of language learning, a child begins to engage in typical babbling; speaking various words that he has heard from his peers such as 'mama' 'dada' 'car' etc. If the word is keyed to external circumstances that the parent approves of, the child will repeat the words in those circumstances because of positive reinforcement. Disapproving reactions from parents will ensure that the child will learn when it is not appropriate to mouth certain words. Observation sentences are fixed words such as 'mama' which are keyed to fixed stimuli in the world, to count as an observation sentence all the members of the community will agree on the correct circumstances for using the word. According to Quine, in the beginning, the child will use observation sentences as unstructured wholes to indicate whatever situation the sentence is keyed to. Other examples of observation sentences are 'It's cold', 'Rabbit', etc. These observation sentences are occasion sentences true on some occasions and false on others. In order to show how we can go from our meagre inputs to torrential outputs, Quine discusses how a child manages to do so as a kind of paradigm.

Firstly, as I have already said, Quine emphasises that a necessary condition of the child learning any language is that he has a pre-linguistic quality space (ibid., 83). This quality space, the exact nature of which is to be determined experimentally, will ensure that the child can be conditioned. Quine further adds that 'Unquestionably

much additional innate structure is needed to account for language learning' (1966, 57). He does not go into the details of precisely what innate apparatus is needed; he merely notes that this is an empirical question. With this proviso in place, he conjectures how a child learns to speak about the world. A child that we are trying to teach language must master observation sentences. When a child has mastered an observation sentence, and has, for example, mastered 'Cat', he knows the situations when he is justified in using the term. When the child learns to use the term, he also learns when not to use the term and hence has internalised a bit of logic; in this case negation. Furthermore, when the child learns to combine the term 'cat' with another term 'white' as in 'white cat', he has internalised the use of conjunction, Quine calls such terms compound observation sentences. A further step is observation categoricals such as 'When it snows, it rains'. Observation categoricals are very useful and can be passed from generation to generation, but according to Quine, they do not commit us to talking of objects. However, a child who has mastered the use of logical connectives such as these has mastered recursion and hence has mastered one of the key tricks of language. In order, however, for a child to be credited with an ontology he needs to have mastered what Quine calls observation predication, an example of which is 'This dog is black'. With observation predication, unlike observation conjunction, one is positing an object and predicating something of the object. In the above case, one is predicating blackness to the dog, whereas in conjunction all one needs is to say that white and cat are occupying the same space.

Once a child is indoctrinated into language in this sense, it is possible for him to become a member of the scientific community. So, for example, by using observation predication, the child begins to have an ontology. When using devices like conjunctions, observation predications and observation categoricals the child can

construct quasi-scientific theories. For example, a child with these tools could say: 'I am cold, and, when it is cold it rains. It is cold, therefore it is raining.' Here the child can make claims about the world and test them to see if they in fact obtain. So the child learning a language is learning a theory which he is constantly adjusting to cope with the contingencies of experience. In this sense, Quine argues that science and philosophy are refined versions of the child's linguistic theory which we are more explicit and precise about.

What this section brings out is a serious area of disagreement between Chomsky and Quine concerning language acquisition. Quine's emphasis on trial and error learning and general intelligence runs entirely counter to Chomsky's views on language learning which focuses on the APS, which involves claims regarding the lack of correction which children supposedly receive by their parents and peers, as well as the data which are available to the child learning various constructions. If true, these claims would demonstrate that Quine's picture of language learning is false. Later chapters of this thesis are dedicated to evaluating whether Chomsky's APS is empirically justified or not.

PART 2: UNDERDETERMINATION AND SCIENCE

In the previous section we considered an area of disagreement between Chomsky and Quine. The disagreement centred on whether Chomsky's poverty of stimulus argument had refuted Quine's conception of language acquisition. For Chomsky the poverty of stimulus argument is a form of UD facing the child as they try to learn their first language. He overcomes this UD by postulating innate syntactic rules. On Chomsky's view, UD pervades all of our scientific theories. In all cases Chomsky claims that the UD is overcome by innate constraints.

Chomsky's attitude to UD can be seen in various different places. When discussing Kripke's rule-following problem, he argues that the various different ways of applying concepts are ruled out by innate constraints. He further argues that we simply automatically follow the plus function as opposed to the quus function. The reason we are following plus instead of quus is, he claims, because of innate rules which we are following.

Furthermore, when considering Goodman's problem of induction, Chomsky again postulates innateness to overcome what he sees as a problem of UD:

And I remember back in 1949 or 1950-approximately then-discussing with him (because I didn't understand it) how it would possibly work unless you assumed innate principles. So take a pigeon- or a person, for that matter. It or she sees that emeralds are green, and that gets entrenched, and becomes a projectable predicate, and so forth. But the same person sees that emeralds are grue; so why is that not being entrenched? The only possible answer that I can think of is that green is somehow a part of the nature of the pigeon and the person, and grue isn't. (2012, 89)

I do not want to discuss the validity of Chomsky's solution to either Goodman's or Kripke's problems here as they are beyond the remit of this thesis. I have introduced Kripke and Goodman to illustrate that when confronted with what he perceives as UD, Chomsky tries to overcome it by postulating innate constraints.

Quine qua externalised empiricist has no difficulty in postulating innate apparatus if it is necessary to explain the behaviour of the organism under study. Thus, when dealing with Goodman's problem of induction, Quine appeals to innate apparatus as a solution to the problem. In his paper 'Natural Kinds' Quine argues that Goodman's problem of induction is overcome through our innate similarity quality space. Thus people are more likely to project 'green' than 'grue' because our innate similarity quality space will make the former projection more natural than the latter one. Quine speculates that our innate similarity quality space will correspond in some

sense to the structure of nature because natural selection will have culled those of our species whose innate similarity quality space radically deviates from the categories of nature (1969, 114-138).

I am not concerned whether Quine's answer to Goodman is adequate or not because this is not the topic of the thesis. The key point to note is that Quine, like Chomsky, has no problem postulating innate apparatus to overcome this particular problem. Quine notes that people do in fact project 'green' instead of 'grue' so he argues that this fact needs to be explained. The possibility of the 'grue' projection was not considered by anyone until the 1950's yet, as Goodman showed, it is a possible projection. So for Quine the best way to explain why this projection is never made is to appeal to an innate similarity quality space which makes certain projections more natural than other ones.

Quine's solution to Goodman's problem of induction demonstrates that he does not reject the logic of Chomsky's poverty of stimulus arguments. If we have behavioural evidence that children exhibit certain regularities, and that these regularities are not learned through trial and error and reinforcement; if it can be shown that the regularities are one amongst many other possible behavioural regularities we are, in Quine's view, justified in postulating innate apparatus to explain this fact. However, the key point is that postulating innate apparatus is only justified if the proposed competency cannot be learned inductively, or is not learned through trial and error, or based on analogy with things people previously know. Goodman's problem of induction fits this bill perfectly. Obviously, induction is not learned inductively, nor is trial and error used as can be seen from the fact that the grue example was not even considered until the 1950's. Since grue is a logical possibility and it was not ever used by anyone before 1950, we can be guaranteed that

people pre-1950 had not learned by trial and error to use the green projection over the grue one. Likewise there is no evidence that people distinguish between possible projections using analogical reasoning. So, from the Quinean perspective, postulating innate constraints to overcome Goodman's problem is fully justified.

It is notable, however, that Quine does not postulate innate constraints to overcome every perceived problem of Underdetermination. A key example of Quine not postulating innate constraints to overcome Underdetermination is in the area of Indeterminacy of Translation. Influenced by Chomsky, most theorists have viewed the IDT as a form of UD. They have argued that this UD can be overcome by postulating innate constraints. Quine famously rejects the equating of IDT with UD on the grounds that even if we work within a fixed though UD total theory of the world problems of IDT still remain. Nonetheless the IDT has a similar logical structure to Chomsky's APS, Goodman's problem of induction, and Kripke's rule-following problem. Yet in the case of the IDT Quine does not insist on innate constraints to overcome the problem; however, with Goodman's problem he has no difficulty in postulating innate apparatus. Quine's seemingly inconsistent ways of dealing with these similar problems needs to be addressed.

I do not want to delve into Quine's IDT too much at this point because I will be discussing it in detail in Chapter 3. In this section I merely want to explicate Quine's attitude to problems of UD. The key difference between the IDT and Goodman's problem of induction is that in the latter case we have evidence to decide between the different projections whereas in the former case we do not. So, for example, if anyone projects grue instead of green, at a certain point in history he will be proven wrong. Hence we have possible behavioural evidence which states that if a person projected grue they would be proven wrong at a certain point in history. So if

nobody projects *grue*, which we can tell is in fact the case based on behavioural evidence, then this fact needs to be explained. Why, if the evidence was consistent with *grue* or *green*, did everyone project *green*? The obvious answer is that the *green* projection is more natural because of our innate predisposition. In the case of IDT things are different. Quine notes that there is no behavioural evidence which can decide between ‘undetached rabbit part’, ‘particular instance of universal rabbithood’ etc. Since there is no behavioural evidence which indicates which translation of ‘*gavagai*’ is the correct one, then to postulate an innate preference for a translation of ‘*rabbit*’ is arbitrary and question begging. I will deal with this point in more detail in Chapter 3. For now, I just wanted to note that while, like Chomsky, Quine is prepared to postulate innate apparatus to overcome supposed problems of UD, he will only do so if the postulation is justified by behavioural evidence.

Another area where Quine does not postulate innate apparatus to overcome problems of UD is in the area of what I will call global UD. Global UD is the underdetermination of our total theory of the world by the data of experience. In the case of global underdetermination Quine speculates that it is possible for a scientist who uses all possible data to be faced with a rival scientist who uses the same data but has a logically incompatible theory. In this case, we will have empirically equivalent systems of the world which are logically incompatible. However, Quine does not respond to this UD by postulating innate apparatus which determines that we accept one theory, while our rival theorists accept a different theory based on their innate apparatus. Chomsky approaches the problem of underdetermination of our total scientific theory from a different vantage point. He assumes from the outset global underdetermination is a fact, and that we nonetheless manage to form accurate scientific theories despite this global underdetermination. To explain this fact he

postulates an innate science-forming faculty. I will now describe and evaluate Chomsky's conception of the science-forming faculty, before discussing in detail how Quine deals with global UD. These evaluations will further reveal differences in how Chomsky and Quine deal with problems of UD.

For Chomsky our language faculty grows in much the same way our arms grow or in the way we reach puberty; no sense, according to him can be given to the metaphor of language learning any more than we can make sense of the metaphor of learning to grow arms (2000a, 7). Our language faculty grows as a result of our innate genetic program and, as such, inclines us to use our language according to these rules¹⁴. From an epistemological point of view, he argues that we are also reliant on a faculty; he calls it our Science-Forming Faculty. It is worth pausing to consider this faculty and its nature. The Science-Forming Faculty¹⁵ is connected with the fact that our ordinary theories of the world are underdetermined. If one were to take any theory of the world it would be supported by a certain amount of data. Chomsky's point is that the data does not just imply one theory; rather there are countless theories which are compatible with the data but which are incompatible with each other. Such underdetermination has the air of a truism about it. Chomsky claims that humans are innately constrained in the types of theories they find plausible and therefore accept. He puts it as follows:

Because we share across the species a kind of science-forming faculty that limits us, if you will, but at the same time provides the possibilities of creating explanatory theories that extend far beyond the evidence that is available...it is worth paying attention to what the scientist is doing when a new theory is created...First of all the scientist has very limited evidence. The theory goes far beyond the evidence. Secondly, much of the evidence that is available is disregarded; that is in the hope that someone else will take care of it. At every

¹⁴ I say our language faculty inclines us to use the rules of language in certain ways: it does not determine us to use it in the same ways. This is because Chomsky thinks that human behaviour is neither genetically determined nor controlled by external stimuli. I will discuss this point in greater detail in Chapter 3.

¹⁵ Hence forth Science Forming Faculty will be referred to as SFF

stage in the history of science even normal science, there is a high degree of idealisation, selection of evidence even distortion of evidence (Magee 1978 p.187).

In other words the SFF constrains the types of theories which humans can formulate. At this point, an obvious question arises: Is Chomsky correct about us having a SFF? Does this faculty help us construct theories which in fact correspond with how things are in the world independent of us?

Chomsky's answer to the above question is as follows:

The successful natural sciences, then, fall within the intersection of the scope of the SFF and the nature of the world. They treat the (scattered and limited) aspects of the world that we can grasp and comprehend by naturalistic inquiry, in principle. The intersection is a chance product of human nature (200b, 82).

So Chomsky is here claiming that it is just a chance product of how our minds are constituted that we can form accurate theories about some aspects of the objective world. While our SFF may help us construct true theories about some aspects of the world, its very scope and limits determine that there are some areas of the world that we cannot comprehend.

Chomsky holds that the SFF is deeply connected to the Problems/Mysteries distinction. He takes the fact that humans are epistemically bounded to be a truism that any naturalistic philosopher will accept, '...We are after all, biological organisms, not angels...' (ibid., 83-84). Creatures created by evolution are bound to be contingent and fallible, having particular cognitive apparatus which works brilliantly in some domains but totally fails in other areas. Rats are a perfect example: there are some mazes which they simply cannot be taught to solve. Chomsky notes this point claiming: 'There is no more reason to suppose humans to be capable of solving every problem they can formulate than to expect rats to be able to solve any maze' (1986b,

231). In another discussion of the SFF he claims, ‘The scope and limits are relative to humans; rats and Martians have different problems and mysteries’ (2000b, 83-84).

Chomsky claims that certain questions such as freewill, why a person makes one choice rather than another, are a mystery, whereas he argues other questions are mere problems, such as the questions of physics and questions about the syntax of natural language. He furthermore claims that we do not have to consider such questions as meaningless; rather we can consider them legitimate questions that the structure of our brains will not allow us to answer. In his Basque lectures he commented:

You can give a naturalist interpretation of such matters, and maybe there is a right question and we just cannot formulate it, because we are just not built that way. So if there is one we may not find it (2009, 47).

Chomsky’s discussion of our SFF commits him to a view of reality which is in some respects Kantian. Both thinkers believe that how we view the world is governed by our innate faculties of the mind. For both, reality conforms to our mode of cognition¹⁶. Furthermore, like Kant, Chomsky believes that there is a ‘thing in itself’ which extends beyond what we *can* know. Clarification of Chomsky’s position is obviously needed. He believes that whatever we know about the world of physics is known because of the innate structure of our minds. He further believes that there are some facts about the world that we cannot know because of the way our minds are structured. The limit on possible hypotheses which we can make is, Chomsky admits, similar to what Peirce called the principle of abduction. However, unlike Peirce, he does not think that there is anything about the theory of evolution which will

¹⁶ There are obvious disanalogies with both thinkers. Most importantly Kant thinks that we can only know the world insofar as it conforms to our mode of cognition. He further thinks that we can never know the real world whereas Chomsky thinks that our mode of cognition means that we can only partially grasp certain aspects of the world while others will forever elude our grasp.

guarantee that our SFF will allow us to answer all the questions we want to answer (2000b, 83).

Chomsky is committed to the view that there is a determinate world which exists independent of our cognition. He also believes that what we know can be known only insofar as it conforms to our mode of cognition, so the *real* world cannot be exhaustively known by humans. He captured this point in his Fara lectures when he used the metaphor of a scientist being like a person looking for his keys at night centring his search around a lamp post; the person is looking for his keys there not because he has lost them beside the lamp, but rather because this is the only place which is lit up. For Chomsky, the scientist achieves success by limiting his researches to the small areas where he can construct intelligible theories. Furthermore, he counsels that we should treat these scientific theories as being more real than the ordinary world of empirical experience.

Ultimately, Chomsky can be construed as believing that humans are born with an innate SFF, which limits and makes possible the type of scientific theories which we can accept. These theories accurately capture certain aspects of the real world, but are ill equipped to deal with certain problems; hence part of the world will permanently remain hidden from us in darkness. Chomsky's view on science can be captured with an image which Locke employed in his *Essay Concerning Human Understanding*. This picture compared a scientist to a person who is positioned within a closet with some light shining in the doors towards him. While the light gives him some understanding of the real world, it is only a fragmentary limited grasp of the

world¹⁷. In other words, the world is partially opaque to us because of the structure of our minds.

THE EVIDENCE FOR THE FACULTY

We saw in the last section that Chomsky compares the faculty of language with the SFF; he does, however, note a distinction between them. He spells out this difference as follows:

In some domains-acquisition of language, object perception, etc.-the growth of knowledge just happens to us, in effect. The mental faculty grows from its initial to its steady state without choice, though not necessarily without effort of willed action. In other domains- the natural sciences, for example- the growth of knowledge involves deliberate inquiry involving hypothesis formation and conformation; guided no doubt by 'abductive' constraints on potential hypotheses as well as other equally obscure factors that enter into choice of idealization and the like. The basic elements of rational inquiry may have some of the properties of such cognitive systems as the language faculty, though the ways in which they are employed is surely quite different: Scientific Knowledge does not grow in the mind of someone placed in the appropriate environment (1980a, 140).

This claim of Chomsky's obviously needs clarification. He is making two key points: (1) the SFF does not just grow when a human is placed in an appropriate environment, and (2) the SFF uses a 'generate and test' model to construct its theories as well as certain abductive principles which Chomsky does not spell out. Since a generate and test model is easily captured by an empiricist model of the mind, he must think that it is the abductive aspect of the SFF which distinguishes it from other theories of how we develop our scientific abilities. This leads to the question of what the nature of this abductive principle is.

Chomsky has rarely explicitly discussed the exact nature of this abductive principle. However, when discussing the problems and mysteries distinction, he

¹⁷ While this Lockean image is useful to illustrate Chomsky's picture of the mind's relation to the world it would be a mistake to otherwise equate Chomsky and Locke's theory of knowledge.

vaguely points towards the type of considerations which our abductive principle forces on us. Here are his reflections on the matter:

Equipped with SFF, people confront ‘problem situations’, consisting of certain cognitive states (of belief, understanding, or misunderstanding), questions are posed, and so on. Often SFF yields only a blank stare. Sometimes it provides ideas about how the questions might be answered or reformulated, or the cognitive state modified, ideas that can then be evaluated in ways that SFF offers (empirical test, consistency with other parts of science, criteria of intelligibility and elegance, etc.) (2000b, 83)

What is central here is the manner he claims the SFF evaluates questions. According to Chomsky, the SFF uses empirical tests, consistency with other parts of science, and a criterion of intelligibility and elegance. He adds an etc. at the end of the paragraph which implies that the SFF uses other criteria; however, we have no way of knowing what they are until he explicates them. Empirical testing is presumably what Chomsky refers to as trial and error, so this is something which can be accounted for by general intelligence. There is no reason to believe that consistency with other parts of science cannot be achieved by an inductive reasoner who has internalised a bit of logic. Our criterion of intelligibility and elegance is something which influences our thinking in aesthetics, as well as our reasoning in politics, etc. There is no reason to believe that our criterion of intelligibility and elegance is limited to a particular faculty of the mind.

The evidence that Chomsky has offered in favour of the SFF is so far extremely limited. He has pointed out that scientific theories are underdetermined by experience. This UD is basically an APS. However, it is meaningless unless he can provide an account of how much data is enough data for a scientist to *learn* from. Until such time that Chomsky provides such a theory, for him to say that such and such scientific theory cannot be learned from experience is meaningless.

Furthermore, Chomsky has never been able to provide a detailed theory of what the nature of the SFF is, and the few details he has provided do not distinguish his views from those of an empiricist. So in this sense he has not provided enough evidence to justify his postulation of a SFF. However there is a different strand of argumentation which Chomsky sometimes uses to support his claim that people have a SFF. He points to the timing of new scientific discoveries and their reception by the public. Obviously people do not develop scientific theories at a fixed age. Chomsky clearly knows this; he is not claiming something absurd like children automatically acquire Newtonian Physics by the age of 5, and Einsteinian Physics by the age of 7. However, he does make claims about the fact that people hit on new scientific theories at the same time, and once these scientific theories are discovered people accept them as obviously true. In this context he is talking about historical time not developmental time. Chomsky claims that Peirce first recognised these points and in *Of Minds and Language* he cites approvingly Peirce's conception of the matter:

Peirce's sense was very straightforward and, I think, basically correct. He says you want to account for the fact that science does develop, and that people do hit upon theories which sort of seem to be true. He was also struck by the fact, and this is correct, that at a certain stage of science, a certain stage of understanding, everybody tends to come to the same theory, and if one person happens to come to it first, everybody says 'Yes that's right' (2009, 3-4)

Here Chomsky is picking out what seems to be a real phenomenon. It does appear that for at least some scientific discoveries two great thinkers may hit upon the same idea at the same time, even though a generation beforehand nobody would have thought of the idea. Some examples from intellectual history are: Leibniz and Newton co-discovering infinitesimal calculus, Darwin and Wallace co-discovering the theory of evolution, and Frege and Russell co-discovering predicate calculus. However, it is hard to see how such facts can be used as evidence for a SFF. A more natural interpretation is that people like Darwin and Wallace hit upon the same discovery

because they were hard working scientists who were able to build on the work of their predecessors. However, the question of how many new discoveries are made possible by prior theories is a question for biographers of scientists and historians of science to answer. Chomsky will need to engage with such thinkers if his speculations are to be given any substance.

Chomsky not only appeals to the fact that great scientific thinkers hit upon the same ideas at the same time to support his theory, he also claims that when such discoveries are made people recognise them as obviously true. Here he is presumably claiming that people find such discoveries obviously true because they interpret them using their own SFF. Now, while I do not want to engage in amateur intellectual history, the historical record strongly suggests that Chomsky is mistaken on this point. Darwin's theory of evolution was not greeted by the scientific community as being obviously true. In fact the theory sparked massive debates and controversies. Newton's infinitesimal calculus was severely criticised by George Berkeley amongst others. In fact most of the great scientific discoveries have been greeted with severe resistance; they have not been just accepted as obviously true upon discovery. This leads us to the question of how Chomsky's speculations on the SFF can handle the people who did not find Darwin's theory of evolution obviously true. Did they have a malfunctioning SFF? Did they have no SFF? Or was the faculty overruled by a stronger faculty in the mind? Chomsky does not even consider such questions. It is obvious that if he wants to construct an accurate theory of a SFF he will need to mesh this theory with actual facts from the history of science and sociological studies of people's reactions to new scientific discoveries. As in the area of language production, when it comes to scientific discoveries and their reception, Chomsky will need to sketch an accurate description of performance in order to justify any claims he

makes about competence. In the absence of such a description Chomsky's claims remain nothing but idle speculation.

Questions about the history of science raise another interesting point. Most historians believe that science was developed in the seventeenth century. When Chomsky discusses the SFF he usually does it in relation to modern science. John Collins in his paper "On the Very Idea of a Science Forming Faculty" claimed that we could view Chomsky's SFF in two different ways. The first way he calls the narrow interpretation; on this interpretation the faculty is responsible for our grasp of science in the modern sense. The broad interpretation claims that our faculty is responsible for our ability to construct rational theories to account for our experiences, so this faculty would account for far more than just modern science.

The narrow interpretation is the one which Chomsky seems to intend us to use. All of the examples which he uses pertain to seventeenth century science. However, if this is the account that Chomsky has in mind, then it is seriously in error. Chomsky concedes that his SFF is not on a par with the language faculty in terms of evidence in favour of it. While UD is used as an argument for both faculties, there is much more evidence which can be adduced in favour of a language faculty than for a SFF. Collins notes that there are four tests which can be used to argue that a particular ability is made possible by a distinct faculty: (1) a faculty-based competence must be uniform across the species, it cannot be a culturally specific capacity; (2) the competence must follow a strict ontogenetic course, explicit teaching must not make a significant difference to the speed of the development of a final competence arrived at; (3) the competence must, to some degree, be invariant over various pathologies, injuries and differences of intelligence, so that a disturbance of a face recognition faculty should

not *necessarily* lead to a disturbance of the language faculty; (4) the competence should reach normal maturity in the face of a *poverty of stimulus* (2002, 135).

Whether we interpret the SFF in the broad or narrow sense will influence whether we think that the SFF meets conditions 1-4. It is plausible to say that all cultures have science in the broad sense of constructing theories to explain, predict and control their experiences. However, science in the narrow sense is definitely culture specific, and did not even exist five hundred years ago. So, taken in the broad sense, we could say that science is species specific and does occur in all cultures. However, science in the narrow sense obviously does not occur in all cultures. In both the broad and narrow interpretations obviously scientific competence does not seem to follow a strict ontogenetic course and teaching does make a difference to it. In both interpretations the competence is not invariant over pathologies. Furthermore, understanding whether the competence arrives in the face of APS is impossible until Chomsky specifies how much stimulus we need in order to learn¹⁸ science in its broad or narrow sense. Until we have such a theory, it is impossible to answer the question or understand whether the stimulus is indeed impoverished.

So, as of yet, Chomsky has not provided us with any good reason to postulate a SFF in either the broad or the narrow sense. He has given us no reason to think that such a faculty is needed to account for our scientific knowledge. His primary reason for postulating the faculty is because he thinks that the faculty can help us overcome UD. However, we will see when discussing Quine that UD can be handled without postulating a mysterious faculty which has little independent evidence to support it.

¹⁸ Sketching a PoS for scientific knowledge seems impossible; it will need to give a plausible story of how Newton or Einstein learned or created their respective theories. It is wildly implausible that they just read them off experience, while it is just as implausible that they knew them innately. Given the difficulty we have in explaining the behaviour of animals and ordinary humans explaining how creative geniuses construct their theories seems impossible.

Quine, like Chomsky, recognises that global UD is a real possibility. However he does not deal with the problem in the same manner as Chomsky. Quine defined UD as follows:

If all observable events can be accounted for in one comprehensive scientific theory—one system of the world, to echo Duhem's echo of Newton—then we may expect that they can all be accounted for equally in another, conflicting system of the world. We may expect this because of how scientists work. For they do not rest with mere inductive generalizations of their observations: mere extrapolation to observable events from similar observable events. Scientists invent hypotheses that talk of things beyond the reach of observation. The hypotheses are related to observation only by a kind of one-way implication; namely, the events we observe are what a belief in the hypotheses would have led us to expect. These observable consequences of the hypotheses do not, conversely imply the hypotheses. Surely there are alternative hypothetical substructures that would surface in the same observable ways (2008, 228).

When trying to make the thesis explicit, Quine found that it was not as obviously true as people such as Chomsky believed. Two theories may be empirically equivalent and logically incompatible; however, Quine argues that in most cases these theories can be made compatible by reconstrual of predicates. According to Quine, a reconstrual of predicates involves:

By a reconstrual of predicates of our language, accordingly, let me mean any mapping of our lexicon of predicates into our open sentences (n-place predicates to n-variable sentences). Thus the predicate 'heavier than' might be mapped to an open sentence 'x is heavier than y', an identity mapping changing nothing, while the predicates 'molecule' and 'electron' might be mapped to the respective open sentences 'x is an electron' and 'x is a molecule', producing our example (*Ibid.*, 228).

So Quine concluded that for genuine UD to occur, we would need two theories which are empirically equivalent and logically incompatible and which cannot be made compatible by a reconstrual of predicates. He further noted that this type of UD is less obviously true than UD considered as ordinary inductive uncertainty.

While Chomsky views UD as a fact, Quine on the contrary views it as a theoretical possibility. So from Chomsky's point of view, this UD must be overcome and he postulates his SFF as a way of overcoming the problem. Quine, on the other

hand, thinks that global UD is only a theoretical possibility, so postulating an expensive piece of cognitive architecture to overcome a theoretical possibility seems to be a bit extravagant. Furthermore, postulating a SFF would result in extreme scepticism about the truths of science. A SFF which is genetically programmed as a result of a-rational evolutionary processes will obviously give no guarantee of being truth guiding when it constrains how we construct our scientific theories. Quine's argument about people who are wrong in their inductions being culled by natural selection will not work in the case of the SFF. Our scientific theories about quantum physics and large scale cosmology will deal with numbers and quantities which are so extreme that natural selection would have been blind to people's cognitive competencies in these areas. So for these reasons Quine would see it as pointless to postulate innate apparatus in the case of global UD.

The global UD which Quine discusses is a strong possibility and hence its consequences need to be clarified. The question is: supposing that two different scientific communities have different theories which rely on identical data, but which could not be made consistent by a reconstrual of predicates, which theory can we call true? Quine has varied at times with his answers to this question. In "On Empirically Equivalent Systems of the World", he claimed that we can take both theories as being true and alternate between them as we manage the flux of reality. While in *Word and Object*, he has claimed that we have no choice but to consider our own theory as being true.

When speaking of 'truth', Quine is obviously not speaking of truth in the sense of correspondence. Quine's primary objection to such a conception of truth is that it leads to an explosion of our ontology, admitting abstract entities such as 'relations' and 'properties'. He correctly notes that based on considerations such as Occam's

razor we should always attempt to keep our theories as simple as possible, and given that Tarski¹⁹ has shown us how to explicate the notion of truth without relying on the notion of correspondence²⁰ we should obviously go with the simpler less cumbersome theory.²¹

Quine speaks of truth in a deflationary sense. To say that the sentence ‘snow is white’ is true is simply to say the following: ‘snow is white’ is true if and only if snow is white. Quine’s gloss on this statement is:

To ascribe truth to the sentence is to ascribe whiteness to snow; such is the correspondence, in this example. Ascription of truth just cancels the quotation marks. Truth is disquotation. (2008a, 422)

Let us imagine that our species had somehow made all possible observations of the universe and we encountered an alien species who had made the same observations but had constructed a different theory than ours. If neither theory can be translated into each other’s language by a reconstrual of predicates, and if both theories are equally simple, we would have a real case of UD.

Chomsky understands UD as the claim that no data simply implies a theory; rather, the type of theory we will construct will be determined by the type of intelligence we are innately born with. So let us imagine that all possible observations are made by our species, and that our innate constitution combined with a series of experiments helps us construct a model of reality which we all recognise as true. Now imagine an alien arrives from a faraway galaxy who, as luck would have it, has a language faculty which is identical to ours, so he can communicate with us easily, but he has a SFF radically different from ours. Upon viewing the same data as us, this

¹⁹ The concept of Truth in formalised languages

²⁰ At least not correspondence in the sense of each word separately corresponding to a fact in the world.

²¹ I am not here suggesting that Quine’s conception of truth is correct, only that he has a sophisticated grasp of the concept, and is aware of and tries to answer other conceptions of truth in the tradition. Whereas Chomsky seems unaware of these problems

alien would construct a radically different theory of the world than us. Furthermore, the alien would consider his theory true. The question is, then, which theory is true?

Both Chomsky and Quine have landed in a similar quandary despite coming at the problem of UD from different angles. Quine recognises the difficult position he has been landed in: if he says that one theory is true and the other false, he must recognise that he seems to have little justification for doing so. Likewise, if he claims that two contradictory theories are both true relative to their own fixed vantage point, then he could be accused of being a relativist about truth. Quine tries to deal with the situation as follows:

Have we... so far lowered our sights as to settle for a relativistic doctrine of truth-rating the statements of each theory as true for that theory, and broking no higher criticism? Not so. The saving consideration is that we continue to take seriously our own particular aggregate science, our own particular world-theory or loose total fabric of quasi theories, whatever it may be. Unlike Descartes, we own and use our beliefs of the moment, even in the midst of philosophising, until by what is vaguely called scientific method we change them here or there for the better. Within our own total evolving doctrine, we can judge truth as earnestly and absolutely as can be; subject to correction, but that goes without saying. (1960, 25).

So we work within our total theory, accepting an absolute standard of truth and modifying our theory as necessary in order to manage the flux of experience as best we can. For Quine, this UD simply has to be lived with. We can do no better than accepting our own theory as true and subject to correction. We have no higher vantage point from which to judge our theory other than from within our own conceptual scheme. Quine's manner of dealing with this difficulty is unsatisfactory because it leaves us with a choice between accepting an intolerable relativism about truth, or simply assuming the truth of our own theory despite the evidence not telling us which one is correct. However, Quine would argue that this uncomfortable choice is simply a fact which is forced on us because of the facts about the relation between

evidence and theory. UD is simply a fact of life which we deal with as best we can within our best theory.

Chomsky with his emphasis on the SFF leaves us in a situation where it seems that two different species with different SFF may construct two contradictory theories. However, the fixed nature of their faculties may leave them in a situation where they cannot even formulate answers as to which theory is the true one. This would make truth relative to our mode of cognition and would lead to a type of relativism which is intolerable. Quine, on the other hand, leaves us with a choice between accepting relativism or arbitrarily choosing to believe that our theory is superior to a rival's theory despite the fact that we do not have evidence to support this belief. Neither solution is wholly satisfactory. However, while Quine's manner of dealing with this problem leaves us in an uncomfortable position, Chomsky's solution is untenable. Firstly, Chomsky's postulation of a SFF does not have sufficient evidence to justify it. Secondly, the SFF leads us to a conclusion where our best scientific theories are considered true only relative to our mode of cognition. Given that the SFF has so little evidence in favour of it, and it leads to such paradoxical results, the obvious solution is to deny that the SFF exists.

In this section, I have discussed how both Chomsky and Quine deal with UD. It was demonstrated that Chomsky does not offer sufficient evidence in support of his postulation of a SFF. It was further argued that he does not consider the relativistic consequences of his postulation of a SFF. Quine's manner of dealing with UD was then discussed. It was shown that Quine is at least aware of the spectre of relativism and has a theory of truth which he uses to try and cope with these problems. My aim was not to evaluate Quine's theory of truth, but rather, to illustrate how both theorists deal with the problem of UD in the area of epistemology. I argued that where

Chomsky's instinct when confronted with UD is to try and postulate an innate faculty of the mind to overcome it, Quine accepts UD as a fact of life, and seeks to build our theories around this fact, and only postulates innate apparatus to overcome the problem if he has behavioural evidence to support it. This point will be explored in greater detail when I discuss the Quine's IDT and Chomsky's APS in later chapters.

PART 3: THE PHILOSOPHY OF LINGUISTICS

SECTION 1: QUINE'S CRITICISMS OF CHOMSKY

In the previous two sections, I discussed how Quine and Chomsky view UD as it pertains to language acquisition and science acquisition. In this section, I discuss both theorists' views on the nature of language. I begin by considering some criticisms of Chomsky's conception of language which Quine put forth in his paper "Methodological Reflections on Current Linguistic Theory". I then discuss how Chomsky replies to Quine's criticisms, and whether his replies adequately answer Quine's criticisms. In the final section of this part, I discuss the relative merits of both theorists' views on language based on their debate with each other.

Rule following Fitting and Guiding:

My distinction between fitting and guiding is, you see, the obvious and flat-footed one. Fitting is a matter of true description; guiding is a matter of cause and effect. Behaviour fits a rule whenever it conforms to it; whenever the rule truly describes the behaviour. But the behaviour is not guided by the rule unless the behavior knows the rule and can state it. This behavior observes the rule (2008a, 215).

Quine begins his critique of Chomskian linguistics by distinguishing between two different types of rules: Fitting and Guiding. He claims that Chomsky uses a third intermediate type of rule; this is a type of rule which Quine claims is an implicitly guiding rule. Quine claims that Chomsky thinks that ordinary speakers of English are *guided* by rules even though these rules cannot be stated by the English speaker. According to this Chomskian picture, we can have two extensionally equivalent

grammars, each of which *fits* the behaviour of the child, neither of which explicitly *guides* the child, and only one of which is *true* of the child. Quine claims that if this intermediate way of following a rule is to be made sense of, we need some type of evidence which will help us decide which grammar the child is implicitly following. He claims that a person's disposition to behave in determinate ways, in determinate circumstances, is the way to make sense of which grammar the person is following. However, obviously these dispositions must go beyond well formedness if we are to use them to explain the distinction, because extensionally equivalent strings are indistinguishable from each other in terms of behaviour. He speculates that such dispositions which may be relevant are those such as the disposition to make certain transformations and not others; or certain inferences and not others. Quine further notes that he has no problem with dispositions; and points out that a body has a disposition to obey the law of falling bodies, while a child has a disposition to obey any and all extensionally equivalent grammars. However, he can make no sense of Chomsky's intermediate notion of rule following. He alludes to the ironic fact that while Chomsky seems to have no difficulty with the obscure notion of implicit guidance by rules, he has serious difficulties with the humdrum notion of disposition.

Chomsky replied to this criticism by first questioning the analogy of a child following the rules of grammar with a body obeying the law of falling bodies. He argued that this is a singularly misleading analogy because the rules of English grammar do not determine what a speaker will do in a given context in the same way that the law of falling bodies will determine, if a person jumps from a building, he will hit the ground at a specified time. According to Chomsky, what the rules of English grammar tell us is that English speakers will *understand* and *analyse* certain sentences in certain ways, and not in other ways. In other words, the linguist in

Chomsky's sense is trying to discover regularities in a person's linguistic competence; the linguist is not after a theory of performance. This reply of Chomsky's has two different strands to it: first, his distinction between competence and performance; second his claim that we can have no scientific theory of performance. His argument that by using idealisations, such as abstracting from memory and performance, we can gain some understanding of subjects like phonology etc. has something to recommend it. Idealisation plays a vital role in any science, and the appropriateness of an idealisation is to be judged by the success of the science which uses it. Given that generative linguistics has had some success, we can tentatively accept his idealisation. However, Chomsky's confident assertion that we cannot have a science of human performance/behaviour has little to recommend it. Quine's claim that people are obeying extensionally equivalent rules if their performance conforms to those rules is a perfectly legitimate claim. And Chomsky has offered us no reason to think otherwise.

Chomsky adds that even if we leave aside the purported disanalogy between the law of falling bodies and linguistic competence, Quine's formulation would still fail because he is guilty of treating physics and linguistics inconsistently. What Quine should have said if he wanted to remain consistent was:

English speakers obey any and all of the extensionally equivalent English grammars in just the sense in which bodies obey the law of falling bodies or the laws of some other system of physics that is extensionally equivalent (*Ibid.*, 188)

Chomsky claims that when put in these terms, we see that linguistics is no more in need of a methodological cure than physics is. Chomsky's argument here is merely repeating the criticism which he made against Quine's IDT argument. He is claiming that yet again Quine is guilty of treating UD as fatal in linguistics but as harmless in physics.

Quine's criticism was that Chomsky was using a third type of rule which was obscure because it went beyond fitting the behaviour of the person, but was not consciously guiding the behaviour of the person. Chomsky notes that people are rarely, if ever, guided by rules in the sense of being able to state the rule they are following. Furthermore, he argues that linguists can go well beyond rules that merely fit the behaviour of people. He says that if we accept the same realistic attitude towards linguistics that we do with physics, then we can say that people are obeying rules which are really encoded in their brains, but of which they are usually not consciously aware. We can use various different pieces of evidence to decide between extensionally equivalent grammars.

To make his criticism of Chomsky's conception of rules explicit, Quine discussed a toy example of two extensionally equivalent grammars which a person could obey. He asks us to imagine a string (abc) which can be divided up as extensionally equivalent grammars (ab) (c) versus (a) (bc). He notes that from a behavioural point of view, we could say that a person is following either of the grammars and that if we ask the native which is the correct grammar he will not be able to tell us. In this situation, he claims that Chomsky's view that the person can be said to be implicitly guided by rule (ab) (c) as opposed to (a) (bc) is senseless.

Chomsky argues that that there is no mystery in deciding whether the child is implicitly guided by rule system (a) as opposed to (b). In, this situation, he argues all we have is a problem which can be solved by the ordinary methods of the natural sciences. Quine had argued that a natural solution to deciding whether (ab) (c) or (a) (bc) was the rule which implicitly guided the people, would be to ask them which rule they followed. He notes, of course, that the people will not be able to tell us. So he argues that since the natives cannot decide which rule they are implicitly following,

and both rules are compatible with the dispositions of the ordinary language user, then the notion of one of the rules implicitly guiding the child is senseless. Chomsky agrees that asking the native will not get us far in deciding between rule (a) and rule (b), and he suggests a different way of dealing with the problem. He cites various different types of evidence which could theoretically bear on the problem. He discusses how the linguist could have evidence which suggests that intonation patterns are determined by structure. This evidence could be derived by studying our own language as well as other languages. He argues that such evidence might bear on the choice between the two proposed grammars. So we might discover that the rules needed for other cases give the correct intonation if we take the constituents to be those of (a) instead of (b). Whether this type of evidence really occurs or not is irrelevant. Chomsky's point is that Quine is incorrect to assume that no evidence on the topic is attainable beyond evidence gained by questioning the natives.

Chomsky cites a quote from Quine's paper "Linguistics and Philosophy" where Quine claims that there is much innate apparatus which will need to be discovered to tell us how the child gets over the great hump that lies beyond ostension. Quine further notes that if Chomsky's anti-empiricism says nothing more than that conditioning is not sufficient to explain language learning, then it is of a piece with his IDT argument. Chomsky notes that if Quine really believes that there is further innate apparatus waiting to be discovered by science, then this casts doubt on his claims in both his "Methodological Reflections" paper and in *Word and Object*. Chomsky asks us to consider the sentence: 'ABC'. Quine had claimed that we have no evidence which can help us determine whether a subject is implicitly following the rule (ab) (c) or (a) (bc). However, according to Chomsky, given that Quine has no problem with innate mechanisms of any sort, then surely it is possible we will

discover an innate mechanism in our species which determines that we follow (ab) (c) as opposed to (a) (bc) or vice versa. So in this respect, he claims, Quine (1969) holds doctrines which run counter to the doctrines he accepts in (1960) and (1970).

Quine would deny this charge of Chomsky's. What Quine has said is that he has no problem with innate apparatus of any sort as long as such apparatus can be made sense of *behaviourally*. When he is discussing the notion of rule following, he is claiming that if two grammars both *equally* fit the behaviour of the subjects, and each subject is not consciously *guided* by these rules, then to claim that the person is implicitly guided by the rule is senseless. The postulation of innate apparatus which ensures that the child follows the rule (ab) (c) as opposed to rule (a) (bc) is pointless unless we have some behavioural evidence to justify such a postulation. The type of evidence which Chomsky claims could be useful to the linguist in deciding whether the subject is following rule A or B, i.e. perceived structure in intonation, is *behavioural evidence* so it does not bear on the type of considerations which Quine is concerned with.

A Chomskian could argue that if Quine would accept evidence such as structure intonation as a way of distinguishing between extensionally equivalent grammars, then he would be committed to accepting the Chomskian intermediate notion of a rule which implicitly guides the subject. Quine would of course rightly deny this. He would claim that, if we assume for the sake of argument that Chomsky's made-up example of structure in intonation pattern is correct, then it would follow that the two extensionally equivalent grammars do not fit the totality of linguistic behaviour of the relevant subjects. He could argue that his flat-footed distinction between the two different types of rules can accommodate the type of example which

Chomsky put forth. Quine could further claim that Chomsky has not fully explained his notion of the third intermediate notion of a rule.

This debate about the notion of a rule is a real point of contention between both thinkers. Chomsky has tried to explicate his intermediate notion of a rule as something in the brain which unconsciously guides the language user. This conception of a rule does not fit in neatly with Quine's two conceptions of a rule; the question is whether Chomsky's third notion captures something real which is missed by the two Quinean notions.

For simplicity's sake, let us consider a particular rule of English (a rule which will be discussed extensively throughout later chapters), the subject-auxiliary rule which states that when forming a question, one must move the main auxiliary to the front of the sentence. According to Chomsky, the subject auxiliary rule is a rule of English (and Spanish) though not of all languages; so it is a rule which requires some experience in order for acquisition of it to take place. Children are only capable of learning this rule because they are genetically programmed to follow certain linguistic universal principles such as the following one: All languages are structure dependent. I will not here go into the justifications which Chomsky offers to support his belief that people follow this rule as I have already done so in Section 1. Instead I will discuss the actual nature of the rule itself.

On the Chomskian conception, the child is guided by certain rules which he does not know and cannot state (unless he has knowledge of linguistic theory). The question we now need to ask is why Quine would have difficulties with these rules, and can such rules be accommodated within his flat-footed conception of rules? Quine's emphasis on reinforcement and induction means that he would more than

likely feel that the postulation of such innate rules is unnecessary²². Nonetheless let us assume that Chomsky is correct that poverty of stimulus considerations dictate that rules such as structure dependence are known innately and that parametric variations of these universal rules result in particular languages being spoken. The question which needs to be asked at this point is whether we can make sense of the notion of ‘rule’ which Chomsky is postulating here?

When Quine wrote his “MRCLT” in the early 1970s Chomsky was operating with a rule-based conception of language. From about 1980 onwards, with the inception of his principles and parameters approach, Chomsky’s conception of language changed. Though the change from a rule-based approach to a principles and parameters approach was a significant empirical advance, it does not have any effect on the criticisms which Quine brings to bear in his “MRCLT”. So let us consider a question such as the following one: ‘Is the man who is happy over thirty?’ which is derived from the statement ‘The man who is happy is over thirty’. When forming this question, Chomsky claims the child begins with the statement and unconsciously applies the rule: move the main auxiliary to the front of the sentence. According to Chomsky, none of this is done consciously; rather, what happens is that a mechanism in the brain which is genetically programmed will interpret the data of experience and construct a model. The mechanism will determine what the rules of the language are, using the universal principles of the language and having the parameters set through experience. When these parameters are set, we can say that the child is unconsciously following various different rules of English, such as the subject/auxiliary inversion.

Suppose we assume (falsely) that Quine would accept Chomsky’s poverty of stimulus argument and agree that rules such as structure dependence are innate and

²² Quine may well be correct on this point; we will discuss whether Chomsky’s innate rules are necessary to explain language acquisition, when discussing poverty of stimulus arguments in Chapter 3.

that these rules determine that ordinary humans will derive rules such as auxiliary inversion when placed in certain linguistic environments. The question we now need to answer is how would Quine characterise these facts? The key point here is that Quine would focus on performance facts; he would ask whether people are disposed to form questions in ways consistent with these rules. Answering these questions would involve studying various different corpuses and seeing how people actually form questions when talking with others, or when writing various different texts. It would also involve constructing various different controlled experiments to see whether the manner in which questions are formed varies in the circumstances of speaking/writing. What Quine would object to is the reliance of Chomsky et al. on the intuitions of others as to what is or is not an acceptable way of forming questions. Quine would claim that the salient aspect of our studies should be the behaviour of the subject, not the intuitions of acceptability or unacceptability by the subjects of various constructions.

However, the next question to ask is what Quine's response would be if the behaviour of the subjects and the acceptability tests lined up perfectly. So Quine would have the following facts to account for²³: poverty of stimulus considerations, subject's intuitions of the acceptability or unacceptability of certain constructions, and subject's behaviour in certain determinate circumstances. In this circumstance, Quine would claim that the person's behaviour fits with any and all extensionally equivalent grammars which capture the behaviour of the subject. He would say that we are justified in claiming that the person's behaviour conforms to a particular rule system (and other extensionally equivalent systems). He would have to object to any postulation of innate apparatus because there is no justification for postulating one

²³ In reality, Quine does not have this collection of facts to account for because as I argue in later chapters, Chomsky's poverty of stimulus argument is dubious, and performance data does not match up perfectly with competence data.

rule system over another one. He would accept any innate apparatus if it was determined by behavioural facts; however, no behavioural facts will help us decide between attributing rule system (A) over rule system (B), as they are extensionally equivalent. So on this Quinean picture, the rules which we claim the behaviour of the child fits do so in the same way the behaviour of physical objects fits certain rules of physics. So rules discovered in this behaviourist manner would easily conform to Quine's flat-footed conception of rules fitting the behaviour of the subject. There would be no need for the Quinean to postulate an intermediate type of rule which implicitly guides the subject.

There is, however a difficulty with this Quinean conception of the behaviour of the child fitting certain rules. The difficulty stems from normative issues. The Chomskian picture conforms to our pre-theoretical intuitions about language in one clear sense. It seems obvious that when I construct a new sentence, the sentence will be grammatical or ungrammatical according to the rules of the language. So, for example, most people would believe that if I construct a new sentence there will be a fact of the matter as to whether the sentence is grammatical or not. However, if we are deriving our rules of language by studying how people actually perform, then at best, all we can speak of is the probability of an utterance with a certain syntactic structure occurring or not. We will have no warrant to claim that the sentence is grammatical or ungrammatical. From a behavioural point of view, we can say that the sentence is atypical but not that it is incorrect. This seems like a serious difficulty with Quine's conception of the nature of rules. He could reply to this concern by claiming that we could test whether the sentence was grammatical by the lights of the linguistic community by asking its members. However, this reply does not solve our problem; rather, it reintroduces the problem in a different guise. If we ask people whether this

or that construction is acceptable, we are testing their intuitions in the same manner that Chomsky recommends. If people's intuitions all agree that a certain construction is ungrammatical, but performance data indicates that the construction is used and accepted in ordinary speech, then we have arrived at an impasse. Furthermore, if people's intuitions of acceptability disagree with their actual performance, then Quine would argue that we would have to give preference to performance data over people's intuitions of acceptability. The problem with this is that if people's intuitions cannot be used to help us decide between a correct or an incorrect linguistic utterance, then we seem to have no way of doing so. All we have are linguistic regularities, and constructions which are irregular from that point of view.

To clarify the above difficulty let us consider a computer. From a performance point of view, the computer exhibits certain regularities. So, for example, if I press the caps lock key when typing in Microsoft Word it results in the words I type afterwards being capitalised. When I press the caps lock key again, the computer no longer types in capital letters. Now imagine if one day I press the caps lock key and the following symbol * appears on my screen *. Imagine I continued pressing the key for a while and the symbol * kept appearing. You can be quite sure that if this happened, nobody would think to themselves 'Strange, the probability of my letters being capitalised when I press the caps lock button has just been reduced'. It is a safe bet that anybody who noticed that the caps lock key being pressed resulted in * being typed would assume that the computer was broken. This assumption would be based on the fact that we know that the computer was designed for a particular purpose which it is no longer achieving efficiently. So we would assume that some part of the computer was broken and set about getting it repaired.

Now in natural language such an approach is possible as well. Take, for example, people with severe schizophrenia, or with some form of aphasia. People with schizophrenia sometimes speak with what is known as word-salad. Such word-salad sentences sometimes exhibit syntactic, semantic and pragmatic deviance. People with some forms of aphasia are sometimes incapable of forming sentences into syntactic units. Analysing sentences from schizophrenics and aphasics is no trivial matter, and understanding the way such sentences go wrong is a flourishing field of psycholinguistics and neurolinguistics. Any theory which claimed that the speech of aphasics and schizophrenics was just statistically unlikely rather than incorrect seems to be seriously deficient. Furthermore, as opposed to the case of the computer, we cannot say in this case that such people are in error because they are behaving contrary to their designers' intentions. Who designed English that we would be sinning against if we speak ungrammatically? Chomsky would answer that people speaking in such a deviant manner are breaking the implicit rules of universal grammar and this explains our judgements that such people are speaking ungrammatically²⁴. Quine, who rejects appeal to implicit rules which govern what sentences we accept as grammatical, will need to tell a different story of how such sentences are viewed as deviant.

In the rough and ready world of ordinary discourse, Quine recognises that intentional idioms are indispensable. It is only when we are trying to limn the true and ultimate scheme of reality that such idioms have no place. It is also in the rough and ready world that people's behaviour (including their verbal behaviour) is viewed as deviant. It is from this rough and ready vantage point that people are judged to be

²⁴ Obviously, there is more to schizophrenic word-salad than ungrammatical sentences; however, I am just focusing on ungrammaticalness because it is directly relevant to the debate between Chomsky and Quine on this particular point.

suffering from schizophrenia and aphasia²⁵. Quine would have no problem with using this pragmatic idiom in daily discourse while using the more precise discourse of neurology and behavioural science when trying to limn the ultimate nature of reality. So in this sense the schizophrenic and aphasic objection does not in any way affect Quine's argument.

To the objection that Quine's flat-footed conception of rule following cannot handle normative notions like correct and incorrect grammar the same reply as above will suffice. First, Quine can point out that in ordinary discourse, when applying the dramatic idiom of intentionality, we can judge that certain statements are unclear, badly structured, pragmatically deviant, etc. However when we are limning the true and ultimate structure of reality, we can do no better than say that people are following whatever extensionally equivalent grammars can be constructed to systematise their utterances. To people such as Chomsky who claim that people follow one true grammar as opposed to any and all extensionally equivalent grammars, Quine would reply that we have in-principle no behavioural evidence which can decide between them. And without such evidence, all we are justified in positing is Quine's flat-footed conception of rule following.

As we have already seen, Chomsky would greatly object to this characterisation because he does not think that syntactic rules can be studied in terms of performance. The important thing to note is that at this point we do not know whether a theory of syntax, semantics, etc. is possible in terms of performance. We do not know whether people's grammatical intuitions match up with their actual linguistic behaviour, though I will consider some corpus data which bears on this point in later chapters. A further difficulty stems from the fact that Chomsky does not

²⁵ The rough and ready vantage point includes the judgements of psychiatric workers, who help themselves to intentional idioms and physical idioms as is useful for their purposes.

typically offer statistical analysis of people's grammatical intuitions, and when studies have been done²⁶ the result has revealed much greater variety in people's intuitions than Chomsky would admit. The central point to take from this section is that, Chomsky's confident assertions aside, we do not know whether it is possible to construct a theory of performance yet, and discovering whether this is possible or not will require empirical research, not rhetoric. This bears on the debate between Chomsky and Quine on the nature of rules. If a science of behaviour is tractable then Quine's conception of the behaviour of various subjects fitting syntactic rules will be the most accurate way of conceiving the facts. However, if it is shown that Chomsky is correct that a science of human behaviour is impossible, then Chomsky's conception of a third intermediate type of rule will be the correct picture. As always, this is an empirical question and I will consider some empirical evidence which bears on the point in later chapters.

SECTION 2: CHOMSKY'S CRITICISMS OF QUINE

In his 1969 paper "Quine's Empirical Assumptions" Chomsky criticised Quine's conception of the nature of language. In particular, he claims that Quine's definition of language is confused. Quine had defined a language as follows:

We are concerned here with language as the complex of present dispositions to verbal behaviour, in which speakers of the same language have perforce come to resemble one another; not with the processes of acquisition whose variations from individual to individual it is in the interest of communication to efface... Meanwhile what is before us is the going concern of verbal behaviour and its currently observable correlations with stimulation. Reckon a man's current language by his current dispositions to respond verbally to current stimulation, and you automatically refer all past stimulation to the learning phrase. Not but that even this way of drawing a boundary between language in acquisition and language in use has its fluctuations, in as much as we consult our convenience in what bound we set to the length of stimulations counted as current. This bound, a working standard of what to count as the specious present, I call the modulus of stimulation (1960, 27-28).

²⁶ For some research on the statistical analysis of people's grammatical intuitions see Lappin and Clark *Linguistic Nativism and Poverty of Stimulus*.

Chomsky interprets the definition of language as ‘a complex of present dispositions to verbal behaviour’ to be a structure that can be represented as a set of probabilities of utterances to be spoken in particular circumstances. He claims that the notion of the probability of a sentence being spoken is utterly useless under any known interpretation of probability. He even goes as far as to claim that the probability of a person speaking a particular sentence such as ‘My arm hurts’ or ‘Tuesday follows Monday’ is no greater than the probability of him speaking a given sentence of Japanese. He goes on to specify that if Quine means ‘probability relative to a situation’ this will add nothing to Quine’s conception if situation is used in its objective sense.

Interpreting Chomsky’s above criticisms is extremely difficult because he offers very little evidence to support his claims. He offers no evidence that the probability of my speaking a sentence such as ‘Tuesday follows Monday’ in a particular situation is the same as my speaking a sentence of Japanese. Furthermore, when criticizing the notion of ‘probability relative to a situation’, he does not specify what he means by situation being used in its objective sense.

He goes on to claim that if a ‘complex of dispositions’ is determined on empirical grounds, then only a few clichés, conventional greetings, etc. have any chance of being associated to the complex defining the language, since few other sentences are likely to have non-null relative frequency. He argues for this by claiming that as the size of a person’s corpus increases, the frequency of particular sentences being spoken will decrease dramatically. Chomsky’s argument seems to be that given the fact that languages are potentially infinite a fact which has been proven mathematically, then the probability of a particular sentence being spoken out of the infinitely many possible ones is virtually zero. So Chomsky argues that if we want to

define a language as ‘a complex of dispositions’, then we will be forced to conclude that language is not only finite but extremely small.

Chomsky notes that Quine seems to vacillate somewhat with his notion of speech dispositions. First Quine formulates his IDT as the claim that ‘manuals for translation can be set up which are compatible with the totality of speech dispositions but which are incompatible with each other’. Chomsky claims that there are two different ways that we can take ‘the totality of speech dispositions’:

(1) We take the ‘totality of speech dispositions’ of an individual to be characterised by probability distributions for utterances under detectable stimulus conditions;

(2) We could understand the notions of ‘disposition’ and ‘situation’ more loosely.

The difficulty with (1) is that except for the null set there is equi probability of my uttering a particular sentence from my language, and my uttering a particular sentence from Japanese.

The difficulty with (2) is that there will be so few similarities between what different people are disposed to say in different circumstances that constructing a manual of translation by these techniques will be impossible.

Chomsky claims that Quine manages to avoid these problems by shifting from speaking of ‘the totality of speech dispositions’ to speaking of ‘stimulus meanings’- that is, dispositions to assent or dissent circumscribed by one narrow experiment. Chomsky objects to this shift because he claims that a person’s disposition to ‘verbal response’ under arbitrary stimulus conditions is different from his disposition to be prompted to assent or dissent to a sentence under a particular experiment circumscribed by Quine. Chomsky even goes as far as to claim that by shifting the ground in this way Quine has trivially established the thesis of IDT, because, if we only allow this narrowly circumscribed experiment as evidence, it will trivially follow

that different manuals of translation can be set up which are incompatible with each other but which are compatible with all the evidence.

Quine rejects Chomsky's reading of what the term disposition means in his IDT argument. First, he rejects Chomsky's claim that he shifts the ground by switching from a disposition to overt speech in particular circumstances to a disposition to assent or dissent to a sentence in a particular circumstance. Quine correctly notes that the former is a subset of the latter, and that focusing on the latter helps him key in on the notion of stimulus meaning. Furthermore, since grasping the stimulus meaning of a sentence is our entering wedge into the language of the native, Quine's focusing on this subset makes translation a much more tractable process. He also notes that focusing on assent and dissent removes the supposed problem of the equi-probability of a sentence of English and Japanese being spoken.

In various different articles written since his "Reply to Chomsky" Quine has criticised Chomsky for his nihilistic attitudes towards dispositions²⁷. Chomsky replied that it is not dispositions that he objects to but rather Quine's definition of language as 'a complex of present dispositions to verbal behaviour'. The difficulty with this definition is that a disposition to verbal behaviour only makes sense if we construe it as the probability of an utterance being used in certain circumstances. However, according to Chomsky, no objective sense can be made of such a notion of probability. While Quine's restriction of the probability of a person answering yes or no to particular observation sentences being put to them may avoid Chomsky's probability objection, it does not save Quine's definition of language. It would be absurd, Chomsky correctly notes, to define language as a person's disposition to answer yes or no to a particular utterance in a particular experimental situation.

²⁷ For some examples of such criticisms see Quine's papers "Linguistics and Philosophy" and "Methodological Reflections on Current Linguistic Theory", as well as his book *The Roots of Reference*.

This objection of Chomsky's is very serious. The claim that at its most basic level, Quine's definition of language is false would have serious implications for his overall epistemological story, which is deeply connected to his picture of language acquisition. However, it is important to note that Chomsky has not proven that language is not disposition to overt behaviour in particular circumstances. From the fact that humans can *in principle* mouth an infinite amount of unique sentences, it does not follow that their *actual* utterances cannot be accounted for in dispositional terms. Establishing what the probability is of an utterance being spoken by a particular person, at a particular time, is an empirical task, and this task will require empirical evidence to determine the answer. Whether Quine's definition of language turns out to be false or true is an empirical matter to be judged by the success of the theory. Unfortunately, Chomsky's polemics against this behaviouristic conception of language resulted in a situation where most cognitive scientists did not even try to test whether language could be understood as a dispositional state. So the debate between Chomsky and Quine on this point remains wide open: it is an empirical question for which we do not as of yet have sufficient evidence to resolve.

It is worth noting that Chomsky's polemics against Quine's conception of language not only foreclosed any possibility of testing Quine's claims, but that his alternative approach leaves us with a picture of language which is incoherent. In his introduction to Chomsky's *Cartesian Linguistics*, James McGillivray sums up Chomsky's conception of language as follows:

Humans seem to be able to produce an unbounded number (different arrangements of words) of sentences without causal antecedent (that are externally and internally 'stimulus free') although perhaps prompted (in Descartes time, 'occasioned') by questions or other factors, while remaining appropriate and rational in what they produce. Clearly their production cannot be the result of some deterministic mechanism that offers the answers it caused to produce, nor even of a mechanism that, given a question and a context, provides one that is a specifiable (and thus bounded) range of answers. If

someone insists on the opposite, they are hereby challenged to come up with the mechanism, for no deterministic system can yield any of an unbounded set of sentences, for a specific discourse context, all of which are coherent and appropriate. If the conceptual difficulties involved in this do not convince, then a lack of success coming up with a science of such a mechanism suggests that those who insist on a deterministic mechanism are on thin grounds indeed. (1966a, 41)

This section is taken from McGillivray's introduction to the new edition of Chomsky's book *Cartesian Linguistics*. Given that Chomsky read the introduction prior to the book being published and used McGillivray's article as his introduction to his book, we can safely assume it accurately represents Chomsky's views. The key point we will focus on is the emphasis on the fact that the language use of ordinary speakers is coherent and appropriate to the circumstance, while at the same time being free from *external* as well as *internal* stimulus control. At other points, McGillivray speaks of language production being appropriate to the situation but being uncaused by stimulus. How anyone could produce a sentence which is uncaused by either external or internal stimulus is an utter mystery. In fact the very idea of such thing occurring is incoherent. Chomsky accepts this characterisation and claims that the productive use of language may be a mystery which exceeds the grasp of a species such as our own.

The important point to note is that when Chomsky criticises Quine for claiming that language use can be explained dispositionally, he is not merely criticising Quine for trying to explain language use in crude empiricist and behaviourist terms; he is criticising Quine for trying to explain language use in any deterministic manner. Chomsky may be correct that we cannot explain language use because of some sort of cognitive closure peculiar to our species. However, he would admit himself that the idea of cognitive closure is merely a conjecture and not something which we know for certain pertains to our use of language. Furthermore,

the evidence for the free use of language is not as overwhelming as Chomsky claims it is. There is the evidence that we can construct an infinite amount of sentences because of the combinatorial aspect of syntax, but this is only an abstract possibility, not necessarily something that any humans as a matter of empirical fact engage in. There is also the introspective evidence which tells us that we can freely use whatever sentence we want independent of internal or external stimulus. However, our own introspective reports should not dictate the type of theory we construct about language use, because introspective reports are notoriously unreliable. Chomsky himself has correctly argued that we should not let our intuitive folk physics, and folk psychology interfere with our scientific physics or psychology. By parity of reasoning he should not let people's folk conception of their freedom constrain scientific theories of the use of language.

McGillivray in the above quote argues that if opponents of Chomsky believe that they can explain the free use of language in a mechanistic way then they should proceed to do so. He further argues that the fact that there has not been any successful science of linguistic use shows that those arguing for a deterministic explanation are on thin grounds. However, McGillivray fails to note that it was because of Chomsky's polemics against people such as Skinner and Quine that most cognitive scientists did not even try to explain language use mechanistically. Chomsky's foundational views on the nature of language won more adherents than Quine's did; however, the important point in this context is that Chomsky won his adherents not with evidence but with rhetoric. Whether Quine's dispositional conception of language can be salvaged is an empirical question which can only be answered by the relevant experiments being done. At this moment in time the debate between both thinkers remains well and truly open.

SUMMARY

This chapter was divided into four different parts. Each part discussed Chomsky and Quine's views on different topics. The primary aim of the chapter was to get clear on where both thinkers disagree with each other, and where their supposed disagreement stems from a mutual misreading of each other. Part 1 was a discussion of naturalism as a philosophical position. I argued that, despite what Chomsky thinks, both thinkers are actually methodological naturalists. So there is no real disagreement between the two thinkers on this particular topic. In Part 2 both thinkers views on language acquisition were discussed. It was argued that here a real difference between both thinkers exists. Quine emphasises reinforcement as playing a primary role in language acquisition, while Chomsky's Poverty of Stimulus Arguments maintain that grammatical structures are acquired with minimal reinforcement or induction. So on this topic there is clear disagreement which will be discussed in detail throughout the rest of the thesis. In Part 3 I discussed Chomsky's and Quine's different views on the nature of Underdetermination. Chomsky typically responds to UD by postulating an innate faculty which overcomes it; this can be seen by his reaction to Goodman's and Kripke's problems. In this vein, to help us overcome the UD of our scientific theory by the data of experience, he postulates a SFF. The evidence for this faculty is extremely weak and I argued for this reason that the hypothesis must be rejected. Having discussed how Chomsky deals with UD in science I then discussed Quine's attitude to UD. Quine's attitude to UD is more nuanced than Chomsky's. When it comes to Goodman's problem of induction he postulates innate constraints to overcome the problem because he has behavioural evidence to justify such a move. However, if Quine has no behavioural evidence to justify his postulation of innate apparatus, then he will not postulate it. If it turns out that in a particular area we have

UD but have no behavioural evidence to justify us in postulating innate apparatus, then Quine argues that we will just have to live with the UD. So to this end, Quine builds his theories around the fact of UD rather than trying to overcome it by automatically and without evidence postulating innate faculties. I introduced both thinkers' differing views on the UD of scientific theory not because it is a main topic of the thesis, but because it illustrates both thinker's views on UD and this will be extremely relevant for the rest of the thesis. The relevance of this will be seen when we are discussing the main topic of the thesis which is both thinkers' views on the nature of language. When we discuss Chomsky's attitude to the IDT, we see that he views it as a problem of UD which can be overcome by innate constraints. However, from a Quinean perspective, since there is no behavioural evidence to help decide between the alternative translation manuals there is no justification for postulating innate apparatus to overcome the problem. Likewise, Chomsky's APS can be viewed as the UD which faces the child learning the syntax of his language. So getting clear on both thinkers' views on UD exposes what both thinkers' attitudes are to the topic. In Chapter 3 and 4 we will tackle the issues of the IDT being a form of UD and the APS being a form of UD and discuss how both thinkers deal with the problem. Part 4 of Chapter 1 is a discussion of both thinkers' views on the actual nature of language and the correct way of studying it. This part consists of debates on whether language is best studied in terms of competence or performance. It is further discussed what the nature of a linguistic rule is. It was argued that the question of whether language can be studied dispositionally has not yet been answered because Chomsky's rhetoric has won over so many adherents that the behaviourist alternative has never been studied empirically.

CHAPTER 3: QUINE CONTRA CHOMSKY: THE QUESTION OF INDETERMINACY.

INTRODUCTION

In the previous chapters we saw the different ways that Chomsky and Quine dealt with Underdetermination. Whenever Chomsky encounters UD he seeks to overcome it by postulating innate constraints; in contrast, Quine treats UD as a fact of life which simply has to be lived with. In this chapter I will discuss the IDT. Chomsky argues that the IDT amounts to nothing more than UD. He argues that the UD in physics is overcome by our innate science-forming faculty, and that UD in language raises no more difficulties than UD in physics. In the case of language, the IDT is overcome by innate constraints imposed by the rules of our language faculty.

Chomsky's interpretation of the IDT has been accepted by the vast majority of contemporary cognitive scientists and they have proceeded to flesh out his proposal that the IDT can be overcome by innate constraints. However, when cognitive scientists are concerned with the IDT, they are typically interested in only one area of it, the inscrutability of reference. They typically argue that the inscrutability of reference is a form of UD in concept acquisition and that this UD can be overcome by postulating innate concepts which we use to learn our first words. In this chapter I will consider the attempt by contemporary cognitive scientists to overcome the inscrutability of reference by postulating innate concepts. Furthermore, I will analyse what effect the supposed overcoming of the inscrutability of reference by postulating innate concepts has on the indeterminacy of translation argument.

PART 1: THE INDETERMINACY OF TRANSLATION-A SYNOPSIS

Quine begins Chapter 2 of *Word and Object* by stating in a loose way what he takes the main point of the IDT to be. He outlines the main claim of the chapter as follows:

The thesis is then this: manuals for translating one language into another can be set up in divergent ways, all compatible with the totality of speech dispositions, yet incompatible with each other. (1960, 27)

In this section I will outline what premises Quine puts forth in *Word and Object* which led him to such radical conclusions about the nature of translation. It will be shown that it is the premise of the Inscrutability of Reference which Chomskians are primarily concerned with because they believe that it is a clear case of UD in word learning which can be overcome by postulating innate concepts. They do not typically consider what effect the overcoming of the inscrutability of reference will have for the IDT argument. At the end of the chapter I will argue that it has little effect on the status of the IDT. However, prior to this, I outline the IDT and consider the experimental research which has been brought to bear on it by thinkers influenced by Chomsky.

Quine asks us to imagine the situation of a field linguist who is confronted by a tribe of a hitherto untouched people. He stipulates that this tribe has never been in contact with any other tribes the linguist may know. So no hints can be given from other tribes which may have been in contact with our newly discovered people. He asks what evidence we can appeal to as we try to translate the native's language into our own. The first thing he thinks we must be able to do is discover how our natives say 'yes' and 'no', because the whole key to radical translation comes from assent and dissent. This itself is no trivial process, because it is well known from anthropological studies that not all cultures use the same gestures as we do to indicate 'yes' and 'no'. When trying to grasp what the native is saying, he argues that we should begin by

trying to translate the sounds which the native regularly uses in the presence of fixed stimuli. So, for example, if the native mouths the sound 'gavagai' every time a rabbit is present in the vicinity, our linguist will assume that the sound is meant to somehow indicate the presence of a rabbit. He will translate 'gavagai' as meaning 'Lo a rabbit'.

Once our linguist has settled on a tentative translation of sounds like 'Gavagai', he will need to test his conjectures, and doing this will require his questioning the native. So our linguist can try saying 'Gavagai' in situations where the native has said it in the past. If our native looks pleased and says 'evet' then we can translate 'evet' as meaning 'yes'. He can then try saying 'Gavagai' in obviously wrong contexts and then learn how they say 'no' i.e. 'wok'. With his tentative translations of 'yes' 'no' and 'Gavagai' the linguist now has his entering wedge into the language of the native. The important point to note is that when trying to learn the meaning of the sounds that the native is uttering, we must begin by keying his sounds to aspects of the external environment. So this approach differs from the methods of seventeenth-century philosophers, such as John Locke, who assumed that words got their meaning by referring to ideas in the mind. The seventeenth-century tradition had it that to discover what the native meant we had to discover the ideas in his mind which he was expressing through language. In contrast, for Quine to discover what the sounds of language mean, we must try to key the sounds to stimulations impinging on the organism. Chomsky believes that seventeenth-century philosophers viewed such matters correctly, and that Quine's new approach was leading philosophers in the wrong direction.

For Quine, the translator's job is not to discover the meaning of sounds in terms of ideas in the mind. Rather, he should try to discover the stimulus meaning of

the terms used in language. ‘Stimulus meaning’ is a technical term and before proceeding, I will give a definition of what the term means. When our native says ‘Gavagai’ we see a rabbit in the vicinity nearly all of the time; likewise, when we say ‘Gavagai’ in the presence of a rabbit they typically assent to what we say. However, Quine warns that what they are assenting to are *stimulations* and not rabbits. He warns this because there will be times when the native will assent to ‘Gavagai’ even though what he sees is only something that looks like a rabbit, or he may dissent because from afar the thing does not look like a rabbit (though it is in fact one). So given that ‘Gavagai’ is to be correlated with a stimulus and not with a rabbit, the obvious question is: what exactly is a stimulus? Quine’s gloss on this is:

A visual stimulation is perhaps best identified, for present purposes, with the pattern of chromatic irradiation of the eye. It is not, however, adequate to think of the visual stimulations as momentary static irradiation patterns. To do so would obstruct examples which, unlike ‘rabbit’, affirm movement. And it would make trouble even with examples like ‘rabbit’, on another account: too much depends on what immediately precedes and follows a momentary irradiation. Better, therefore, to take as the relevant stimulations not momentary irradiation patterns, but evolving irradiation patterns of all durations up to some convenient limit or modulus. (*ibid.*, 31-32)

Visual stimulations are not the only stimulations which we receive; there are others corresponding to all the senses, such as auditory stimulations. So Quine believes that once we have such an empirical criterion of meaning in place we can then begin to talk about translation. Typically, translation was meant to discover the ideas that the two languages shared; now, with our new externalized empiricism, we could say that ‘Gavagai’ could be translated as meaning ‘Rabbit’ if they shared the same non-verbal stimulation.

Once Quine had put the notion of stimulation as being what translations shared, he had effectively replaced meaning with stimulus meaning. He defined stimulus meaning as follows:

We may begin by defining the affirmative stimulus meaning of a sentence such as ‘Gavagai’, for a given speaker, as the class of all the stimulations (hence evolving ocular irradiation patterns between properly timed blindfolding) that would prompt his assent. More explicitly, in view of the end of sec 7, a stimulation Θ belongs to the affirmative stimulus meaning of a sentence S for a speaker if and only if there is a stimulation Θ' such that if the speaker were given Θ' and asked S, and then were given Θ , and then were asked S again, he would dissent the first time and assent the second. We may define the negative stimulus meaning similarly with assent and dissent interchanged, and then define the stimulus meaning as the ordered pair of the two... The imagined equating of ‘Gavagai’ and ‘Rabbit’ can now be stated thus: they have the same stimulus meaning. (ibid., 33)

Now, bearing in mind Quine’s definition of stimulus meaning, let us return to our field linguist engaging in radical translation.

An important qualification of the above description needs to be made. Quine claimed that it could be argued that ‘Gavagai’ and ‘Rabbit’ shared the same stimulus meaning. However, because of what he calls intrusive information, such strict equation of ‘Gavagai’ with ‘Rabbit’ is not necessarily justified. This is because what counts as the stimulus meaning of ‘Gavagai’ will vary depending on collateral information. For example, a native may assent to ‘Gavagai’ when there is movement in the grass in an area where rabbits are known to play. This will then count as part of the stimulus meaning of ‘Gavagai’. However, to the linguist who does not have this collateral information, this movement would not correspond with his stimulus meaning of ‘Rabbit’. So we would have to say that the two terms have similar but not identical stimulus meanings. In raising this difficulty, Quine is alerting us to the fact that observational sentences such as ‘Gavagai’ get their meaning partly as a result of collateral information resulting from the natives’ overall theory of the world. This is, of course, a reiteration of his holistic theory of meaning which he argued for in “Two Dogmas”.

In order to do a good translation, Quine claimed that we need to discover what sentences of the language count as occasion and standing sentences, and even more

importantly, whether the sentences are observation sentences. An occasion sentence is a sentence which commands assent and dissent only if queried after prompting of appropriate stimulation. Quine counts 'Red', 'Gavagai' and 'It hurts' as occasion sentences. A standing sentence is a sentence which will command assent or dissent independently of whether there is a prompting stimuli. Obviously, for the purposes of translating natural language, occasion sentences will be the primary evidence which the linguist will have to go on.

The easiest type of occasion sentences to translate will be the ones which are what Quine calls observation sentences. He notes that observationality is social. A term counts as observational if most people assent to it given the same stimuli. There are grades of observationality which go from 'Red'- 'Rabbit'- 'Bachelor'. He grades 'Red' the most observational and 'Bachelor' the least observational. One of the main reasons for such a gradation between levels of observationality is that sentences which are more observational take less collateral information to understand them. Occasion sentences like 'Bachelor' require so much collateral information to understand them that we would not even count them as being observation sentences.

For Quine, the notion of observation is closely connected to the notion of *the modulus*. The modulus is what Quine calls a working standard of the specious present (*ibid.*, 28). Sometimes if we increase our modulus, we will be able to increase our notion of observationality. This can be seen if we think back to our native assenting to 'Gavagai' in an area where rabbits are normally seen. If we increased our modulus to a long enough time frame to include past stimulations of rabbits in that area, their notion 'Gavagai' would be similar to our 'Rabbit'; hence observationality would increase. Quine claimed that such an increase in modulus would increase the observationality even of terms such as 'Bachelor'.

One of the key points that Quine is trying to dislodge in his IDT argument is the view that when a translator constructs a correct translation what he uncovers is the meaning of the term being translated. He wants to externalise this process and try to key language to stimulations which prompt the verbal behaviour of the subject. A correct translation of an observation sentence will occur when one can say that the two observation sentences share the same stimulus meaning, though because of intrusive information, this convergence of stimulus meaning cannot be total, as we mentioned above. We could improve our situation if we increased our modulus; however, if we consider a case like 'bachelor' a further difficulty occurs which cannot be overcome by increasing the modulus.

Typically the stimulus meaning of 'Bachelor' and 'Unmarried man' will be the same for each person so we can say that they are stimulus synonymous. However, the stimulations which will prompt assent to 'Bachelor' will differ for each person. But we can check their objectivity by noting the ease of communication which occurs between our speaker, for whom 'Bachelor' and 'Unmarried man' are stimulus synonymous, and the rest of our community when they use these terms. The beauty of intrasubjective stimulus synonymy is that it works for bilingual speakers as well. Quine notes that 'Solerto' and 'Bachelor' will have the same stimulus meaning for a bilingual speaker. So he argues that if our linguist wants to learn non-observation sentences such as 'bachelor', he can do so by learning the foreign language like a child does and later map the words of his new language onto his old language. The approach of learning the foreign language in the same manner as a child does is one which cannot be made sense of until we have a worked out theory of how our child learns his first language. Quine sketches a theory of how a child learns his first language in Chapter 3 of *Word and Object*; I will evaluate how his theory stands up to

empirical research done over the last twenty years throughout the rest of this thesis. The empirical evidence at this stage does not as of yet conclusively support his conception of how a child learns its first language nor does it refute it. So Quine's claim that a linguist can learn non-observational terms like the child learns them is senseless until such time as we have enough empirical evidence to show how children do in fact learn non-observation sentences.

This difficulty that I have sketched is not one which would overly bother Quine because he claims that there is a behavioural way in which we can test which sentences are stimulus synonymous for which people. His approach is as follows:

He need merely query the sentences in parallel under random stimulations until he either hits a stimulation that prompts assent or dissent to one sentence but not to the other, or else is satisfied that he is not going to. A visiting Martian who never learns under what circumstances to apply 'Bachelor' or 'Unmarried man' either, can still find out that 'Bachelor' for one English speaker does not have the same stimulus meaning as 'Bachelor' for a different English speaker and that it has the same as 'Unmarried man' for the same speaker. He can, anyway, apart from one difficulty: there is no evident reason why it should occur to him thus blindly to try comparing 'Unmarried man' and 'Bachelor' (Ibid., 47).

However, he warns that even if we could construct such behavioural tests and hence discover behavioural evidence for intrasubjective stimulus synonymy, such tests would not free us from the spectre of intrusive information. One example which he thinks illustrates this point: is that of a Himalayan explorer who learns to apply 'Everest' to a distant mountain seen from Tibet and 'Gaurisanker' to one seen from Nepal. Quine asks us to imagine that later the peaks are discovered to be identical. This empirical discovery would mean that the stimulus meaning of 'Everest' and 'Gaurisanker' would forever coincide for the speaker. The point in this case is that, as a result of an empirical discovery, two words which had different stimulus meaning can end up having the same stimulus meaning. So even in the case of intrasubjective stimulus synonymy, collateral information is extremely relevant to fixing the stimulus

meaning. This is extremely important because it shows that the meaning of a word is not divorced from beliefs about the world. The two are deeply connected. It reinforces Quine's view, sketched in Chapter 1, that the idea of analytic sentences which are held true purely because of meaning, independent of collateral information is senseless.

Thus far we have been speaking of the stimulus meaning of observation sentences, but further differences emerge when we try to move from observation sentences such as 'Gavagai' to 'gavagai' treated as a term. Quine stressed throughout *Word and Object* that he believes that words are learned by abstracting them out of observation sentences. He claimed that it would be a mistake to assume that because we can cash out observation sentences in terms of stimulus meanings, we can do the same for terms. As he put it:

Stimulus synonymy of the occasion sentences 'Gavagai' and 'Rabbit' does not even guarantee that 'gavagai' and 'rabbit' are coextensive terms, terms true of the same things. For consider 'Gavagai' Who knows but what the objects to which this term applies are not rabbits after all, but mere stages, or brief temporal segments, of rabbits? (ibid., 53)

This point about the reference of terms is the most famous part of Quine's IDT argument. It is called *The Inscrutability of Reference*. It is the inscrutability of reference that thinkers influenced by Chomsky have focused on as an instance of UD. Given that the stimulus meaning of 'Gavagai' typically occurs in situations where a rabbit is present it is natural for us to think that the term 'gavagai' refers to rabbits. However, Quine correctly points out that we only make the above assumption because we assume that the native is enough like us to have a short term for 'rabbit' and not one for 'undetached rabbit part', 'particular instance of universal rabbit hood', or 'it rabbiteith' (a feature placing term like it rains). We use what Quine calls analytic hypotheses to break down observation sentences into their component words. For

Quine, these interpretations of terms amount to nothing more than imposing our own ontology onto the natives. The stimulus meaning of 'Gavagai' leaves it underdetermined whether 'gavagai' refers to a rabbit or to the other possibilities which I have mentioned. When discussing analytical hypotheses Chomsky makes the following point:

There can surely be no doubt that Quine's statement about analytical hypotheses is true, though the question arises why it is important. It is to be sure undeniable, that if a system of "analytical hypothesis" goes beyond the evidence then it is possible to conceive alternatives compatible with the evidence... this situation in the case of language, or "commonsense knowledge", is, in this respect no different than the case of physics. (1969, 144)

Here we see Chomsky comparing the forming of analytic hypothesis to translate the terms of a native's language with constructing an underdetermined theory in physics. We saw in the chapter above that Chomsky's solution to underdetermination in physics is to postulate innate apparatus to overcome the UD. In fact, it was shown that Chomsky's solution to UD in all areas is to postulate innate apparatus to overcome the problem. Since he views the IDT argument as a form of UD he argues that we overcome the IDT because of our innate language faculty. Our analytical hypotheses where we speculate that the term 'gavagai' refers to rabbit as opposed to 'particular instance of universal rabbit hood' etc are limited by our innate concepts, which we share with all members of our species. The categorisations that the native as well as the linguist makes will be subject to the same constraints forced on us by our innate language faculty. The innate constraints on interpretation will also determine that children will interpret the meaning of their peers' words using their language faculty and so will avoid making the possible interpretations which Quine talked about in his IDT argument.

PART2: INDETERMINACY OF TRANSLATION AND UNDERDETERMINATION

Most contemporary cognitive scientists have followed Chomsky in arguing that the IDT amounts to nothing more than UD and that this UD can be solved by postulating innate concepts. In particular they are interested in problems of inscrutability of reference and argue that the fact that children do not fall foul of the inscrutability of reference can be explained by the fact that children are born with innate concepts.

In his book *The Language Instinct*, Stephen Pinker presents the inscrutability of reference as a form of UD:

A rabbit scurries by, and a native shouts, 'Gavagai' What does gavagai mean? Logically speaking, it needn't be 'Rabbit'. It could refer to that particular rabbit (Flopsy, for example). It could mean any furry thing, any mammal, or any member of that species of rabbit (say, *Oryctolagus*), or any other member of that variety of that species (say, chinchilla rabbit). It could mean scurrying rabbit, scurrying thing, rabbit plus the ground it scurries upon, or scurrying in general. It could mean footprint-maker, or habitat for rabbit fleas. It could mean the top half of a rabbit, or rabbit meat on the hoof, or possessor of at least one rabbit's foot. It could mean anything that is either a rabbit or a Buick. It could mean collection of undetached rabbit parts, or 'Lo Rabbit hood again, or 'It rabbiteth', analogous to 'it raineth'. The problem is the same when the child is the linguist and the parents are the natives...Figuring out which word to attach to which concept is the Gavagai problem, and if infants start out with concepts corresponding to the kind of meanings that languages use, the problem is partly solved. (1994, 156)

Pinker then goes on to cite the research of the psychologist Ellen Markman to show that it is indeed the case that children are born with concepts which constrain the types of meanings that words can be given. It is Pinker's view that Quine's problem is a problem of UD which is solved by the fact that the concepts that children are born with will place innate constraints on the type of meanings that children can attach to words.

The child psychologist Paul Bloom draws a similar conclusion to Pinker:

These problems of reference and generalization are solved so easily by children and adults that it takes philosophers like Quine and Goodman to even notice that they exist. If we see someone point to a rabbit and say 'gavagai', it is

entirely natural to assume that this is an act of naming and that the word refers to the rabbit and should be extended to other rabbits. It would be mad to think that the word refers to undetached rabbit parts or rabbits plus the Eiffel Tower. But the naturalness of the rabbit hypothesis and the madness of the alternatives is not a logical necessity; it is instead the result of how the human mind works. (2000, 5)²⁸.

The Harvard linguist Cedric Boeckx echoes the claims of the above thinkers:

Yet, if you think about it, the Gavagai situation is the one we all faced as children trying to acquire the meaning of words. How did we guess that elephant refers to that big grey animal with a long trunk? Because someone pointed to the animal and said elephant? But how did you know what exactly was being pointed at? Surely the finger couldn't point at the whole elephant; it was your cognitive bias that interpreted the act of pointing in that way. (2010, 41)

So the above quotes clearly indicate that, within the realm of cognitive science, it is common-place to follow Chomsky and view the inscrutability of reference as a problem of UD, and indeed as a problem which can be solved by postulating innate concepts. Pinker, Bloom and Boeckx are speaking only of the inscrutability of reference as it affects the child learning his first words. Chomsky goes further: he is arguing that the IDT is a form of UD which can be overcome by postulating innate concepts. He does not explicitly demonstrate how the IDT is overcome by innate concepts; rather, he merely points out that the situation in physics raises the same problems as the IDT problem. In the case of physics the problem is overcome by a science forming faculty, whereas in the case of the IDT the problem is overcome by our language faculty. His reasoning is that is that our innate concepts will ensure that when we are acquiring our first language, we will not think that a word like 'gavagai' refers to undetached rabbit part, etc. Rather, we will discover that the word merely labels our innately known concept of 'rabbit'. The fact that our innate concepts will determine that our words have fixed meanings means that the IDT is incorrect in claiming that there are no facts about translation.

²⁸ While Pinker's and Bloom's arguments on this narrow point are similar, obviously this does not imply that they hold the exact same view on the nature of the mind.

It could be argued that by postulating such innate concepts, these thinkers are merely begging the question against Quine by assuming the very determinacy of meaning which he is arguing against. However, Pinker et al. argue that they are not merely postulating innate concepts in order to overcome indeterminacy. Rather, they claim that there is independent evidence which supports their belief in innate concepts, and this evidence offers reasons to believe that UD will not be something that affects a young child who is learning his first language.

PART 3: INDETERMINACY OF TRANSLATION AND EXPERIMENTAL RESEARCH

SECTION 1: EVIDENCE FOR INNATE CONCEPTS OF BODY AND MIND

Empirical research in psychology indicates that children as young as four months already display conceptual knowledge. Jean Piaget and Vygotsky have done some of the most detailed studies of young children's interactions with other people and objects as they learn to speak. More recently, psychologists have been trying to build on and expand on the work of these great pioneers. The work of Elizabeth Spelke and Susan Carey stands out as the best of this new wave of research on child cognition. Jean Piaget's experimental research led him to believe that young children at the age of five months do not have an adequate knowledge of objects and that they have to pass through a variety of developmental stages before they will acquire the full concept of an object. Piaget noted that children playing with a ball will not search for it if it runs out of their field of vision. He drew the conclusion that they do not have knowledge of object permanence at this age. As early as the seventies, psychologists such as Bower were questioning Piaget's view by claiming that these babies do not search for the missing objects because of a problem of coordinating their movement. And since the seventies, it has been common-place to try to test children's concept of

an object by using experiments that do not require coordinated sequences of actions.

Bower's studies have made surprising discoveries: (Here I am citing from Baillargeon, R., Spelke, E, and Wasserman's paper "Object permanence in infants"):

Bower's studies have yielded four findings that seem to have provided evidence for object permanence in infants well below 9 months. First, 7 week old Infants were found to discriminate between disappearances that signalled the continued existence of an object (e.g. Gradual Occlusion), and disappearances that did not (e.g. gradual dissolution or sudden implosion). Second, 2 month old infants were found to anticipate the reappearance of an object that stopped behind the screen, 'looking to that half of the movement path the object would have reached had it not stopped' Third, five month old infants were found to show disruptions in their tracking when an object was altered while passing behind the screen: they tended to look back at the screen, as though in search of the original object. Finally, 5-month-old-infants were found to reach for an object that had been hidden by darkening the room. (1985, 195)

However, Bower's experiments are not conclusive evidence that children are born with concepts of object permanence. Piaget himself claimed in a personal correspondence with Bower that he could interpret some of these experiments in a way which was consistent with his own theory (ibid., 97). Because of the inconclusiveness of the debate between Bower and Piaget,²⁹ Baillargeon et al. set up more rigorous tests to show that humans are born with concepts of object permanence. Quine, like Piaget, also argued that Bower's experiments did not necessarily show that a 4-month-old child has a concept of an object:

True, an infant is observed to expect a steadily moving object to reappear after it passes behind a screen; but all this happens within the specious present, and reflects rather the expectation of continuity of a present feature than the reification of an intermittently absent object. (1969, 24)

Here Quine is interpreting Bower's famous experiment to mean not that the young child has a concept of an object, but as indicating that the young child expects

²⁹ Piaget is not only a strong proponent of the view that the empirical evidence for innate concepts is weak; he also refuses to grant any significance to Chomskies' poverty of stimulus argument. So Piaget is, in effect, one of the strongest opponents of both the logical and empirical argument for innate concepts. Chomsky and Piaget had a public debate in the 1970 and this debate is recorded in the Book: Piattelli-Palmarini, M EDS *Language and Learning: The debate between Jean Piaget and Noam Chomsky*.

continuity in his experiences. Baillargeon et al.'s experiment aimed to test this view of Quine and Piaget's.

They began by testing whether children have what they called The Solidity Principle, which was the principle that a solid object cannot pass through another solid object. An experiment was set up where an object was presented to the child and then occluded. Then another object was rolled towards the occluded object. If the child exhibited surprise that the object moved through the space which seemed to be occupied by first the object, then this would exhibit two things: first, that the child had grasped the Solidity Principle; second, that the child understood object permanence, because he knew the objects still existed when occluded. Baillargeon et al performed this experiment on four month old children. An unargued assumption of the experiment was that the child was exhibiting surprise when he stared at the event longer than when it did not violate the solidity principle and the object permanence principle. When completed the experiment showed that four-month-old infants do seem to reason according to the principle of solidity, and the principle of object permanence.

Quine explained Bower's experimental results by saying that the child expected the object to return out the other side of the screen because of an expectation of continuity. The child saw the object-shaped thing move towards and behind a screen. So he expected this object-shaped thing to continue moving in the same direction as it started from; hence, the surprise when it did not move out the other side of the screen. For Quine this in no way indicates that our child has a sophisticated concept of an object. However, expectation of continuity does not explain what is going on in Baillargeon et al.'s experiment. The child is exhibiting two beliefs: first, the belief that object-shaped things do not disappear when occluded; second, that

object-shaped things are solid, and hence cannot be moved through. Hence the surprise the child exhibits when the object appears on the other side of the screen as though it had passed through the original object. Obviously, expectation of continuity does not explain what is going on in Baillargeon et al's experiment. So to account for both Bower and Baillargeon's experimental results, Quine needs to postulate at a minimum that 4 month old children expect object-shaped things not to disappear when occluded, to move in undeviating paths unless stopped by another object-shaped blob, and to have solidity. A child with knowledge of this sort seems to be clearly demonstrating a complex concept of an object, and a theory which, contrary to what Quine believes, precedes the learning of the syntax of quantification by about 2 years.

However, one might be wondering at this point how this experiment is relevant to the IDT. The relevance of the experiment should be clear if we consider some parallels which the radical translator encounters and which the child learning his first words encounters. As we saw above, Quine argues that 'Gavagai' considered as an observation sentence has a clear stimulus meaning and can be correlated with non-verbal stimulation. The same situation occurs when the child says 'Mama'; this is an observation sentence which has a clear stimulus meaning. According to Quine, the word is correlated with a scattered portion of what is going on in the child's environment when he speaks. For Quine, the difficulty arises when we try to break down observation sentences into terms and predicates etc. When we tried to break down 'Gavagai', the observation sentence, into 'gavagai', the term, our problems begin. Our breaking down the observation sentence involves us using analytical hypotheses which go far beyond the empirical data. Quine notes that no behavioural evidence can decide whether 'gavagai' refers to 'undetached rabbit part', 'particular instance of universal rabbithood', 'rabbit fusion' etc. If we want to translate the words

of the native we must unjustifiably impose our ontology onto the natives. The situation of the child is similar, though there are some notable differences. When the linguist is trying to translate terms like 'gavagai' he ends up imposing his own ontology onto the native; however, according to Quine, the child has no ontology. So when the child tries to move from using 'Mama' as an observation sentence to 'mama' as a term he will have little to go on. The child cannot impose his own ontology because he supposedly has none. As the child begins to gradually learn the apparatus of quantification and syntax, through induction, reinforcement etc, then he will gradually acquire the language of his peers. However, if Baillargeon et al.'s experiment is correct, then when children are learning to speak, their innate concept of object will determine that the underdetermination facing the child learning his first language will be massively reduced.

There is also experimental evidence of children having theories of mind (Baron-Cohen), as well as the theory of objects which Baillargeon et al discovered. These empirical discoveries indicate that prior to learning their first language, the child has an ontology consisting of objects, agents, and causality etc. which will partly determine how the child interprets the speech of the adults they engage with. So it is typically assumed that the child's innate concepts of objects and agency will massively limit the underdetermination facing the child as he learns his first language. However, while suggestive, this supposed evidence for innate concepts is derived from the study of pre-linguistic children so it does not directly deal with issues of the reference of words. In the next section I will discuss an experiment which was done on children who are in the early stages of learning their language. The experiment was designed to test whether Quine's claim that children only develop an ontology after they have mastered the syntax of quantification is true.

SECTION 2: ONTOLOGY AND LANGUAGE LEARNING: AN EXPERIMENTAL TEST

The experimental research that I outlined above purported to demonstrate that children have innate concepts of object and agency. If the evidence is taken at face value, it implies that these innate concepts will ensure that the UD facing the child learning his first words will be decreased. However, this argument runs contrary to Quine's claim that we are only justified in attributing an ontology to children after they have mastered the syntax of quantification. In this section, I will consider an experiment constructed by Soja, Carey, and Spelke which aims to test Quine's claim.

In their 1990 paper "Ontological categories guide young children's inductions of word meaning"³⁰ Soja, Carey, and Spelke tested whether the ontological distinction between objects and non-solid substances conditions the projection of word meanings prior to the child's mastery of count/mass syntax. Quine denied that children make any ontological commitments prior to learning the syntax of quantification which helps them master the count/mass noun distinction. Here Quine is making an empirical claim that prior to grasping the count/mass distinction, an agent like Mother, a property like Red, and a non-solid substance like Water are on a par. According to Quine, it is only when a child has mastered the apparatus of divided reference through grasping the syntax of quantification that the child can distinguish these substances. Soja et al. set out to test these empirical claims of Quine's. Before outlining their experiment I outline Quine's views on language acquisition.

QUINE'S POSITION

Quine claims that when young children mouth words such as 'Mama', 'Water', or 'Red', we are in no position to state that they are using the words as terms which refer to the same things which we refer to by the sounds.

For though we may fully satisfy ourselves that the child has learned the trick of using the utterances 'mama', and 'water' strictly in the appropriate presences, or as a means of inducing the appropriate presences, we still have no right to construe these utterances in the child's mouth as terms, at first, for things or substances. (1969, 7)

According to Quine, from our own mature perspective, we have come to view the child's mother as a body which revisits the child from time to time, and water as a scattered object. However, from a behavioural perspective, we have little justification for imputing this ontology onto the child. After making this negative point about our lack of justification for imputing our mature ontology onto the child, he then goes on to make a positive point about the nature of the child's ontology.

But the mother, red, and water are for the infant all of a type: each is just a history of sporadic encounter, a scattered portion of what goes on. His first learning of the three words is uniformly a matter of learning how much what goes on about him counts as the mother, or as red, or as water. It is not for the child to say in the first case 'Hello mama again', in the second case 'Hello another red thing' and in the third place 'Hello more water'. They are all on a par: Hello more mama, more red, and more water. (*ibid.*, 7)

Here Quine is clearly claiming that young children who use words, such as 'Mama', 'Red' and 'Water' are not distinguishing them in terms of being respectively Objects, Properties, and Non-solid substances. His reason for arguing so is that we have no positive behavioural evidence to support the claim that children make such distinctions, and in the absence of such positive evidence, there is little reason to impute such a rich ontology to young children. Quine's claim is that we should only attribute to children the ability to distinguish between Objects, Substances and

Properties when we have behavioural evidence which supports us making this distinction.

Progressively, however, the child is seen to evolve a pattern of verbal behaviour that finally comes to copy ours too closely for there to be any sense in questioning the general sameness of conceptual scheme. For perspective on our own objectifying apparatus we may consider what steps of development make the difference between 'mama'-babbling infant who cannot be said to be using terms for objects, and the older child who can. It is only when the child has got on to the full and proper use of individuating terms like 'apple' that he can properly be said to have taken to using terms, and speaking of objects. Words like 'apple' are not words like 'mama' or 'water' or 'red' are terms whose ontological involvement runs deep. To learn 'apple' it is not enough to have learned how much of what goes on around you counts as apple; we must learn how much counts as an apple, and how much as another. Such terms possess built in modes of individuation. (ibid., 8)

Now Quine acknowledges that the child may learn 'apple' in the same way that he learns 'mama' or 'red' but he goes on to say that the child will never master 'apple' in its individuating use until he gets on with the scheme of enduring physical objects. And in order to get on with the scheme of enduring physical objects, the child will need to master the apparatus of identity, difference etc. Quine claims that to be able to tell if the child has got the trick of individuation down, we need the following:

How can we ever tell if the child has got the trick of individuation? Only by engaging him in sophisticated discourse of 'that apple', 'not that apple', 'an apple', 'same apple', 'these apples', 'another apple'. It is only at this level that a palpable difference emerges between genuinely individuating use and the counterfeits lately imagined. (ibid., 9)

It is at this stage that Quine claims that we are justified in attributing an ontology to the child. Prior to that, attributing an ontology to the child is making an unsupported conjecture which is not justified by the facts. For Quine, our child learns the adjectives 'same', 'another', 'an', 'that', 'not that' contextually. First the child gets used to various longer phrases which contain them, and he gradually develops appropriate habits in relation to the component words as common parts and residues of those longer forms. He further speculates that the contextual learning of all of these

various different particles goes on simultaneously, so that we gradually adjust them to each other as a coherent pattern of usage is evolved (*ibid.*, 10).

So the story of child ontology as Quine tells it is that the young child, as he learns to mouth words in intersubjectively appropriate circumstances, does not yet have an ontology. The child's words just represent scattered portions of what goes on and do not distinguish between Objects, Properties and Substances. At this stage of the child's development the child uses an innate similarity quality space to categorise his experiences:

If the child is to be amenable to such training, however, what he must have is a prior tendency to weight qualitative differences unequally. He must, so to speak, sense more resemblance between some stimulations than between others. Otherwise a dozen reinforcements of his response 'Red', on occasions where red things were presented, would no more encourage the same response to a thirteenth red thing than to a blue one; and a dozen reinforcements of his response 'Mama', on occasions dominated by the mother's face at various angles, would be just as inconsequential. In effect therefore we must credit the child with a sort of pre-linguistic quality space. We may estimate relative distances in his quality space by observing how he learns. (1960, 83).

So Quine's picture of a child learning language and the ontology which is implicit in this language involves pared down assumptions according to which we attribute to the child no more than is necessary to explain his verbal behaviour. Quine treats the babbling which a child begins to emit at the age of 12 months as a form of operant behaviour which is omitted rather than elicited. He claims that the family of the child will reinforce the child's verbal behaviour (such reinforcement made possible by the child's pre-linguistic quality space) in such a manner that the child's use of observation sentences such as 'mama' will reliably distinguish between 'mama' portions of the environment, and 'non-mama' portions. However, at this stage we cannot credit the child with having an ontology; from the point of view of external verbal behaviour we have no reason to attribute to the child a concept of 'Mama' as a name of a spatio-temporal object, as opposed to being a name of a mere mass term

like 'Water'. It is only when we engage the child in discourse and he can answer the questions using terms such as 'not that mama', 'same mama', 'another mama' etc., that we are justified in attributing to the child a concept of 'mama' as an object as opposed to scattered portion of mama environment etc.

TESTING QUINE'S CLAIMS

Soja et al. conducted their experiments to test Quine's claim (1960, 1969) that young children only develop an ontology after they have grasped the syntax of quantification. Contrary to Quine, they claimed that young children have a distinction between different ontological categories prior to grasping the syntax of quantification, and that in fact these ontological categories constrain the process of language learning. They distinguished their views from Quine's in the following way:

According to Quine, then, when children hear a new word, the meaning they assign to it is determined by procedure 0:

Procedure 0: Conclude that the word refers to aspects of the world that share salient properties of the perceptual experience when the word is used. (1991, 182)

Soja et al. proposed a different view of the procedures children use when they learn a new word; their procedure assumed that the child had ontological categories prior to learning the syntax of quantification.

Procedure 1 Step 1: Test to see if the speaker could be talking about a solid object; if yes,

Step 2: Conclude the word refers to individual whole objects of the same type as the referent.

Procedure 2 Step 1: Test to see if the speaker could be talking about a non-solid substance; if yes,

Step 2: Conclude the word refers to portions of substance of the same type as the referent. (ibid., 183)

Soja et al. proposed an experiment which would decide between these two different proposals about how children learn new words.

One way for Soja et al. to test whether Quine was correct, or whether they were correct, was to test how children generalised when they learned words for

different objects. If children could generalize prior to a grasp of the count mass syntax this would be evidence that Quine was wrong. Markman and Watchel constructed an experiment to test this possibility. In that experiment, children were taught a new word in the presence of an unnamed object, and were then asked to find more objects of this kind. It was found that children generalized to what adults would consider whole individuated objects, as opposed to perceptually similar entities which were not objects of the same type. For example, children generalized a word that was initially applied to one object to a new object of the same shape as opposed to a new object of a different shape but the same material, and in preference to a new object consisting only of a part of the original object. Soja et al. have claimed that this experiment does not distinguish between Quine's view and their conjecture for two reasons: (1) the experiment was done on children aged 3 ½, and children of this age have already mastered the syntax of quantification³¹. (2) The experiment can be made consistent with both Quine's and Soja's views. It may seem that the fact that the child generalizes based on shape as opposed to substance shows that they are following Procedure 1, namely: check whether the data is referring to an object if so conclude the word refers to individual whole objects of the same type as the referent. However nothing in the experiment rules out the interpretation that the child is applying the word using Procedure 0: Conclude that the word refers to aspects of the world that share salient properties of the perceptual experience when the word is used. For example, Landau et al. have made this Quinean interpretation of the experiment, claiming that it shows that the child generalizes his first nouns based on shape, for example, book-shaped, clock-shaped etc. (ibid p.184).

³¹ Soja et al. cite the experimental work of Gordon P. "The acquisition of syntactic categories: The case of the count/mass distinction" as evidence that children prior to the age of 2½ do not have knowledge of a count/mass distinction

Now obviously, to avoid this problem one needs to put certain measures in place. First, one needs to ensure that the experiment is done on children who are below the age of 2 ½, the age at which children master the syntax of quantification. Second, one needs to test how children generalise words to non-solid substances as well as to objects. In the experiment of Markman and Washoe, some people interpreted the experiment by claiming that children were generalising words according to shape using some sort of innate perceptual quality space. If people are right that children generalise names by using Procedure 0, then children will generalize names based on shape whether the name originally refers to an object or a non-solid substance.

THE EXPERIMENT

Twenty-four 2-year-olds from the Greater Boston area were recruited and randomly placed into two groups (informative syntax groups, and neutral syntax groups), with equal numbers of boys and girls in each group. Each testing session began with two familiar trials: one object trial and one substance trial. The stimuli in the familiar object trial were a blue plastic cup, a white Styrofoam cup and cup pieces. The stimuli in the familiar non-solid substance trial were peanut butter and Play-doh. These trials followed the same format as the unfamiliar trials described below. The two familiar trials were followed by eight unfamiliar trials: four object trials and four substance trials which were intermingled. The subjects were tested on each trial on two separate occasions. Eight novel words were used: 'blicket', 'stad', 'mell', 'Coodle', 'doff', 'tanninn', 'fitch', and 'tulver' (ibid., 187)

TEST 1: AN UNFAMILIAR OBJECT TRIAL IN THE NEUTRAL SYNTAX CONDITION

The test involved presenting the child with an unfamiliar object, e.g. a plumbing T-shaped pipe, and giving the child a name for the object, e.g. blicket. In the neutral syntax condition the child is told 'This is my blicket'. The experimenter then continued to talk about the object using 'my' 'the' and 'this' as determiners. She and the subject then manipulated the object. The object was then placed to the side and two other sets of objects were then presented to the subject. One set consisted of objects of the same sort as the original but made of a different material, e.g. a plastic T shape; the other set consisted of objects of the same material but a different shape i.e. bits of metal. The experimenter then said 'Point to the blicket'

TEST 2: AN UNFAMILIAR SUBSTANCE TRIAL IN THE NEUTRAL SYNTAX CONDITION

The child was shown an unfamiliar substance, and told 'This is my stad'. The experimenters referred to the substance using only the determiners 'my' 'the' and 'this'. The experimenter and the subject talked about the substance and played with it. In the presentation of test substances, the subject was shown two substances, the original and the new one, and told 'Point to the stad'. The original substance was in the alternative configuration, whereas the new substance was in the configuration used originally with the named substance. There were four pairs of substances: (1) Dippity-do (a setting gel), and lumpy Nivea (a hand cream mixed with gravel), (2) Coffee (freeze dried) and Orzo (a rice shaped pasta), (3) Sawdust and leather (cut to tiny pieces), (4) Crazy foam and Clay. Of each pair one member was named and the other was used as the alternative to the original in the test presentation. Each member served in both roles across subjects.

TEST 3: OBJECT AND SUBSTANCE TRIALS IN THE INFORMATIVE SYNTAX CONDITION

This condition differed from the neutral syntax condition only in the determiners and quantifiers used when naming the original stimulus. The experimenter introduced an object trial in the informative syntax condition with ‘This is a blicket’ and used ‘A blicket’ and ‘Another blicket’ in subsequent discussions. Substance trials in the informative syntax condition were introduced with ‘This is stad’ and in subsequent discussion the experimenter continued to omit determiners or use ‘some’ or ‘some more’. This was the only difference between the different informative and uninformative trials. In the familiar word trial subjects differentiated the object and the substance trials as predicted.

WORD LEARNING TRIALS

Subjects differentiated the two types of trials. Responses were consistent with shape and number on the object trials, and were not consistent with shape and number in the substance trials.

WHAT THE TEST SHOWS

If before the child has grasped the syntax of quantification the child differentiates in the above manner, this shows that the child is not generalizing the word-based perceptual similarity, but is doing so based on the type of object he is presented with. So, for example, if he was generalizing according to an innate perceptual similarity quality space which focuses on shape then why does this not work for substances? The answer is because the child recognises that objects and substances are distinct ontological categories. Soja et al. summed up their results as follows:

In sum, the children chose according to object type when the stimulus was an object and according to substance type when the stimulus was a non-solid substance. There was no effect of the syntactic context: performance was neither facilitated nor hindered by the additional syntactic information.

The data from Experiment 1 show that different inferences about the meaning of a newly heard word are drawn according to the ontological status of its referent. If the word refers to an object, the child's projection respects shape and number, and ignores texture, color, and substance. If the word refers to a non-solid substance, the child's projection ignores shape and number, respecting texture, color and substance. (ibid., 192)

From this experiment, Soja, et al. claim to have shown that Quine's view of how children learn language is incorrect because the experiment shows that, contrary to what Quine claims; children do indeed have a distinction between different ontological categories prior to grasping count/mass syntax. It also shows that these innate ontological categories are what help a child learn a language and not the apparatus of quantification. Soja et al.'s experiment purports to have shown that children learn words according to ontological distinctions which they exhibit knowledge of prior to learning a language. So their experiment, along with the experiment of Baillargeon et al., purports to show that Quine is wrong in his story about how children develop their concept of an object.

Again, the relevance of this experiment to the IDT should be apparent. While Soja et al. purport to have shown that children have an ontology prior to learning their first language, Baillargeon et al. purport to have shown that young children interpret the meaning of new words using this pre-linguistic ontology. Quine claims that linguists construct their analytical hypotheses by imposing their ontology onto natives. These two experiments try to demonstrate that a child interprets the sounds of his peers by imposing his own innate concepts. So according to this view, there are facts about meaning which are imposed by our universally shared concepts.

SECTION 3: THEORY OF MIND AND TRIANGULATION

Research by Baillargeon et al. which I discussed in Section 1 Part 3 of this chapter indicates that four-month-old children seem to have implicit knowledge of the following five principles about objects:

- (1) The principle of solidity.
- (2) The principle that objects are three-dimensional.
- (3) Objects tend to move on undeviating paths.
- (4) Objects move continuously through space and time.
- (5) Objects only move when contacted by another object. (Children make exceptions to this principle when they are viewing intentional objects)

This work of Baillargeon et al. indicates that children from a very young age have an intricate knowledge of objects; on the other side of the coin the work of people like Simon Baron-Cohen³² indicates that prior to learning language, normal children view other people as intentional agents with beliefs, desires etc. Furthermore, there is evidence that such children who lack this theory of mind suffer severe language learning difficulties.³³ The evidence for a pre-linguistic theory of mind has been summed up succinctly by the psychologist Paul Bloom:

What understanding do prelinguistic children have about the minds of others? Consider first sensitivity to what other people are attending to. By around nine months, a baby will naturally follow its mother's line of regard (Butterworth, 1991), and will also follow her pointing gestures (Murphy and Messer 1977), at about the same age, babies can monitor their parents' emotional reactions to potentially dangerous situations and react accordingly. For instance when seeing a spider, a baby will be less likely to approach it, if its mother seems fearful than if she seems happy (Zarbatany and Lamb 1985) and when babies are uncertain or hesitant, they check what their mother is looking at and how she is reacting (Bretherton, 1992)... (2000, 67)

Bloom then wonders whether the gaze following is really an indication that the child has an implicit theory of mind, as Baron-Cohen believes it does, or whether the gaze is just an automatic orientating device that has nothing to do with intentional

³² See Simon Baron Cohen: *Mindblindness: An Essay on Autism and Theory of Mind*.

³³ See for example Cohen's book *Mindblindness*, and Kuhl, P. "Is speech learning gated by the social brain?" *Developmental Science* 10 pp 110-120

attribution. And he cites an experiment which seems to indicate that Baron-Cohen was indeed correct:

One way to address this question is to ask what sort of stimuli elicit gaze following in babies. A study with 12-month-olds by Johnson, Slaughter, and Carey (in press) reports an intriguing finding. When exposed to a robot who acts contingently with them through beeping and light flashing, but that has no face, babies will nonetheless follow its gaze (the orientation of the front, reactive part of the robot) treating it as if it were a person. But they do not do so if a faceless robot fails to interact with them in a meaningful way. This suggests that gaze following is applied to entities that give some sign of having intentional states, regardless of their appearance, and supports the view that gaze following is related, at least for twelve-month-olds to intentional attribution. (ibid, 62).

There is, then, some compelling evidence from both a logical and an empirical point of view which seems to indicate that children have pre-linguistic concepts which they use to help them learn their language. Furthermore, it is clear from the empirical studies that children exhibit knowledge of intentional agents and object constancy prior to their learning their first words at the age of 12 months. Furthermore, the child Psychologist Ellen M. Markman has provided evidence that when children do begin to learn words, the interpretations they give to the sounds they hear are fixed within certain parameters.

This evidence so far indicates the following: prior to the age of 12 months when young children begin to learn their first words, they exhibit a grasp of concepts of agency, object, substance etc. Young children also do not need to grasp the syntax of quantification to have ontological commitments; in fact there is evidence that young children (below the age of 2 ½) have an ontology which constrains the meanings that they can give to words. This evidence is typically used to argue that the inscrutability is overcome by our innate concepts. I will now discuss a final piece of evidence which was discovered by the psychologist Ellen Markman which purports to show how the pre-linguistic ontological constitution of the child will ensure that the IDT is overcome.

SECTION 4: MARKMAN AND WORD LEARNING

Markman begins her book *Categorisation and Naming in Children: Problems of Induction* (1989) by discussing Quine's IDT problem and claims that the young child is faced with the same problem as the adult linguist. Her empirical research is conducted to the end of seeing how the child does in fact overcome the problem. She begins by discussing the research of Piaget on young children, which indicates that young children categorise objects in terms of thematic relations (e.g. a spider and a web are categorised together because they are both part of the same theme: the spider building a web). Piaget drew the conclusion from his research that young children have no taxonomic categories (e.g. Car and Bus are put together because they are both vehicles) and categorise the world entirely in terms of thematic categories. Markman agreed with Piaget that children do indeed prefer thematic relations when categorising objects. However, she strongly disagreed with Piaget's claim that young children cannot categorise taxonomically and cites research from Susan Carey which indicates that, contrary to what Piaget suggests, young children are able to categorise objects taxonomically. Nonetheless, whatever the capabilities of children to use taxonomic categories, it is accepted by all researchers that children do prefer to categorise things according to thematic relations.

The fact that children prefer to organise their experiences according to thematic relations poses a serious problem for theorists who claim that innate constraints on interpretations will limit the problem of IDT. Markman summed up the problem as follows:

Children seem to readily learn terms that refer to object categories. Their vocabulary is filled with words such as 'ball' and 'dog', simple concrete nouns referring to object categories. Yet children often notice and remember thematic relations between objects more readily than an objects category. How is it that children readily learn labels for categories of objects if they are attending to relations to objects instead? To take a concrete example, imagine a mother

pointing to a baby and saying, 'baby'. Based on the classification studies, we should assume that the child will be attending to the baby sucking on the bottle, or to the baby being diapered. Why, then, doesn't the child infer that 'baby' means something like 'baby and its bottle' or 'baby and its diaper'? (1989, 27)

Therefore, if we accept the results of the categorisation tests, and if we further accept that the IDT is overcome by pre-linguistic biases, then it seems to follow that children will think that words like 'baby' will pick out thematic events like 'baby sucking on his bottle', and this view runs counter to our intuition of what words typically mean.

Markman proposed a solution to this problem. She claimed that when children are interpreting new sounds, they will typically assume that the new sound refers to a whole object. Markman then constructed a series of experiments which were designed to test this proposal. Markman's proposal is as follows:

Children may be biased to treat novel labels as referring to whole objects (the whole object assumption) and to treat them as referring to objects of the same type (the taxonomic assumption). (ibid p, 27)

Upon testing these assumptions the experimental evidence verified that these assumptions were indeed correct.

EXPERIMENTAL EVIDENCE

Markman's experiment was done on children between the age of two and three, and was designed to test whether children exhibit what she called the taxonomic and whole object assumption. Her prediction was that children should interpret novel labels as labels for objects of the same type, as opposed to objects that are thematically related. And to test this prediction, she organised a test which compared the way children organised objects when they were given a novel object label and the way they organised objects when they were given no label. There were two conditions in the experiment: a no word condition, and a word condition. In the two conditions the children were shown a target picture; they were then shown two other pictures and asked to select one of them as being the same as the target (ibid., 27).

NO WORD CONDITION:

The children were introduced to a hand puppet and told to put the picture they choose into the puppets mouth. The child was then shown a picture such as a picture of a dog, and then told ‘Look carefully. See this? Find one that is the same as this.’

NOVEL WORD CONDITION:

This condition was identical to the no word condition with one exception. In this condition the child was told that the puppet could talk in puppet talk. They were told to listen very carefully to find the right picture. The puppet gave the target picture an unfamiliar name and used the same name in the instructions for picking a choice picture. For example, the puppet might say, ‘See this? It is a sud. Find another sud that is the same as this sud.’(ibid., 28)

The results of the experiments in the two different conditions are instructive. In the no word condition, the child who had to choose between another category member and a thematically related object, often chose the thematically related object. They selected other category members a mean of only 59% of the time, which is no better than chance. In the novel word condition, the children selected categorically the vast majority of the time. They chose the other category member a mean of 83% of the time, which is better than chance.

So these results supported Markman’s prediction that when young children are learning the meaning of a new word they switch to taxonomic understanding of the meaning of the word. In other words, the child will assume that ‘Gavagai’ refers to rabbit as opposed to rabbit running in the grass.

However, this experiment is not directly relevant to the claims which Quine is making in *Word and Object*. When Quine is talking about the inscrutability of reference he typically makes the following point:

Point to a rabbit and you have pointed to a stage of a rabbit, to an integral part of a rabbit, to the rabbit fusion and to where rabbit-hood is manifested. Point to an integral part of a rabbit and you have pointed again to the remaining four sorts of things; and so on around. Nothing not distinguished in stimulus meaning itself is to be distinguished by pointing unless the pointing is accompanied by questions of identity and diversity. (1960, 53)

Markman's experiment indicates that children will make a whole object assumption in the new word condition 83% of the time. The relevance of this fact for Quine's concerns is questionable. If a child interprets new words using the taxonomic assumption instead of the thematic assumption, then this is an interesting fact about child psychology but it does not refute Quine's argument. If a child does make the taxonomic assumption for novel words he hears and continues to make this assumption when he uses the word, this will not overcome Quine's concerns. The child may not use 'Gavagai' to refer to rabbit running in the grass, however nothing in the experiment can tell us whether the child uses the term to refer to an undetached rabbit part or a particular instance of universal rabbithood. The point of Quine's thought experiment is that we can have no behavioural evidence which can indicate which of these interpretations is the correct one. We can try and find out if we engage the children or native in questioning, for example, are you speaking about a rabbit or an undetached rabbit part? However, in the case of the native we can only ask them when we have learned their language and hence have made assumptions about reference which will be hidden from experience. While in the case of child, we will only be able to ask them after they have learned our language and hence have adopted our ontology. So Markman's experiment does not really touch the issues which concern Quine. In Markman's experiment, it can be decided whether the child is making a thematic choice or a taxonomic choice depending on what objects they pick. However, Quine's example of 'gavagai' cannot be decided in such a manner because

every time the child picks out rabbit he will also be picking out an undetached rabbit part etc.

PART 4: EVALUATING THE RELEVANCE OF THE EXPERIMENTS

We have seen that a number of contemporary linguists and psychologists view the IDT as a form of UD which faces the child when he is trying to learn his first words. These theorists argue that innate constraints overcome this problem of UD. Merely postulating innate constraints to overcome the problem is question-begging, so the theorists need independent evidence to support their postulation of innate concepts. I have reviewed some different experiments which are typically cited as evidence to support the view that the UD facing the child is overcome because of innate constraints.

The first experiment indicates that children at the age of four months have a concept of what it is to be an object which includes beliefs that it is solid, three-dimensional, tends to move in undeviating paths, moves continuously through space and time, and only moves when contacted by another object. This complex concept of an object exists prior to a child learning language and is clear evidence that the child has ontological commitments prior to learning language and the syntax of quantification. The second experiment showed that before the age of two and a half when they have mastered the syntax of quantification, children could, contrary to what Quine claimed, distinguish between mass terms and count nouns. The experiment proved that when learning new words for solid substances children generalised based on shape and number, and when learning new words for non-solid substances generalised based on texture and colour. This experiment showed that children can distinguish between these substances prior to learning count nouns and

mass terms, and in fact this pre-linguistic ontology helps the child learn the distinction between mass and count nouns, not vice versa.

The third experiment showed that when children are learning a new word they are constrained to make a taxonomic assumption when deciding what the word refers to. This experiment is interesting in showing that when children are learning a new word they seem to be constrained in how they interpret the word. However, the experiment does not deal with the concerns of Quine, because even if the taxonomic assumption is in place, this still leaves the question open of whether 'gavagai' refers to undetached rabbit part, particular instance of universal rabbit-hood etc.

We saw above that thinkers like Pinker, Bloom and Boeckx follow Chomsky in thinking of the IDT as a form of UD which can be overcome by innate concepts. They typically cite the experimental work of Baillargeon, Spelke, Soja and Markman as evidence that children are born with the innate concepts which will ensure that children will overcome the problems raised by the IDT. However I argue now that these experiments are of limited significance when it comes to evaluating Quine's philosophical position.

When Quine speaks of ontological commitment only being possible when we have mastered the syntax of quantification he is primarily talking about an explicit worked out ontology. A child having implicit beliefs about the behaviour of objects is of limited philosophical significance. It is only when a creature begins to speak about this and that object, and to make claims which are true or false about objects, that we can really evaluate their ontic commitments. My primary concern is whether any of the three experiments offer evidence that the IDT viewed as UD is overcome by innate constraints.

The three experiments offer no evidence that the UD is overcome for children learning their first language because of innate constraints. Markman's experiment gives us no reason to believe that a child can decide between 'gavagai' in the sense of rabbit as opposed to undetached rabbit part, particular instance of universal rabbithood, or rabbit fusion etc. So her experiment is not relevant to the questions which Quine is raising when it comes to the Inscrutability of reference. It is typically believed that the evidence provided by people such as Baillargeon et al. etc. is evidence of innate concepts, and, indeed, the authors of these papers typically argue in this manner. However, the fact is that they only provide evidence for pre-linguistic concepts, not for innate concepts. So while they show that the UD of learning words is overcome by pre-linguistic concepts, the UD still remains for the pre-linguistic concepts, and to postulate innate concepts to overcome the UD is to merely beg the question against non-nativists.

So, for example, if children have pre-linguistic concepts which determine they will interpret 'rabbit' as referring to rabbit then by claiming that the concept *RABBIT* is innate we have solved the inscrutability of reference for a child learning his first words. His innate concept of *RABBIT* will determine that 'rabbit' will mean rabbit. Furthermore if our innate concepts determine what the words mean, then this will mean that our analytic hypotheses' are fixed from the start. However, the empirical evidence indicates that this argument does not work.

First, the experimental evidence which I reviewed above merely shows that children have pre-linguistic concepts; they do not demonstrate that children have innate concepts. If we assume that a child has a pre-linguistic concept of *GAVAGAI*, the question arises of how the child learns this concept. If the concept of *GAVAGAI* is supposed to refer to Rabbit as opposed to undetached rabbit part, particular instance

of universal rabbithood, or rabbit fusion, then one is faced with the question of what evidence the child used to decide amongst these different meanings of the concept? Quine's point was that there was no behavioural evidence which can decide between the various possible meanings of *GAVAGAI*. So if there is no evidence which the child can use to decide between the various different meanings of *GAVAGAI*, then Chomskians are faced with an APS for concept acquisition. If they want to argue for determinacy about the meaning of the concepts, they need to overcome the APS. A possible way to overcome the APS is to argue that children are born with innate concepts. However, to argue in this manner is to merely beg the question against Quine by assuming the very determinacy of meaning which is in question here.

To be clear, the issue here is not innateness per se, but innate concepts. Human intuitive physics differs from those of apes in some respects, and other creatures in many respects.³⁴ The differences between these creatures are obviously to do with their innate genetic endowment. However, nativists typically claim that certain intuitive theories of object behaviour, the behaviour of agents etc. is written into the genetic code of most if not all humans. And they claim that these theories will grow in the mind if children are situated in any normal environment. Non-nativists claim that humans differ from other animals because they are innately wired up to be able to recognise statistical regularities in their environment which will help them learn the structure of the world they are born into. The important point to note here is that the experiments of Soja et al. and of Baillargeon et al. do not tell us whether children learn their theory of object behaviour statistically or because they are born with innate concepts of objects etc. Quine argues that we learn our concepts through observation and trial and error. However he notes that this mode of learning will not determine

³⁴ See Tomasello: *The Cultural Origins of Human Cognition*, and Marc Hauser: *Wild Minds*.

what the meaning of *GAVAGAI* is. The experimenters who I reviewed above all argue that such concepts are innate; however, as we have seen, the evidence does not support their claims on this point. Without such evidence in place, they can simply state that innateness overcomes the UD but it remains just that, a statement, a question begging-assertion rather than an empirical fact.

To show that we are born with innate concepts as opposed to learning them statistically, an a-priori proof that such concepts cannot be learned statistically in a trial and error manner is needed. However the three experiments offer no such evidence. So they have not demonstrated that the UD facing the child is overcome by innate constraints.

It is also important to note that the types of experiments which they constructed are in principle blind to Quine's concerns. Quine is claiming that from a behavioural point of view there is no fact of the matter as to whether 'gavagai' refers to undetached rabbit part, particular instance of universal rabbithood, rabbit fusion etc. None of the experiments above can tell what 'gavagai' refers to, nor can any behavioural evidence in principle, because, as Quine says, every time we point to one of them we are pointing to the others as well. So the experiments do not really affect Quine's IDT argument.

The situation is this: one of the premises of the IDT, the inscrutability of reference, says that there is no fact of the matter as to whether 'gavagai' refers to undetached rabbit part, particular instance of universal rabbithood, it rabbiteth, etc. One can avoid this conclusion by saying that we know intuitively that there is a distinction between the various potential references of the term, and we know this because we are innately constrained to interpret 'gavagai' as referring to rabbit. However, this move merely begs the question against Quine. Likewise it could be

argued that Quine is begging the question against Chomsky et al. by assuming that there is no determinacy of meaning. It is the old philosophical chestnut of one man's modus-ponens being another man's modus-tollens.

At this point our debate between the two thinkers amounts to mutual question-begging against each other on the question of innate concepts. We have seen that if concepts are learned statistically through incremental data processing, then there can be no evidence to distinguish between the various interpretations of 'gavagai'. If, on the other hand, we assume that children are born with innate concepts which distinguish between the various different interpretations, then our problem is supposedly solved. However, given that behavioural evidence cannot decide between the various different interpretations, the postulation of innate concepts that solve the problem amounts to merely assuming without evidence that the problem does not exist.

It could be argued that this interpretation of mine is very unfair because it ignores the fact that from an intuitive point of view people believe that *RABBIT* refers to rabbit as opposed to undetached rabbit part, etc. Nativists could argue that our intuitions are good evidence for innate concepts. On this view, while Quine and Chomsky may be engaging in mutual question begging, since Chomsky's theory is consistent with our intuitions and Quine's is counter-intuitive, here we should accept Chomsky's theory until we have further evidence to indicate otherwise. However, this position is extremely tenuous. Chomsky has long argued that we should not let our intuitive theories of how the world works interfere with our scientific theories of the nature of the world. Thus in various different places Chomsky talks about how our intuitive conception of the nature of physics and psychology should not hold back our scientific theories of the mind and body. Likewise, our intuition that we have a

determinate meaning for terms like ‘gavagai’ should not dictate that we accept the intuition as an intuition of something real. If we have no empirical evidence to support the intuition then we can safely explain away the intuition. So it seems clear that Quine and Chomsky are begging the question against each other on this issue.

We saw above that Jerry Fodor has constructed arguments for innate concepts. It might be wondered at this point if his arguments can be used to offer independent evidence for innate concepts which could help the child overcome UD. This is a question that I will not consider here; it was demonstrated above that Fodor argues for innate concepts which are entirely different than the kind which Chomsky argues for. Furthermore, Fodor’s views of the nature of concepts are also at odds with the views put forth by Carey, Soja et al. So, given that Fodor holds a view of concepts incompatible with those of Chomsky, and Carey et al. it is doubtful that his conception can help their arguments. So I will bracket the question of whether Fodor’s views can help the efforts to claim that children overcome UD through innate concepts.

In the next section, I argue that even if children were born with innate concepts which helped them overcome UD in concept learning this fact would have little significance for the IDT.

PART 5: THE APTNESS OF THE ANALOGY

The case of our own small child and that of the adult foreigner are superficially alike: each of these persons has mastered a word attesting to the presence of red. But the two cases differ in that the reference of the foreigner's word has yet to be assessed; whereas the reference of the child's word has yet to be acquired. Assessment of the reference of the foreigner's word awaits only our systematic English manual of translation of his elaborate language... The child, on the other hand, is too young to have acquired any apparatus, English or otherwise, whereby to distinguish among these various possible references. We can credit him with the knack of responding distinctively to red episodes, but there the credit stops. (Quine: 1974, 83)

In this chapter, I have been reviewing some empirical research on the nature of concepts which shows that Quine is incorrect about how children develop their ontology. But nothing in this research tells us much about facts of translation; in fact it is only the strained analogy with the situation of the field linguist and the situation of the child that makes us think there is any connection at all. We have seen above that thinkers such as Chomsky, Pinker, Markman, and Boeckx all write as though the two situations were identical. It is my contention, however, that this analogy should be dropped, that we should not compare the situation of the field linguist with the situation of the radical translator. Firstly, the field linguist differs from the child in that the field linguist has an *explicit* theory of the world which will affect how he interprets the sounds of the native, whereas the child supposedly operates using an *implicit theory* which will automatically map certain sounds onto certain interpretations. The linguist will have a total network of beliefs about the world which he will bring to the translation situation. These beliefs will help him translate what the native is saying, but they will also ensure that we end up translating them according to our own lights. It could be argued that this situation is no different than the situation that the child finds himself in, and that the same solution is possible in

terms of positing innate constraints on the interpretations the translator can make. Furthermore, it could be argued that these constraints are justified, given that the translator and the native are both members of the species homo-sapiens and so are subject to the same innate constraints. Such a move, I argue, is not justified by the best evidence that we have available to us at the moment. Firstly, as we have seen above, we do not as of yet have sufficient evidence that innate concepts exist, for all that contemporary cognitive science has shown is that children have complex concepts from a very young age. Secondly, even if Homo-sapiens are, indeed, born with innate concepts, this would have limited significance for questions of translation.

Let us assume (wrongly) that contemporary cognitive science has proven that humans are born with innate concepts. So when our native begins to converse with the linguist, there will be a plethora of innate constraints which will in some ways determine how communication between the two takes place. These constraints have a lot in common with the types of constraints which Markman, Soja, etc talk about. Thus, one such constraint will be fact that the linguist will automatically track the eye movements of the native and try to triangulate with the natives on shared objects of experience. The linguist will use the language of the eyes to try out translations of the native's utterances by mouthing 'Gavagai' while looking at the rabbit. But obviously such innate constraints do not entirely determine what an adult is referring to when he uses the terms.

Consider Quine's claim about the nature of adult conceptual schemes:

Each man is given a scientific heritage plus a continuing barrage of sensory stimulation; and the considerations which guide him in warping his scientific heritage to fit his continuing sensory promptings are, where rational, pragmatic.
(1953, 46)

An adult speaks about reality and communicates with others based on what he has learned about the world through his scientific education, religious education, social

education, etc. When an ordinary adult speaks, he will speak in a language which reflects the ontological commitments of the community which he was socialised in. The discourse of any community will be shot through with assumptions which are justified in light of the total theory of the experts and the citizens of the community. It is possible that a child *may* be born with a folk biology which determines that it sees a moving furry object with eyes as an agent. And an adult may have the same innate bias wired in his brain; however, such a bias will not determine the speech of the adult. Imagine, for example, to borrow Putnam's idea that it had recently been discovered that rabbits were remote-controlled spies from Mars. In that situation, people may be innately biased to see rabbits as agents but would have scientific evidence that this bias was in fact false. Various theories of the world which are learned through pain-staking trial and error will influence what people think they are referring to when they refer to certain entities. When we are trying to understand the discourse of a native, we have no way of knowing what sort of ontological commitments they are adhering to. Do they have the same information about rabbits being remote-controlled spies as we do? The natives could be Platonists, or Nominalists, they could be Dualists or Monists. Without knowing their ontology prior to translation, we will have no choice but to assume that their ontology is the same as ours. This will mean, of course, that we are translating their language according to our own lights. The experiments of Markman, Soja and Spelke are not really relevant when it comes to facts about translation. The native and the field linguist may both be members of the species *Homo-sapiens* and hence will presumably learn language subject to the same constraints as scientists like Markman, Soja et al. claim. Nevertheless, the fact that we develop sophisticated theories of the world which go far beyond the crude ontology of children means that understanding the constraints on

language learning will give us little to go on in constructing a translation manual when we are trying to interpret our native.

Quine is quite explicit on this point: the IDT argument is separate from questions of how a child learns his first language. When Quine is speculating on how a child learns his first language, he is doing so because he believes that understanding how humans in fact develop their ability to refer will help us better understand the process of reference. And understanding our process of reference better will help us better grasp the process of ontic commitment.

We have seen in this chapter, that despite the fact that most cognitive scientists believe that the IDT has been overcome by the postulation of a language faculty, the evidence indicates otherwise. None of the experimental evidence demonstrates that children are born with innate concepts. So neither the Inscrutability of Reference nor the IDT is overcome by innate concepts. The fact is that, as of yet, we do not have enough evidence to indicate whether children are born with innate concepts. Furthermore, even if children are born with innate concepts, this fact would have no relevance for the IDT unless it could be shown that the innate concepts people are supposedly born with cannot be changed by experience.

CHAPTER 4: POVERTY OF STIMULUS ARGUMENTS: THE EVIDENCE

PART 1: CHALLENGES TO THE APS

SECTION 1: IMPLICATIONS FOR CHOMSKY AND QUINE

In Chapter 1 I detailed some of the central differences and similarities in Chomsky's and Quine's views on the nature of language and science. I argued that both thinkers were methodological naturalists who held virtually identical views on the philosophy of science. However, I also argued that when it comes to the details of how children learn their first language there is a substantive difference between them. The primary difference between them centres on the role that they think reinforcement plays in a child learning his first language. The Quinean picture of a child learning his first language involves the child using his innate babbling instinct as he mouths various different words. The parent reinforces these emissions positively and negatively until the child's pattern of verbal behaviour is battered into the external shape of his social environment. As Quine put it in *Word and Object*:

People growing up in the same language are like different bushes trimmed and trained to take the shape of identical elephants. The anatomical details of twigs and branches will fulfil the elephantine form differently from bush to bush, but the overall outward results are alike (1960, 8)

Linguistic nativist Chomsky disagrees with this Quinean picture. He thinks that the outward shape of language results not from the child's faltering attempts at speech being corrected by his peers, but from the child using his innate universal grammar to structure the data of experience which the child contingently encounters.

The issue between Chomsky and Quine on this point is a purely empirical one. In the last twenty years much detailed evidence has emerged which can be used to decide between the two theorists. The central idea around which nativism was built has been poverty of stimulus arguments. Chomsky has argued that children display knowledge of language, and that this knowledge is not provided by the environment,

and therefore it must be innate. In Chapter 1 I used the subject-auxiliary inversion rule to illustrate how poverty of stimulus arguments work. In their paper ‘Empirical Assessment of Stimulus Poverty Arguments’³⁵ Geoffrey Pullum and Barbara Scholz call the subject-auxiliary inversion rule the paradigm case which nativists use to illustrate poverty of stimulus arguments. They cite eight different occasions that Chomsky uses the example (Chomsky 1965, 55-56; 1968, 51-52; 1971, 29-33; 1972, 30-33; 1975, 153-154; 1986, 7-8; 1988, 41-47). They also cite other Chomskian thinkers (including linguists such as Lightfoot, 1991, 2-4; Uriagereka, 1998, 9-10; Carstairs-McCarthy, 1999, 4-5; Smith, 1999, 53-54; Lasnik, 2000, 6-9; and psychologists such as Crain, 1991, 602; Macrus, 1993, 80; Pinker, 1994, 40-42, 233-234) who here endorsed the claim. They argue that this supposed instance of an APS is being passed around and repeated over and over again. No surprise, then, that without having knowledge of Pullum and Scholz’s article, I chose the subject-auxiliary inversion as my paradigm example of an APS. I had fallen into the same pattern of passing around this well-worn example. In this chapter, I examine whether the APS argument (as applied to syntactic knowledge)³⁵ actually works. I will discuss how, if sound, the APS affects Quine’s view of language learning. It will be shown how Quine’s theories on language acquisition will be affected if the APS turns out to be false.

Pullum and Scholz (2002) show that poverty of stimulus arguments are used in a variety of not always consistent ways in the literature. Having surveyed some of the literature on the APS, they isolate what they believe to be the strongest version of the argument. The argument they construct is as follows:

³⁵ In Chapter 3, I considered Quine’s IDT as a form of APS and considered whether it could be overcome by innate knowledge. In that instance, I was considering the APS in the area of concept acquisition.

- (A) Human infants learn their first languages either by data-driven learning or by innately-primed learning.
- (B) If human infants acquire their first languages via data-driven learning, then they can never learn anything for which they lack crucial evidence.
- (C) But infants do in fact learn things for which they lack crucial evidence
- (D) Thus human infants do not learn their first languages by means of data-driven learning.
- (E) Conclusion: Humans learn their first languages by means of innately primed learning.

This gloss on the APS is one that Chomsky would accept as an appropriate schematisation of the APS, though Chomsky believes that there is more evidence to support a belief in innate domain-specific knowledge than the APS³⁶. Pullum and Scholz claim that the key to evaluating the soundness of the APS is premise (C) which is the empirical premise of the argument. So to evaluate the argument, they study the linguistic environment of children. Their aim is to check if there really is no evidence provided by the environment which the child can use to formulate a hypothesis of a particular rule. For example, will a child be presented with data such as ‘Is the man who is at the shop happy?’, which can help them learn the subject-auxiliary inversion? Pullum and Scholz’s research programme involves searching the *Wall Street Journal* corpus to discover if constructions which Chomsky claims a person could go much or all of their life without encountering are, in fact, more frequent than he would lead us to believe. However, prior to discussing what the evidence tells us about the frequency of the sentences, I first want to discuss what Quine would make of the APS as discussed by Pullum.

The first premise of Pullum’s version of the APS is that a child learns language either by data-driven learning or by innately-primed learning. Quine

³⁶ Independent of his APS argument, Chomsky cites various different strands to support his belief in a language faculty. He speaks of the supposed universals which exist in all the languages of the world. He also points out that the language faculty grows at a fixed rate and that general intelligence is not affected by the loss of the ability to speak and understand language. These points have all been contested in the literature. However, I will not discuss them here as my main concern is with the evidence which Chomsky uses to support his APS claims.

maintained throughout his entire philosophical career that our linguistic abilities are not distinct from our overall theory of the world. In fact, he has consistently maintained that learning a language is learning a theory of the world, and, furthermore that learning a scientific language is learning a more explicit regimented form of ordinary language. According to the picture presented in *Word and Object*, a child begins by babbling various different sounds and has these sounds reinforced in various different ways. Through the process of conditioning and reinforcement, the child eventually learns when it is appropriate to use which sounds. According to Quine, at this stage the child has not learned any concepts. Quine argues that through processes such as analogical reasoning, abstraction etc., children eventually learn to structure some of these sounds into syntactic units. It is only after we have mastered this syntax and can then speak of certain objects as being the same as or different than other objects, that we can be said to have grasped the concept of an object, and learned to speak about objects in the world. The important point is that for Quine, the processes which a child uses to learn a language are the same as the processes he uses to learn about the world. So Quine would not accept that language is learned by innately-primed learning (in the sense of innate domain-specific knowledge). The question of whether a child learns his first language by data-driven learning is a more complicated question on the Quinean picture.

On some versions of data-driven learning, the child is presented as a passive observer of verbal behaviour. From when they are born (strictly speaking, when in utero as well), children are bombarded with verbal behaviour. So, on one data-driven learning picture, children (unconsciously) observe the various different patterns of verbal behaviour; circumstances of occurrence, order of occurrence, tone used etc. and unconsciously construct a model of the language they are presented with.

Quine does not deny that the child uses such statistical methods to organise the data of experience; so in this sense he agrees with the statement that a child learns by data-driven learning. However, for Quine, the word 'data' has a much wider meaning than mere models constructed based on observed linguistic regularities. For Quine, an important part of the data is the type of reinforcement that the child receives. The child elicits utterances and receives various different types of reinforcement, either negative or positive depending on the appropriateness of the utterance. So on the Quinean picture, as the child is learning his first language, he might be reinforced for putting forth a question such as 'Will Mama feed me?'. Now suppose the child had been constructing questions by moving the first auxiliary of various statements to the front of the sentence, and suppose further that the child had been positively reinforced for this behaviour. Given this state of affairs, the child will continue to emit questioning behaviour like this until he receives negative reinforcement. Now suppose that the Quinean child wants to ask a more complicated question; suppose he wants to discover whether the sentence 'The man who is tall is sad' is true. The child will continue along the pattern of previous questions and will turn the statement into a question and ask 'Is the man who tall is sad?*. On the Quinean picture, this questioning behaviour will be negatively reinforced. The child will continue to try different constructions based on past experience and reinforcement until eventually their language output is moulded into the shape of the child's community. So for Quine, the data the child learns from is not merely observation of the statistical patterns of the language he is exposed to, but also includes the ways various constructions are reinforced negatively and positively. The important point is that in order for Quine to accept premise (A) of Pullum and Scholz's reconstruction of the

APS, he would have to understand data in a wider manner than that obtained by mere passive observation.

Quine would accept premise (B) as long as data-driven learning is considered in this expanded sense (reinforcement, plus statistical regularities in the environment). Premise (C) is the crucial empirical premise: but infants do in fact learn things for which they lack crucial evidence. The “parade case” of linguistic nativists where children display knowledge where they have not been provided crucial evidence is the subject-auxiliary inversion. Quine emphasises induction, analogy, and reinforcement as the primary tools in language learning. He has never endorsed the claim that children have knowledge of the rules of language for which they have not received data from their linguistic environment. For Quine, any sentence a child utters is either learned inductively from the child’s PLD or is constructed through an analogy with previously heard utterances in the PLD. Through induction, analogical reasoning and reinforcement, the child will eventually arrive at the language of his peers. So Quine would deny the truth of the crucial empirical premise (C). Furthermore, premise (C) is a crucial test of Quine’s theory of language acquisition: if it could be demonstrated that a child has knowledge of a rule of language which was not learned by experience, analogy or reinforcement, then this would demonstrate that Quine’s theory of language acquisition is seriously incomplete.

By reviewing the *Wall Street Journal* corpus, Pullum and Scholz have provided evidence that the constructions which Chomsky claims a child will never be exposed to in their lifetime do, in fact, occur. They used the *Wall Street Journal* because it is easy to obtain and free. People have justly complained that the *Wall Street Journal* is obviously not a representative of the type of data a child will be exposed to. To this they have replied that since Chomsky claimed that the type of

sentences which a child needs to be exposed to in order to learn structure dependence rules are so rare that a child can go much or all of his life without encountering them, then the *Wall Street Journal* is therefore evidence that Chomsky is wrong on this point. Geoffrey Sampson, in his book *The 'Language Instinct' Debate*, has provided evidence that the type of constructions which Chomsky claims are vanishingly rare occur in children's books. Furthermore, he has searched the British National Corpus (including a search of child-parent interaction) and found hundreds of examples of the relevant constructions³⁷.

Pullum et al. think that since they have shown that premise (C) is not in fact true, then the overall argument, while valid, is not sound and therefore the argument for linguistic nativism does not go through. Let us assume that Pullum is correct and that Chomsky's argument is not sound: what then are the implications of this for Quine?

As we have seen, Quine thinks that children learn language through data-driven learning, in a broad sense. One of the primary objections to the Quinean picture of language learning is that negative reinforcement does not play the role in language learning that Quine thinks it does. A wide variety of experimental evidence has been put forward by psychologists who claim that this evidence shows that children are not corrected when they speak ungrammatical sentences (see for example, Marcus 1993, 53-85; Gropen, J., S. Pinker, et al., 1989, 203-57; Crain, S. and M. Nakayama 1987, 113-25). At a superficial level, this seems to show that Quine's picture of language acquisition is incorrect. The empirical evidence seems to indicate that the picture of a child mouthing constructions such as 'Is the man who tall is sad?''*, and receiving negative reinforcement is, in fact, incorrect. Therefore, one

³⁷ I will deal with the frequency of the constructions later in the chapter and whether there are enough constructions for a child to learn the rules. For now I want to focus in a schematic way on what these findings mean to a Quinean picture of language learning

could conclude that even if Pullum is correct that the child is exposed to some examples which help the child learn the structure-dependent rule, this view will not help the Quinean conception of language learning.

However, it does not automatically follow that because explicit reinforcement is not involved in language learning that a more subtle kind of reinforcement is not used. Whether or not reinforcement is explicitly used in learning complex grammatical utterances, it is unquestionable that children do receive positive reinforcement for speaking. When a child begins to speak first, every utterance is encouraged and rewarded with affection. In *Word and Object*, Quine notes that any reinforcement that the child receives will be concomitant with a variety of different stimulations. As he writes:

The original utterance of 'Mama' will have occurred in the midst of sundry stimulations, certainly; the mother's face will not have been all. There was simultaneously, we may imagine, a sudden breeze. Also there was the sound of 'Mama' itself, heard by the child from its own lips. (1960, 81)

So, for Quine, the effect of the reinforcement will be that the child will repeat the word in the presence of Mama's face, in the presence of a mild breeze, and upon hearing the sound mama. However, the child will not receive reinforcement for saying 'Mama' in the presence of a sudden breeze, so will eventually stop emitting this behaviour. The child will, however, receive reinforcement for saying 'Mama' in the presence of mama, and for repeating the word 'Mama' upon hearing someone near him speak it. One helpful consequence of this type of reinforcement, according to Quine, is that the child who is being reinforced for repeating 'Mama' when someone says 'Mama' will, from the parent's points of view, appear to be engaging in mimicry. If the child can recognise that he receives reinforcement not just for sounds used in certain appropriate contexts but also for mimicking the behaviour of his peers, then he will have had a very useful tool reinforced. The child will have realised that it pays to

listen to his peers and to try to imitate their behaviour. To this end the type of statistical abilities postulated by people such as Lappin and Clark will be obviously useful in helping the child learn his first language. Furthermore, if what is being reinforced is mimicking behaviour, then the fact that certain sentences which Chomsky claims do not occur in the data do, in fact, occur, this will obviously be of vital importance for Quine's theory. Obviously, Quine's mimicking theory will only work if the child experiences the constructions which he displays knowledge of. All of this is schematic. While it does not show that Quine's theory of language acquisition is correct, it shows that recent research which purports to show that Chomsky's APS arguments in syntax do not work, can also play a role in supporting Quine's theory of language acquisition.

SECTION 2: RECENT CRITICISMS OF THE APS

The first criticism that I will consider is a logical argument which has been put forth by Geoffrey Sampson. In his *The 'Language Instinct' Debate*, Sampson claims that it doesn't matter whether there is data which refutes Chomsky's APS, because the argument is self-refuting in itself. He attributes to Chomsky the following claim: 'Language has certain properties no evidence of which is available in the data to which we are exposed to when learning the language.' He then asks how Chomsky can possibly know this. The adult's conscious knowledge of the properties of the language is based on observations of the language, but Chomsky claims that such observations are insufficient to determine the properties of the language. Therefore, if there is a grammatical rule which a language learner rarely or never encounters in their data, then there seems no reason why a linguist would encounter such a

grammatical rule either. Both the language learner and the linguist are exposed to the same data which Chomsky claims will not determine the rules of the language at all.

In essence, what Sampson claims is that if there is *no* evidence in the data from which a child can learn the rule, then there is *no* evidence in the data which justifies the linguist in postulating the rule. Hence, for Sampson, the APS is self-defeating. One possible way for the linguist to overcome this difficulty would be to claim that he uses his innate knowledge of grammar as well as observation to discover that the rule obtains. However, Sampson correctly notes that to argue thusly is to beg the question against your opponent. So he concludes that the APS is either self-defeating or a mere question-begging stipulation.

A key aspect of Sampson's argument is his emphasis on Chomsky's claim that children will never encounter certain constructions in their experience which could help them learn the relevant rule. He quotes the following statement of Chomsky's given in a 1980 lecture:

The child could not generally determine by passive observation whether one or the other hypothesis is true, because cases of this kind rarely arise; you can easily live your whole life without ever producing a relevant example...you can go over a vast amount of data of experience without ever finding such a case... (1980, 121)

In the above quote Chomsky claims that sentences which confirm the subject-auxiliary inversion rule are virtually never encountered. Sampson then asks rhetorically: if such sentences are never encountered, what reason would we have to say the rule exists?

The main difficulty with Sampson's argument is that the data a professional linguist is exposed to obviously far exceed the data a language learner would encounter. A child from a professional background will be exposed to about 30

million word tokens by the time they are three years old³⁸. A linguist from a similar background (assuming that he has completed a PhD and is around 27) will have been exposed to about 240 million word tokens³⁹. So a typical linguist will have encountered at least nine times the number of words that a typical child has. Obviously, if we accept Chomsky's claim that a child can go much or all of his life without encountering the relevant constructions, then the data the child is exposed to will be irrelevant. However, when Chomsky and other linguists discuss APS examples, what they typically state is that the data is insufficient for the child to learn the general rule, not that there is no data at all. So, bearing in mind the fact that a linguist is exposed to at the very least nine times the data that a language learner is, it is quite possible that the linguist will be exposed to enough examples to learn of the existence of the auxiliary inversion rule, while the child *may* not have been exposed to enough data to learn this rule from his PLD. Furthermore, the linguist will have access to other languages with which to compare his data from English. He will have conversational partners to discuss his findings with, and will have access to thousands of books and articles detailing the discoveries of other linguists. This evidence indicates that the linguist will have been exposed to much more than nine times the amount of linguistic data which your average child is. On these grounds, it is clear that Sampson's argument is inconclusive at best. To show that Chomsky is making a claim that is self-refuting, Sampson needs to demonstrate that a child and a

³⁸ See Hart and Risley's book *Meaningful Differences in the Everyday Experiences of Young Children 1995* for a discussion of the linguistic input an average child is exposed to.

³⁹ Hart and Risley's data refers to children hearing spoken language around them. An educated adult who is an avid reader would have access to much more data per year than the data spoken to him by his peers. So a twenty seven year old would have at least nine times the data a three year old has learning his language in terms of heard linguistic utterances. However if you count the data the adult receives from reading, the data the adult is exposed to would be much higher than nine times that of a three-year old child.

professional linguist are exposed to the same amount of linguistic data. Such a claim is of course patently absurd.

While Sampson's argument does not work as a demonstration that Chomsky's APS is self-refuting, it does reveal a real weakness in Chomsky's APS. Chomsky is making claims about the child's PLD for which he has provided no evidence. So Sampson's argument at least demonstrates the necessity of Chomsky providing evidence for the controversial claims he is making.

Pullum and Scholz did the first detailed study of how often sentences relevant to the structure-dependent APS appear in the data a child is exposed to. As I discussed above, they began by making the logic of the APS explicit by structuring it as a logical argument. They isolated the third premise which claims that data relevant to learning the structure-dependent nature of language do not occur enough in the child's PLD for him to learn the relevant construction. They set out to test this claim by checking a corpus of linguistic text; they used the *Wall Street Journal* as their corpus because it was freely and easily available.

In order to test how often a construction is encountered by a child learning a language, it is first necessary to test how much linguistic data a child is exposed to. Pullum and Scholz relied on the work of the psychologists Hart and Risely, who in their 1995 *Meaningful Differences in the Everyday Experiences of Young Children*, detailed the amount of linguistic data a child is exposed to. Hart and Risely documented the vocabulary development of forty-two children aged 1-3. The authors noted the production and use of language of the children as well as the language they were exposed to. They also noted that the amount of linguistic data a child is exposed to depends greatly on the socio-economic class that they belong to. According to their study, a child from a professional household will have been exposed to about 30

million word tokens. A child from a working-class family will have been exposed to 20 million word tokens. And a child from a family on welfare will have been exposed to 10 million word tokens.

Pullum and Scholz also report findings from Hart and Risely's book which indicates that 30% of the speech directed at children is in the form of interrogatives. Hart and Risely also estimate that the mean length of utterances directed to children is four words long. Pullum and Scholz then argue that if we take the statistic of a child whose family are on welfare, being exposed to 10 million word tokens, divided into sentences four words long, we arrive at the conclusion that the child is exposed to 2.5 million sentences every three years. And furthermore since 30% of those sentences are interrogatives, we can argue that the child is exposed to seven hundred and fifty thousand questions every three years, i.e. a quarter of a million questions per year. In their research of the *Wall Street Journal* they discovered that the questions relevant to learning the structure-dependent rule occur in 1% of interrogatives in the corpus. From this they conclude that a child will typically be exposed to seven thousand five hundred relevant examples in three years. This means that the child will be exposed to two thousand five hundred examples per year; therefore on average the child will be exposed to seven relevant questions a day. They conclude their paper by asking if seven relevant questions a day is enough to learn such a rule. Furthermore, they correctly claim that if nativists think that it is not, they need to explicitly set out a learning theory which shows why it is not.

The obvious objection to the above argument is that Pullum gets his data from *The Wall Street Journal*, and such data is hardly representative of the linguistic experience of the child. Pullum cites some research which shows uniformity across linguistic texts as evidence that *Wall Street Journal* may, in fact, be representative of

the child's linguistic experience. However, the fact that the Hart and Risely research claims that child-directed sentences are typically four words long shows this to be incorrect. The average length of sentences in the WSJ will obviously be much longer than four words. Geoffrey Sampson's research taken from the British National Corpus (not available in America at the time Pullum and Scholz were writing) uses samples of speech between child and parent as well as the ordinary speech of adults, so it avoids some of the difficulties of Pullum and Scholz's research.

In his (2002) paper "Exploring the Richness of the Stimulus", Sampson largely agrees with Pullum and Scholz's research; however, he claims that while his research is complementary to theirs it is not subject to the same objections. He sampled the normal conversational speech which people typically have with each other and which a child is routinely exposed to. To this end he used the British National Corpus (henceforth BNC). He used the demographically sampled speech section of the BNC which he claimed contains 4.2 million words. This section of the BNC was constructed by giving recording equipment to individuals selected to be representative of the national population with respect to age, social class, and region (2002, 3). By exploring this corpus, Sampson aimed to avoid the criticisms directed at Pullum and Scholz which claimed that their corpus did not accurately represent the data a child is exposed to when learning a construction.

Sampson begins his discussion by making a terminological point. Whereas Pullum and Scholz use the term 'auxiliary verb' as something that can be the main and sole verb of a clause, Sampson calls a verb 'auxiliary' only if it is followed by another verb. For this reason, while Pullum and Scholz would call the following sentences auxiliary inversions, Sampson would call them 'verb-fronting questions'.

VERB-FRONTED CONSTRUCTIONS

Here we will discuss what Pullum and Scholz refer to as ‘auxiliary-initial clauses’. Poverty of stimulus theorists claim that children typically will not hear examples of questions formed by fronting verbs which in the corresponding declarative statements are preceded by complex constituents. Sampson aimed to test what people actually say when speaking to each other. He did this to help him understand whether the poverty of stimulus theorists were correct. However, when trying to analyse the data he found an unexpected complication. There are two different types of verb-fronting sentences, both of which Pullum and Scholz include in their WSJ search. These different types of constructions occur in radically different magnitudes in spoken speech.

The first type of verb fronting is of the following form:

(1) Will those who are coming raise their hands?

(1a) Those who are coming will raise their hands

Sampson reminds us that in the above constructions, the complex constituent is the subject of the fronted verb. So he calls sentences 1 and 1a verb-fronting sentences which involve complex preverbal subjects.

The second type of verb fronting has the following form:

(2) If you do not need this, can I have it?

(2a) If you do not need this, I can have it.

Sampson reminds us that in 2 the main clause is preceded by an adverbial clause. He calls sentences like 2 and 2a “verb-fronting sentences” involving initial adverbial clauses. He first begins to consider questions of the form of 2 which he calls initial adverbial clauses.

INITIAL ADVERBIAL CLAUSES

Sampson searched for adverbial initial clauses in the BNC-demographic (which contains 4.2 million words). He claimed that his search was not exhaustive because such an exhaustive search would be extremely difficult with this grammatical pattern and the BNC corpus. He did not offer any reasons why this particular grammatical pattern would make an exhaustive search so difficult. However, he did claim that such a detailed search was not necessary since Chomsky had claimed that 'a person might go through much or all of his life without being exposed to a relevant construction', and that therefore finding any examples of the constructions would refute Chomsky.

In attempting to find such examples, Sampson targeted cases where the adverbial clause begins with *if*. He found twenty-two clear cases of initial adverbial clauses. He furthermore claimed that Wh-questions could also be considered relevant. Wh-questions also involve moving an auxiliary of the main clause, rather than one in the preceding adverbial clause. And he claimed that that if this class is relevant, then he had a further twenty-three cases. However, he realised that counting Wh- questions would be controversial, so he only counted the twenty-two constructions which he found for initial adverbial clauses.

Sampson uses Hart and Risely's estimates of how many words a person is exposed to every three years. He takes the figure they provide that a working class person is exposed to twenty million words every three years. This choice itself is controversial; there is no reason to focus on the stimuli that a working class child is exposed to rather than the stimuli that a professional child is exposed to, or the stimuli that a child from a family on welfare is exposed to. If we accept that children from linguistically deprived backgrounds develop normal linguistic abilities, then the figure

of ten million should be used because children develop such abilities despite only being exposed to this amount of linguistic data. Furthermore, if the relevant constructions do not occur in the data, and children display competence of the rules, then this shows that the rule must be innate. However, Sampson would probably reply to this that the argument relies on the untested assertion that people from linguistically deprived environments have languages as richly structured as those of ordinary members of the linguistic community. Sampson has long argued against the dogma of convergence, the view that all speakers from all societies speak languages which are equally complex. He holds that if we are to establish that children from linguistically deprived environments have language as complex as their better educated colleagues, then we will need evidence to support this claim. And he holds further that nativists have so far not provided us with any evidence of this kind.

So to avoid begging the question against either nativists or anti-nativists, it is best to start, as Sampson does, with Hart and Risely's figure of twenty million words every three years. So let us work out the numbers. Sampson found twenty-two constructions out of a corpus consisting of 4.2 million words. Using Hart and Risely's data, we can estimate that the average length of each construction for a child up to three is four words long. So we can estimate that Sampson's 4.2 million words amounts to about 1.1 million sentences in the corpus. Hart and Risely estimate that a working class child will be exposed to five million sentences (of four words long) in the first three years of their lives. So if Sampson finds the relevant data twenty-two times out of 1.1 million sentences, then we can expect that he will find at least one hundred and ten examples in five million sentences. This would work out at about

thirty seven relevant examples per year. So a child could expect to encounter a relevant construction at least once every ten days.⁴⁰

The question which Pullum and Scholz raise in their paper can be fruitfully asked of Sampson's results: is one example every ten days enough for the child to learn the construction? The nativist who is claiming that innate domain-specific knowledge is the only explanation for our competence in the relevant construction owes us an answer as to why we cannot learn it from one example every ten days. Typically nativists have not met this challenge; they have merely pointed to the supposed poverty of stimulus as evidence that the construction must be innate. However, likewise, if anti-nativists claim that the relevant construction can be learned using some kind of data-driven learning, then they owe us a model of how this is done. Assessing whether such constructions can be learned by experience will require mathematical models of how learning from such few constructions is possible. Other possible tests may involve developing computer programmes which can learn from this amount of data. Such programmes have been developed already. So, for example, Clark and Eyraud (2007), Perfors et al. (2006), Reali and Christiansen (2005) have all developed programmes which can learn from less data than discovered by Pullum, Scholz and Sampson. I will review these models at the end of this chapter. Ultimately what we have learned from this data is that Chomsky's confident assertions that children cannot learn certain constructions from the data they experience have not been justified with enough evidence.

⁴⁰ Like Pullum and Scholz, Sampson interprets his corpus data using Hart and Risely's figures. One difficulty with this is that Hart and Risely estimate that the sentences a child encounters are 4 words long. The few examples Sampson published from his corpus research contain sentences which are 10 words long. So doubt could be cast on whether his corpus is representative of the child's PLD. Obviously much further research is needed to clarify this matter; however, like Pullum and Scholz, Sampson is to be applauded for beginning this research into the PLD instead of ignoring it like Chomsky.

The other type of verb fronting which Sampson discusses is the type of construction where the complex constituent is the subject of the fronted verb. An example of this type of construction is:

(1) Those who are coming will raise their hand.

(2) Will those who are coming raise their hand?

Here Sampson found some surprising results. Sampson discovered that on this point Chomsky was correct. In the 4.2 million word BNC, Sampson found no constructions of the relevant kind. However, he did not view this as providing support for Poverty of Stimulus theorists. He claimed, on the contrary, that the reason that the construction did not occur in the BNC is because the construction is not an idiom of ordinary English *speech*. It is rather an idiom of *written* English.

Sampson's search of the speech-directed portion of the corpus showed that the relevant construction never occurred in 4.2million word tokens. He did not do an exhaustive search of the written-language section of the BNC; instead he merely provided examples from random searches of the corpus. Here are some of the examples he found:

(14a) Did the fact that he is accompanied by a doctor on the campaign trail help to lose him last week's TV showdown with Clinton? CAT.00742 (*Punch* magazine, 1992)

(15b) Did Mr Mortimer, 69, who has an Equity card, enjoy himself? CBC.08606 (*Today* newspaper, 1992)

(16c) 'Is the lady who plays Alice a child or a teenager?' asked my six-year-old' B0300647. (*Alton Herald* newspaper, Farnham, Surrey, 1992)

(17d) Is a clause which is known to be unenforceable in certain circumstances an unreasonable one? J6T.00908 (R. Christou, *Drafting Commercial Agreements*, Longman, 1993)

(18e) Will whoever is ripping the pages out of the stony new route book please grow up. CG2.1379 (*Climber and Hill Walker* magazine, George Outram and Co., Glasgow, 1991)

(2002, 18)

Sampson thinks that these examples show that children do not typically form questions using auxiliary fronting when speaking. He argues that constructing questions using auxiliary fronting is restricted to written questions. However, he does not provide any evidence as to how often such constructions occur in written work. His primary point is that, in order for an APS theorist to use a lack of examples in speech of yes/no questions formed by fronting a main-clause verb as evidence for innate knowledge, they have to rule out the possibility that children learn the rule from written language. The fact that people will judge certain constructions as grammatical, despite not encountering them in spoken language, is not that important if the person has encountered them in written language. Here, in short, Sampson is shifting the burden of proof onto the APS theorist to show that the child cannot learn such rules from written language. And in absence of such a proof, he is assuming that the APS does not hold.

What Sampson and Pullman and Scholz have shown is that the “parade case” of APS as put forth by Chomsky does not offer clear evidence at all. Obviously much more research is needed on the topic. The important point to note is that this APS has been shown to be incorrect in claiming that children learn a particular rule in the absence of experience. Hence, this particular APS does not establish that Quine’s

conception of language learning is incorrect. The question of the viability of Quine's story of how the child learns his first language remains open. Nor of course can this APS be used to support Chomsky's claim that we need to postulate innate domain-specific knowledge to explain language acquisition.

SECTION 3: NEGATIVE EVIDENCE

In order to achieve a complete picture of what contemporary evidence tells us about the debate between Chomsky's and Quine's picture of language learning, we will now need to evaluate what the state of play is in regard to the issue of negative evidence. Most linguists believe that the issue of negative evidence is crucial to understanding language acquisition. The issue of negative evidence centres on the fact that children do not typically encounter ungrammatical sentences which are marked as such. A child will not, for example, hear a sentence such as 'Is the child who beside the man is happy?' along with a tag to indicate that the sentence is deviant. So the question arises as to how children know that these sentences are ungrammatical. The children are not presented with these sentences and told they are ungrammatical. Nor (so the theory goes) do they produce these ungrammatical sentences only to be systematically corrected by their peers. Hence, it is argued that the only way to explain how a subject tested by a linguist can clearly tag certain sentences as grammatical, and certain as ungrammatical, is to postulate innate domain-specific linguistic knowledge.

A key premise in the above argument is that children are not systematically corrected for their grammatical mistakes by their peers. This claim goes back to the experimental research of Brown and Hanlon (1970). In particular, Crain and Nakayama (1987) test the claim that children try out grammatical theories and weed out the false ones through explicit teaching from their peers.

Crain and Nakayama elicited yes/no questions from children between the ages of 3;2 and 5;11 in response to prompts such as *Ask Jabba if the boy who is watching Mickey Mouse is happy*. They found that (with different frequencies at different ages) children sometimes produced correct forms such as (15a) and they sometimes produced various incorrect forms, one example being (15b). However, they never produced the kind of incorrect form predicted by the ‘structure-independent hypothesis’ such as (15c). They offered this as support for the theory of innate linguistic knowledge⁴¹. Below are the two examples which children sometimes produced, and the structure-independent hypothesis which children never reproduced:

- (15a) Is the boy that is watching Mickey Mouse happy?
- (15b) Is the boy who’s watching Mickey Mouse is happy?
- (15c) Is the boy who watching Mickey Mouse is happy? (2002, 20)

Examples like (15c) are the type of production which one would predict based on the structure-independent hypothesis. However utterances like these do not ever occur. In particular they claim that this shows that children are innately predisposed to prefer structure-dependent rules to organise the data of experience (Sampson 2002, 20). So here the issue is that children do not try out constructions like (15c) and have them criticised by their peers. They automatically construct questions like (15a) and (15b) which are structure-dependent rules. That is, independent of poverty of stimulus considerations, Crain and Nakayama claim to have shown that children do not construct structure-independent rules, and so do not receive any negative evidence that sentences (15c) are ungrammatical. If we add to Crain and Nakayama’s claim the fact that children do not hear sentences like (15c) spoken yet know that they are ungrammatical, we have an argument for innate knowledge based on a lack of negative evidence.

⁴¹ In this section I am following Sampson’s (2002, 20) reconstruction of Crain and Nakayama.

The argument of Crain and Nakayama is of vital importance to our thesis. It offers support to Chomsky's claim that children are born with an innate language faculty. It also contradicts Quine's picture of a child learning the rules of syntax through positive and negative reinforcement. Obviously, if children do not utter constructions such as (15c), then Quine's claim that such constructions are shown to be incorrect through negative reinforcement must be false⁴².

The supposed lack of negative evidence in the instance of auxiliary inversion may not be as damning to Quine's picture of language learning as it appears. The data which Pullum and Scholz gathered from the *Wall Street Journal* indicates that in the case of subject-auxiliary inversion, children encounter about seven relevant constructions every day. Using statistical reasoning,⁴³ the child exposed to this type of experience from passive observation alone would within a few days of birth have evidence that the structure-dependent hypothesis was superior to the structure-independent one. A child who was unconsciously analysing the data of experience would then not even try the structure-independent rule of (15c) though he may try structure-dependent rules such as (15a) and (15b).

However this argument fails when one takes account of Geoffrey Sampson's work. His data shows that examples of verb-fronted sentences (excluding cases where the subordinate clause precedes the main subject which we discussed above) in actual speech is zero. So the child cannot learn the rule inductively. If Crain and Nakayama are correct that children never try out the barred interpretation, then this indicates that they are correct that the rule is innate.

⁴² Quine does not anywhere discuss auxiliary inversion; however, his constant emphasis on the fact that language is a social art in which people's utterances are beaten into shape through reinforcement from their peers shows that he thinks that all our linguistic rules are structured through positive and negative reinforcement from our peers.

⁴³ Later in the thesis, I will discuss Reali and Christianson's mathematical models which demonstrate that it is possible for a child to learn structure-dependent rules with even less data than Pullum and Scholz discovered.

So it could be argued that Crain and Nakayama's result combined with Sampson's show that Quine's conception of language acquisition is incorrect. A critic could argue that Sampson and Crain and Nakayama's research shows that Quine is incorrect because they show that induction and negative and positive reinforcement do not play any role in learning that this particular rule is incorrect. However, given that Crain and Nakayama's experiment only relates to the subject-auxiliary inversion, it could be argued that it is obviously not equipped to rule out stimulus/response learning entirely either. However, there is a more fundamental reply which could be raised to this experiment. The reply is Geoffrey Sampson's and uses data from how people actually speak which is used to cast doubt on Crain and Nakayama's experiment.

Sampson has raised objections to this experiment based on his discoveries that children do not typically form questions using auxiliary inversion. He correctly notes that based on his corpus research children would not be expected to reply in the manner they do in the experiment. Crain and Nakayama use the fact that children never try out (15c) to support their claim that children do not use the structure-independent hypothesis. However, Sampson points out that in ordinary speech, as revealed by his corpus analysis, people do not use auxiliary inversion to form questions. According to his analysis, children should form the question in the following manners:

(16a) Is he happy, the boy who's watching Mickey Mouse?

(16b) The boy who's watching Mickey Mouse is happy, isn't he?

Sampson correctly notes that, since we know from the corpus analysis that children do not typically form questions like (15a,b) in speech, it is odd that children would answer in this way in the experiment. He points out that Crain and Nakayama only

give figures for children's 'correct' question formation, so it is impossible to tell whether children tried out '(16a), and (16b)'. He further speculates that the fact that children use an idiom of speech not ever found in ordinary discourse may indicate that the children were primed for the experiment. Sampson's discussion does not refute Crain and Nakayama's experiment; however, it does demonstrate that the experiment is far from conclusive. So Crain and Nakayama's experiment does not refute Quine's trial and error position of language acquisition. The Crain and Nakayama experiment would need to be replicated and done in different cultures to be viewed as anything more than a suggestive idea. The other experiment which is typically offered as evidence that children do not learn their language through trial and error is Brown and Hanlon's paper.

Brown and Hanlon's 1970 paper purports to show that children do not learn by explicit instruction. However, their paper has negligible impact on Quine's position, because behaviourism is not committed to reinforcement being explicit. In fact contemporary research in psycholinguistics supports the view that much language instruction is implicit rather than explicit. In the section below I will enumerate experimental research which bears on this point.

SECTION 4: EVIDENCE OF NEGATIVE EVIDENCE

In their 2003 paper "Adult Reformulations of Child Errors as Negative Evidence", Chouinard and Clark constructed an experiment which was designed to test whether adults were implicitly instructing their children about the rules of their language. The test aimed to discover if adults were using side sequences and embedded constructions as ways to correct the children's utterances. On pages 9 and 10 of their paper they gave the following examples of side sequences and embedded corrections.

An example of a *side sequence*: (indented sequence is the correction).

(1) Roger: now-, um do you and your husband have a j-car.

Nina: have a car?

Roger: yeah.

Nina: no-

(Startvik and Quirk 1980: 8.2a 335)

An example of *embedded corrections*.

(2) Customer in a hardware store looking for a piece of piping:

Customer: Mm, the *wales* are wider apart than that.

Salesman: Okay, let me see if I can find one with wider *threads*.

(looks through stock) how's this?

Customer: Nope, the *threads* are even wider than that.

They claimed that adults made use of side sequences and embedded corrections to correct children's errors and to keep track of what the children meant to say. The adult reformulations indicate to children (a) that they have made an error, (b) what the error was, and (c) the form needed to correct the error.

In their experiment they set out to test the following four claims:

(1) Negative evidence is available in adult reformulations.

(2) Negative evidence is available to children learning different languages, and for different types of errors.

(3) More reformulations are available to younger children.

(4) Children detect and make use of the corrections in reformulations. (2003, 12)

METHODS USED IN THE EXPERIMENT

The experimenters got their data from five corpora in the CHILDES Archive. Three of the children were acquiring English (Abe from the Kuczaj corpus, Sarah from the Brown corpus, and Naomi from the Sachs corpus) and two were acquiring French (Philippe from the Leveille and Suppes corpus and Gregoire from the Champaud corpus) (Ibid., 13). In order to analyse child errors, the experimenters included all spontaneous child utterances in the transcript, with the exception of utterances with unintelligible speech and child utterances preceded or followed by unintelligible speech on the part of adults. The experimenters first tested whether the children's strings were adequate. If the string contained an error, they categorised what sort of an error it was, i.e. morphological, syntactic etc. They then checked

whether the next adult utterance was a reformulation. The utterance was a reformulation if it repeated in corrected form the portion of the child's utterance which had contained an error. They further coded the correction by noting whether it was side sequence or an embedded correction. They finally checked whether children took up the repeated change that had been made, rejected it or tacitly accepted it.⁴⁴

For the analysis of conventional child utterances, they took a random sample of 200 utterances for every age slice for each child. They identified all the error-free child utterances in the sample and tabulated how many of the constructions were replayed by the adult in the next turn. If the adult just repeated what the child said, they called it a *replay* (ibid., 14). They got two different researchers to code each transcript and they agreed on their codes 90% of the time. Where the two researchers disagreed, they resolved their disagreement by discussion.

Once they had coded the transcripts, they coded each of the lines for detailed analysis. For each of the children, they enumerated the total coded lines and the total number of erroneous utterances. They then divided the data into age slices to track developmental trends.

RESULTS

The following is a list of their results which bear on the five hypotheses which they put forth at the beginning of their paper.

(1) Negative evidence is available in adult reformulations.

They devised a table to represent the four different age slices of the three children in the English corpora, and they divided the adult replies into conventional and erroneous. They found that adults repeated erroneous utterances far more than they repeated conventional utterances. On average, they repeat erroneous utterances more

⁴⁴ The study used only adults who were the children's parents.

than twice as often as they repeat conventional utterances. More interestingly, the percentage of corrections of erroneous utterances is extremely high. In the age slice 2.0-2.5, (of the three English-speaking children) the following pattern was observed: Abe had 67% erroneous utterances reformulated, Sarah had 65%, Naomi had 48%. In the French corpora, Philippe had 67% reformulations, and Gregoire had 60%. So in the age range 2.0-2.5, most of the children had at least 60% of their utterances reformulated, and Naomi, who had the lowest number of reformulations still received almost 50% corrections of her reformulations. Out of the other age-slices, the lowest number of reformulations for incorrect utterances was for the ages of 3.6 to 3.11. Here there was not enough data to complete the French reformulations; however, the reformulations for the English speakers were as follows. Abe received 28% reformulations for incorrect utterances, Sarah 41%, and Naomi 20%. So it is certain that children do receive reformulations of incorrect utterances. Even in the worst case that of Naomi between the ages of 3.6-3.11 20% of her incorrect utterances were corrected. Below, I discuss whether 20% correction is enough for a child to learn various different rules. Of the corrections given side sequences made up the majority of corrections the children heard as opposed to embedded corrections. Chouinard and Clark (*ibid.*, 21) claimed that of the five children Abe, Sarah, Naomi, Philippe and Gregoire, the amount of side sequence corrections respectively was 57%, 70%, 70%, 73%, and 62%. In other words the reformulations in the majority of cases were designed to check what the child had meant.

NEGATIVE EVIDENCE IS AVAILABLE FOR CHILDREN LEARNING DIFFERENT LANGUAGES, AND FOR DIFFERENT TYPES OF ERRORS

The study found that negative evidence was available for each of the children whether they were French or English. Furthermore, negative evidence was used at

comparable rate whether the error was phonological, morphological, lexical or syntactic. And again, reformulations occurred at a much higher rate than repeating of conventional utterances.

MORE REFORMULATIONS ARE AVAILABLE TO YOUNGER CHILDREN

In general, this prediction was shown to be correct. Adults tend to decrease their reformulations as children get older and make fewer mistakes. However, there was one exception to this trend: as Naomi got older, her errors were reformulated more. So this question needs to be looked into further.

CHILDREN DETECT AND MAKE USE OF THE CORRECTIONS IN REFORMULATIONS

Obviously, just because adults use reformulations it does not follow that such reformulations are understood and used by the children. Evidence that children understand and use such reformulations can only be discovered by noting how the children respond to them. Chouinard and Clark discuss four possible ways that children could respond to an adult reformulation: (1) they can *take up* the reformulation explicitly by repeating it and, in doing so, correcting at least part of their original utterance; (2) they may overtly *reject* the adult's reformulation, thereby signalling that the parent has misinterpreted what the child intended, and when the parent tries a different reformulation the child may accept it; (3) after hearing a conversation, they may *acknowledge* it at the start of the next turn in the conversation. (4) they can simply *continue* with the conversation without overtly acknowledging the change or taking it up; such continuations could be counted as tacit acceptances of adult reformulations. Overall, the responses where children acknowledged a reformulation or repeated new information, alongside those where they either took up or else rejected the formulation were as follows: Abe 56%-72%, Sarah 25%-38%,

Naomi 39%-100%, Philippe 39%-75%, Gregorie 25%. By any standard, this shows that children do attend to reformulations a sizable percentage of the time.

GENERAL DISCUSSION

Four of the children used in the experiment had a parent who was college-educated; however, as Chouinard and Clark acknowledge, it is unclear whether this experiment will generalise across social classes.⁴⁵ Furthermore, only two cultures were used in the experiment, so it is unclear whether the experiment can generalize across cultures. Chouinard and Clark further discuss the oft cited evidence from Ochs and Schieffelin (1984), who claim that in some cultures negative evidence is not presented to children, because the child's parents do not interact with the child till they are competent speakers. Their paper is of central importance because it is usually offered as key evidence in favour of the nativist's argument for innate domain-specific linguistic knowledge. Usually nativists will point to Brown and Hanlon's 1970 paper, as well as Crain and Nakayama's 1987 paper, to demonstrate that children do not receive explicit negative evidence. However, typically anti-nativists point out that, while there is evidence that children do not receive explicit instruction, they do in fact receive implicit instruction. To this nativists reply that, while this may be so in our culture, it is certainly not so in all cultures. To demonstrate this point, they cite Ochs and Schieffelin's paper. It is claimed that since *all* of the members of the various different human cultures learn a language, and only *some* receive negative evidence, then negative evidence is not a key factor in learning language. This objection is clearly relevant to the work of Chouinard and Clark as they only considered two cultures in their experiment. However, Chouinard and Clark consider this objection (*ibid.*, 39) and offer a criticism of it. Ochs and Schieffelin had claimed that in Kaluli

⁴⁵ See Pinker 1994 on the Negro families who do not speak to their children.

and Samoan cultures parents do not converse with children who are not yet competent users of language. They claim that, contrary to what is typically believed, Ochs and Schieffelin's paper is in fact largely consistent with their own findings. It is true that in the Kaluli and Samoan cultures adults do not converse with children who are not yet competent speakers of the language. However two points need to be made about this. Firstly, the fact that adults do not engage in conversation with children of this age may not be as important as sometimes considered if older children in the community converse with the younger children of the community.⁴⁶ Secondly, even if parents do not converse with children who are not yet competent users of language, it does not follow that they do not correct their language use. Chouinard and Clark cite a section of Ochs and Schieffelin which indicates that children of the Kaluli do indeed receive negative feedback:

Kaluli mothers pay attention to the form of their children's utterances. Kaluli correct the phonological, morphological, or lexical form of an utterance or its pragmatic or semantic meaning. (1984, 293)

Chouinard and Clark claim that in the Kaluli culture feedback takes the form of adults telling the children what to say on different occasions. So, for example, the adult would prefix a child's statement with the instruction 'elema' (meaning 'say like that'). If the child makes a statement which is grammatically incorrect when talking to another person, the adult will face that person and say 'elema' followed by the grammatically correct utterance. So clearly in this culture explicit instruction is used to teach the child how to speak. However, contrary to what is typically reported, Ochs and Schieffelin do not provide evidence against anti-nativist theories of language acquisition.

⁴⁶ Whether the older children of such a community do, in fact, engage with the younger children in a manner in which they can use to learn the rules of their language is an empirical question. The point is that Ochs and Schieffelin's paper does not rule out this possibility.

The overall conclusion of Part 1 of this chapter is that contemporary evidence is still largely consistent with the picture of language acquisition sketched by Quine in the 1960's. The arguments by Chomsky and those influenced by his paradigm have not shown that Quine's view of language acquisition is incorrect. However, there are some nativists who claim that Pullum, Scholz, and Sampson, et al. are guilty of attacking a straw man. It is claimed, that the APS they criticise is not really the APS which Chomsky uses. In Part 2 of this chapter I will consider whether the alternative conception of the APS argued for by John Collins avoids the counter arguments by Pullum et al. I will further analyse how this alternative APS affects Quine's picture of language acquisition. I will then discuss Chomsky's latest attempt to defend the poverty of stimulus argument and will demonstrate that it fares no better than earlier versions of the argument.

PART 2: CONCEPTIONS OF THE APS

INTRODUCTION

In this part of Chapter 4, I consider an argument by the philosopher and Chomsky scholar John Collins which purports to show that Pullum et al. have misunderstood the nature of Chomsky's APS. Collins directs his arguments against Fiona Cowie, a philosopher who in her book *What's Within*, uses Pullum and Sampson's data to argue against Chomsky. Since Cowie uses Pullum's reconstruction of Chomsky's APS Collins's reply to her can be taken as a reply to Sampson, Pullum and Scholz.

I will illustrate precisely what Collins takes the real APS to be. I will then show that it is Collins who misconstrues the APS, not Pullum et al. I will therefore argue that, contrary to what some Chomskians claim, Pullum et al. raise serious difficulties with the research program of generative linguistics. Furthermore these

difficulties with Chomsky's programme show that despite the rhetoric employed by Chomsky, Quine's alternative conception of language acquisition is still very much a live option.

SECTION 1: THE REAL APS?

John Collins's paper "Cowie on the Poverty of Stimulus" is a review of Fiona Cowie's book *What's Within*. Collins focuses on the sections of her book which criticize Chomsky's poverty of stimulus arguments. He argues that Cowie has seriously misunderstood the aims and methods of Chomsky's paradigm in linguistics. Much of Cowie's criticisms of Chomsky are derived from data from a 1996 paper which Pullum gave called "Learnability, Hyperlearnability and Poverty of Stimulus". In this paper, Pullum argued that the data which a child is exposed to when learning his first language is not as impoverished as Chomsky and his followers claimed. While Cowie presents Pullum as claiming that his data refutes nativism, Pullum in a later paper (2002 co-authored with Scholz) chastised Cowie for this, and argued that their data only show that more research is needed into the child's PLD, not that nativism has been refuted. I am including Collins's criticism of Cowie in this section, because his reply to Cowie's use of Pullum's data indicates that he thinks that both Cowie and Pullum have misconstrued the nature of the APS. Other authors, for example Legate and Yang (2002), try to meet Pullum and Scholz's challenge by showing that the data Pullum and Scholz discovered is, in fact, insufficient to learn the structure-dependence rule. Collins, in contrast, argues that Pullum's construal of the APS is incorrect and that because of this his data does not cast doubt on the *real* APS.

Thus, for example, Collins writes:

Cowie, to be fair, does have Pullum's reconstruction of the 'Chomskian argument' in mind. Pullum presents the argument so as to refute it, but Cowie

finds it an 'irresistible target', for it is 'so much more clearly and forcefully stated than the nativists own versions'. The nativists' versions are not 'clearly and forcefully stated', I have suggested, because no-one serious is interested in knock-down arguments; there are certain empirical and theoretical constraints and a substantive proposal to satisfy them. (2003., 21-22)

I turn next to Collins's criticism of Pullum's APS

COLLINS ON PULLUM AND SCHOLZ'S VERSION OF THE APS

Pullum and Scholz's version of the APS is admirably clear. It isolates the key empirical premise of the argument, and proceeds to analyse the empirical data to check whether the key premise is in fact correct.

Collins however does not view Pullum and Scholz's APS as representing Chomsky's *real* APS. One of his first criticisms of Pullum and Scholz's construal is that it implies that Chomskians are searching for a knockdown argument which will prove that language is innate. So, on their conception, a key premise of the APS is that a child has knowledge which he could not have learned from his environment. According to them, if this key premise is proven correct, then Chomsky has a knockdown argument for innateness. We have seen in Part 1 of this chapter that Chomsky does indeed use auxiliary inversion to support his claim that children know a rule of language that they could not have learned from their linguistic environment. Collins does not deny that Chomsky has sometimes argued in this manner. However, according to Collins, when Chomsky argues thusly, he is merely using the auxiliary inversion as an expository device to indicate the way linguists' reason. Chomsky did not intend to use it as a knockdown argument against the empiricist.

Critics of Chomskian nativism have expressed frustration with this mode of arguing. They argue that every time a candidate is presented as evidence for nativism, and is then shown to be inadequate, the nativist replies that *this* example was not the

real argument for nativism. Thus Cowie vents her frustration at this perceived dishonest mode of nativist argumentation:

The nativist- say, Chomsky- articulates a version of the argument. The empiricist counters it by pointing to its evidential short-falls and/or its failure to do justice to empiricism's explanatory potential. But no sooner is one rendition of the APS cut down than myriad other variations on the same argumentative theme spring up to take its place. For every non-obvious rule of grammar (and most of them are non-obvious), there is an argument from poverty of stimulus standing by to make a case for nativism. And for every such argument (or at least for the ones I have seen), there are empiricist counter examples of exactly the kinds we have reviewed in this chapter, waiting, swords at the ready, to take it on. (1999, 203)

When Cowie speaks of a line of rules being put forth and refuted by the empiricist, she is clearly thinking of Pullum and Scholz's version of the APS. Such postulated rules are the subject-auxiliary inversion, subject (verb-object) asymmetry, anaphoric one etc. The frustration which Cowie feels is that the subject-auxiliary rule was used by Chomsky on countless occasions, and by many others influenced by him. It is unquestionably used as *the* paradigm example of an innate rule,⁴⁷ so to be told that it is not the real argument is frustrating to say the least. However, Cowie massively overstates the strength of her position here. The paradigm example of auxiliary inversion is taken to indicate that a child has an innate preference for structure-dependent rules. Pullum and Scholz have shown that Chomsky has overstated the strength of his position by claiming that occurrences which would indicate that the structure dependence rule is the correct one are vanishingly rare. From the fact that Chomsky has overstated the strength of his position, it does not follow that the empiricist has proven that he is wrong. Pullum and Scholz themselves correctly claim that their data is suggestive at best, and they call for further studies into the PLD. Their research certainly is not a knockdown argument against nativism; it is rather a

⁴⁷ The innate rule is obviously structure dependence not the auxiliary inversion rule, which is of course not a rule of all languages.

timely reminder that, polemics aside, the case for nativism has not been proven. In this sense, Legate and Yang (2002) is an attempt to provide justification for the nativist research programme by situating it in a comparative setting. They consider a construction that both sides admit is learned (the use of null subjects) and analyse the PLD empirically to see how many examples of null subjects the child is presented with when learning the rule. They then compared the result of this with the amount of times the child is presented with the subject-auxiliary inversion. Based on their comparison, they claimed that the child was presented with less than half the amount of evidence in the case of auxiliary inversion. So they argued (incorrectly in my view) that the data the child was presented with was not enough for the child to learn the structure-dependent rule of question formation. I do not want to discuss Legate and Yang's paper here. I merely want to show that Cowie's claim that each innate rule postulated by Chomsky et al. has been refuted by attention to the PLD is false. On the contrary, all that has been shown is that more attention to the PLD needs to be given by both nativists and empiricists. Legate and Yang's paper can be seen as a nativist response to this request by empiricists such as Pullum.

Collins obviously would not deny that more research needs to be done on the PLD. He would presumably welcome Yang and Legate's attempt to answer the criticisms of Pullum and Scholz. However, he would also argue that Cowie and Pullum and Scholz have misunderstood the nature of Chomsky's APS and, that replies such as Legate and Yang's, while useful, concede too much to the empiricist by accepting their reconstrual of the APS. Using Pullum and Scholz's neat deductive reconstrual of the APS, Cowie seems to view nativists as merely putting forth an empirical premise into the argument, only to have it refuted, and to then put a different empirical premise in response to this and so on. Each time the empirical

premise is shown to be false, another one is added as quickly. Collins argues that this way of viewing the matter badly misconstrues things; on his view of the matter, the APS is not the deductively neat argument that Pullum and Cowie et al. think that it is.

Collins's reconstrual of the APS is less aesthetically pleasing than Pullum's deductive version. However, it does represent a type of APS which Chomsky has used from time to time. Nonetheless, I will argue that Chomsky primarily uses the APS which Pullum et al, critique.

Collins begins by noting certain features of our linguistic competence. He notes the obvious fact that all humans (bar congenital defect) acquire *a particular* language, while no other animal does. Furthermore, if you move children from their birth place to another country, they will end up speaking *a different* language, while all other animals will speak no language no matter where they are brought up. He acknowledges that these considerations do not militate towards nativism; on the contrary, they merely show that the language we do acquire must be acquired as a result of some innate species-specific machinery. This fact is merely a truism and is accepted by both sides of the debate. It in no way shows that innate domain specific knowledge is required for a child to learn his first language.

However, the above facts do indicate a problem which any linguistic theory worth its salt must solve. The problem in a nut shell is this; how do we construct a theory of language acquisition which is both descriptively and explanatorily adequate.

Collins puts it in the following way:

The descriptive adequacy, therefore, of a general theory of linguistic competence would appear to involve a delineation of the seemingly infinite variety of languages upon which a child may fixate. On the other hand, if our general theory is to be explanatorily adequate, then we need to explain how a child may fixate on any point in this infinity without any such point being favoured prior to the child's exposure to language. (2003, 30)

Any theory of linguistic competence needs to deal with this criterion. Collins correctly notes that these constraints do not of themselves tell us what (1) the child's initial state is, (2) what his final state is, nor (3) what data the typical child is exposed to which helps it move from the initial state to the final state. If there were only one language, for example English, we could answer questions (1) and (2) instantly. The child's initial state would be English and his end state would also be English. According to Collins, question (3) would also be answered, because the child would need no data to decide amongst languages, as there would only be one language that the child could represent.

Obviously there are more languages than English. In fact, if one considers English in different epochs, the English spoken by Chaucer, by Shakespeare, by Orwell then one must face the fact that English itself consists of more than one language. Rough estimates make it 7000 languages today⁴⁸. If one wants to explain a child's linguistic competence, then he needs to account for the fact that a child born in a different place or time could learn any of the 7000 languages spoken today. Furthermore, the child could learn any of the different languages spoken in other eras, or the various possible languages to be spoken in the future. So a descriptively adequate theory will have to account for children's linguistic competence which enables them to acquire the different types of possible or actual human languages. What is the initial state of the child that makes it possible for him to grasp any of these languages he is exposed to? If in order for the child to learn a language he needs the capacity to represent the rules of the language, then the more languages there are

⁴⁸ When I say that there are an estimated 7000 languages known today, I am speaking of E-languages. Given the difficulties of individuating E-languages which Chomsky has repeatedly discussed, it is unclear how to calculate how many different languages there are known at a given time or how many have been known or are possible to know. Internal to Chomsky's theories the question should be rephrased as how many I-languages can be derived from UG based on permitted parametric variation. Answering such a question is obviously impossible until we have a definitive worked out conception of UG.

which the child *can* learn, the more inclusive must be the child's initial ability to represent the grammar. The child will of course also need to have the capacity to represent the rules of possible human languages. So we will need finer and finer data in the particular child's environment to help the child decide which language he is supposed to learn. This data will also need to be so detailed that it stops the child from keying into other languages that it is possible for him to learn.

The difficulty with this approach is with the necessity of postulating richer and richer data to explain the child zeroing in on their grammar. The reason that this is a problem becomes apparent when we consider the data which a linguist has at his disposal when he is trying to discover the nature of UG or of a particular I-language such as English. The linguist has as much data on the grammar that he could wish for. He has the ability to reflect on it theoretically. He can compare the language with a variety of other languages. Yet, despite all of this data, the linguist still cannot discover the rules which govern the English language. Collins claims that if we argue that the child learns his language through data-driven learning we will be claiming that the child who learns English has enough data to figure out what linguists have been unable to figure out over the last 2000 years:

But here's the rub! The linguist has as much data on the grammar of English, say, as he could wish for, he also has the capacity to reflect on it, theoretically or otherwise, and the advantage of comparing it with data from other languages, but he still cannot figure out the grammar of English – that is *inter alia*, we have linguistics for! If, then, we content ourselves with the bland remark about nativism, we are led to think of the child who successfully acquires English as having enough data to figure out what self-reflective linguistic inquiry has been banging its head against for the last couple of millennia. Something is wrong. (ibid., 4)

He argues that the only explanation for the child achieving what linguists cannot achieve through thousands of years of inquiry is the postulation of children being born with innate apparatus:

What the child's innate equipment is required to do, it seems, is actively constrain its 'choices' as to what is part of the language to be attained. But no child is wired to target any particular language: the child can make the right 'choices' about any language with equal ease. This suggests that children must begin with 'knowledge' specific to language, i.e., the data to which the child is exposed is 'understood' in terms of prior linguistic concepts as opposed to general concepts of pattern frequency, say. If this is so, then we can see how a child may acquire a language even though the data itself is too poor to determine the language: the child needs no evidence for much of the knowledge it brings to the learning situation. In crude terms, children always make the right 'hypotheses' as a function of their genetic endowment. Thus, since the child can fixate on any language in the face of a poverty of stimulus about each language, and all languages are acquirable, children all begin with the same universal linguistic knowledge. This is the poverty of stimulus. (ibid., 5)

SECTION 2: THE STRUCTURE OF COLLINS' APS

So Collins's reconstrual of the APS is as follows:

- P1: Language is either acquired through data-driven learning or innately primed learning.
- P2: All human children acquire language.
- P3: No non-humans acquire language.
- C1: Therefore language is acquired because of a unique property of human children not shared with non-humans.
- P4: The range of languages it is possible for human children to acquire is infinite.
- P5: All linguists using data-driven learning have not discovered a complete grammar of one language.
- P6: All human children with less data available have acquired a particular language.
- P7: Therefore either human children are smarter than linguists or human children do not acquire language through data-driven learning.
- P8: Human children are not smarter than the linguists.
- C3: Therefore human children do not acquire language through data-driven learning.
- P9: If human children do not acquire language through data driven learning, then the fact that the child acquires a particular language as opposed to other possible languages cannot be explained through data-driven learning.
- C4: Therefore human children acquire their particular language through innately primed learning.

The first three premises of Collins's argument are correct and the conclusion is true as well. However, I have serious difficulties with Premise 4. Nothing in either Chomsky or Collins's argument has proven that the number of languages it is possible for humans to acquire is infinite. A more sensible claim for premise 4 would be that there

are an extremely large number of languages which it is possible for people to learn. Furthermore, it is difficult to see how this claim can be fitted into the overall structure of the argument. Or rather, it is obvious what role the claim is meant to play in the argument, but it is difficult to fit this role into our argument schema. The role it plays is that once we have shown that the child does not learn his language through data-driven learning, then it is difficult to see how the child arrives at the particular language he does, as opposed to the countless other possible languages he is capable of learning. The Collins/Chomsky solution is that the child is born with certain universal principles which are subject to parametric variations, and this explains the possible languages which humans can learn. So the child is born in the initial state UG and his experiences trigger various different parameters which results in the child arriving at his steady state, his I-language, i.e. English, French etc. If one took Premise 4 and Premise 8 out of the argument, it would still go through as valid because of Premise 1 and C3. Given that the overall argument could go through without P4 or P8 one may want to ask why the premises are in the argument in the first place. The answer is, that without the premise, our theory will not be explanatorily adequate, i.e., it will not explain both the diversity of languages acquired and the mode of acquiring them. So we will want our meta-argument to express that our object argument is designed to meet the criterion of descriptive and explanatory adequacy.

METACRITERION: An argument for an innate language faculty must match the criterion of descriptive and explanatory adequacy.

With our object argument and our meta-criterion in place, we have Collins's APS ready to evaluate. The argument has at least seven premises, some of which are uncontroversial, like P2, and some more controversial like P4. Some of the premises

are disjunctive, and may seem controversial because they leave out alternatives and assume that certain processes can only occur in one of two different ways.

Premise 9 and C4 are key aspects of the argument. Premise 9 states that if the child does not acquire language through data-driven learning, then without innate domain-specific knowledge, we cannot explain how the child arrives at the particular language he does as opposed to the countless other languages it is possible for him to acquire. This is certainly true. If the child does not learn from the PLD then there is no reason bar innate constraints that he would target the correct language. C4 states that the child does not learn his language from the PLD so must therefore key in on the correct language through innate domain specific knowledge. C4 is derived from P5-8 which states that using data-driven learning linguists working over 50 years have not converged on the correct grammar for English. Given that each child acquires English in a few years with much fewer data available to them, it follows that unless children are smarter than linguists they did not learn their language from the PLD.

Overall, the argument as set out by Collins is not very convincing. It does not amount to a deductive proof or a knockdown argument as Collins acknowledges. The argument aims to set out the facts of language acquisition which we need to build our theory around. For example, P5-8 does seem to indicate that unless children are smarter than the thousands of linguists working on generative grammar over the past fifty years, then they cannot be learning the language from the PLD. However, given that children born in different linguistic environments do arrive at different languages, the PLD is obviously a factor in how children learn their language. The tension between these facts of language acquisition is what a linguist needs to accommodate. However, to do this, one needs to set out what the structure of each I-language is, what they have in common, and where they differ. The theory of Principles and

Parameters, which states that children are born with a UG that consists of fixed principles some of which are subject to parametric variation, aims to accommodate this fact. On this theory, the different parameters are set by experience, while the universal principles are innate.

Hence, in Collins' view, the APS does not depend on every child lacking this or that datum. As Collins construes the APS, it depends on the fact that linguists have access to a much greater PLD than children do. Yet children can quickly arrive at the grammar of their language while linguists over generations have failed to isolate the correct grammar of any language. Nonetheless, Collins is not claiming that facts about the PLD are unimportant for the generative grammarian. Rather, he is claiming that we can only sensibly interpret the importance of each particular datum in the light of facts about UG and the particular I-language of a particular speaker. As set out by Collins, the APS is overcome by the postulation of a UG. This UG consists of invariant principles that a child is born with and that are subject to parametric variation, depending on the experiences of the child. Thus a principle of UG would be that all phrases have a head and a complement. The child is born knowing this. However, it is the child's experiences with their PLD which determine whether the phrases are head first (English) or head last (Japanese).

So the order of explanation would be the following. First, discover the structure of particular languages and try to ascertain what principles are shared by the different languages of the world. Second, discover the ways these languages differ from one another, and construct a theory in terms of parametric variations that can explain these differences. When one has the bones of the principles and parameters theory set up, one is then in a position to explain how this or that datum results in a particular construction being acquired. As presented by Collins, the APS is a set of

considerations which leads one to postulate a UG subject to parametric variation. The particular details are to be formulated within linguistic theory as the various different principles and parameters are discovered. Each discovery will either tell for or against the solution to the APS put forth by Chomsky et al.

Collins's version of the APS does not work. The disparity between the child's ability to learn from their PLD and the linguist's ability to construct an explicit grammar of this language need not be explained in terms of innate domain-specific knowledge. We do not need to claim that a two-year-old child is smarter than teams of linguists researching grammar over thousands of years. Nor do we have to claim that the child has more data available to him than the linguist. The difficulty with Collins's argument is that it equates an organism's ability to acquire a competence in x , with an organism's ability to form an *explicit theory* of his competence in x . It does not follow that because an organism has difficulty in forming an *explicit theory* of a particular competence x by extensively studying datum Y that competence in x cannot be acquired from datum Y . Collins's argument therefore fails because it unjustifiably equates an *explicit theory* of a competence with an *implicit ability* to acquire a competence. It is possible, for example, that children use unconscious statistical abilities which help them learn the rules of their language from our PLD. These statistical abilities may not be accessible to consciousness. Our ability to unconsciously detect patterns in our environment may outstrip our ability to construct explicit theories about these patterns. This bare possibility could turn out to be empirically false; however, whether it is or not is an empirical question. Collins's argument, as he stated it, gives us no reason whatsoever to hold that innate domain specific knowledge must be wired into the child.

SECTION3: WHICH APS DOES CHOMSKY WORK WITH?

Having argued that Collins's version of the APS is not a particularly strong argument, I now want to consider whether Collins' or Pullum et al.'s versions of the APS correctly characterize Chomsky's APS. While, Chomsky does seem to argue from the same general considerations which Collins has outlined above, he also joins these arguments with APS's of the kind that Pullum and Cowie consider. Any intellectually satisfying characterization of Chomsky's APS must explain why he felt it was not only necessary to argue from the general considerations like the ones Collins points to but also uses APS's like the ones Pullum et al. critiqued.

In his paper "Linguistic Nativism", Collins uses the Principles and Parameters model of language acquisition which Chomsky developed in the 1980's. When discussing the APS in this era, Chomsky constantly made unsubstantiated claims about the PLD. These claims need to be noted and outlined if we are to really understand Chomsky's APS. Pullum and Scholz (2002) refer to the auxiliary inversion as the experimental crux of the APS. They cite numerous different places that Chomsky and his followers have used auxiliary inversion as an example of the APS. Collins admits that Chomsky does indeed argue like this in various different places. However, he claimed that when Chomsky argues like this he is not making a claim about the PLD; but is rather setting up a challenge to the empiricist. Collins claims that Chomsky wants to ask the empiricist what is it about the child's PLD which helps him converge on the correct grammar, and why this data is insufficient for the linguist to learn the same grammar. In this section I will examine Chomsky's actual writing to see if this interpretation of his APS is correct.

Since the 1980's, Chomsky has been labelling the problem of language acquisition as Plato's Problem. He characterises this problem by quoting Bertrand

Russell's question 'How comes it that human beings, whose contacts with the world are so brief and personal and limited come to know as much as they do' (1986a, xxv). Chomsky argues that these questions arise in the particular sphere of language acquisition in the same way that they do in general epistemology. He claims that when it comes to language acquisition, the solution to the problem is to postulate innate knowledge. He calls the APS Plato's Problem because he feels that Plato's discussion of the child displaying knowledge of geometry which he has not previously been taught is a good instance of an APS. It is interesting how he characterises the situation of the slave:

This experiment raises a problem that is still with us: How was the slave boy able to find truths of geometry without instruction or information? (1986b, 4)

The key words here are '*without instruction or information*'; Chomsky repeatedly claims that children know various principles and rules for which they received no instruction or information. He further claims that this is the APS and it is overcome by the postulation of innate knowledge. In this context, whether a particular construction is in the PLD or not is extremely important. Likewise, it is important to determine whether the child gains instruction explicitly or implicitly through positive and negative reinforcement. On this construal, the APS does not appear to be merely a challenge to the empiricist but rather to be an explicit claim about the PLD which is either true or false. I recognise that Chomsky's vague sketch of what he thinks the APS is need not correlate with how he uses the APS in his linguistics. So in order to evaluate how Chomsky *uses* the APS as opposed to what he *states* the APS is, we will need to situate the APS within the context of him describing the rules of a language.

In his *Language and Problems of Knowledge* (hence forth LPK), Chomsky considers both English and Spanish in detail. He tries to distinguish what rules they share from those that exist only in one language and hence are presumably learned

from the PLD. Throughout his discussion, he makes claims about the nature of the PLD the child has available and the order that the child will learn the data in. One of the first examples Chomsky offers of Plato's Problem in LPK concerns a-phrases and reflexive pronouns in Spanish. He begins his discussion by illustrating a particular rule of natural language. He then considers how a child using analogical reasoning would apply this rule to other constructions. He claims that a child reasoning using analogy would create constructions which are incorrect by the lights of the native speakers of the language. So here Chomsky's arguments bear directly on Quine's model of language learning. Quine had claimed that analogy along with induction and reinforcement play a key role in language learning. However, Chomsky is here claiming that when the details of language are looked at closely, we see that a learning model based on analogy will make incorrect predictions about the type of sentences ordinary language speakers will find grammatical. He then claims that the child will not try out the false constructions which are derived by analogical reasoning only to receive negative reinforcement. He claims further that the child receives no data from his environment which helps him learn the correct rule. However, he concludes that the rule must be innate. This claim again runs contrary to Quine's views on how a child learns his language.

In LPK, pg 12-20 Chomsky illustrates what he believes to be a clear case of Plato's Problem. He begins by discussing simple sentences of Spanish, giving their direct translation in English, and a paraphrase of the translation in ordinary English.

The first sentences he discusses are:

(1) Juan arregla el carro.

'Juan fixes the Car.'

(2) Juan afetia a Pedro.

Juan shaves to Pedro.

'Juan shaves to Pedro.'

Chomsky notes that sentences (1) and (2) illustrate an interesting fact about a language such as Spanish. He points out that in Spanish, while an animate object of a sentence is preceded by a preposition 'a' (to), an inanimate object such as 'el carro' does not need a preposition before it. He claims that this feature of language is not shared by similar Romance languages such as Italian. He then goes on to consider more complex sentences involving causative constructions, which also feature the verbs 'afetia' and 'arregla'.

(3) Juan hizo (arreglar el carro).

Juan made (fix the car).

'Juan had someone fix the car.'

(4) Juan hizo (afeitar a Pedro).

Juan made (shave to Pedro).

'Juan had someone shave Pedro.'

It should be noted from above that the subject of the complement clause is unexpressed, and so is interpreted as someone unspecified. However, Chomsky notes the subject may be explicitly expressed as in (5)

(5) Juan hizo (arreglar el carro a Maria).

Juan made (fix the car to Maria).

'Juan had Maria fix the car.'

We can see the difference between the English and the Spanish versions of the proposition in (5). In Spanish the subject of the embedded clause is an adjoined propositional phrase (a Maria), whereas in the English sentence Maria appears before the verb. Chomsky asks us to try and construct an analogue to (5) using the phrase *afeitar a Pedro*, instead of *arreglar el carro*. Doing this we get (6):

(6) Juan hizo (afeitar a Pedro a Maria).*

Juan made (shave to Pedro to Maria).

'Juan had Maria shave Pedro.'

Here Chomsky notes that sentence (6), constructed on analogy with sentence (5), is an unacceptable sentence. So a child using the Quinean process of analogical synthesis would in this case construct a grammatically deviant sentence. However, this fact in

isolation tells us nothing about how a child learns this fact of Spanish. An analysis of children's PLD and a study of children's linguistic performance would be needed before we rule out a Quinean conception of language learning. Whether Spanish children try out successive '*a phrases*' in speech, only to receive negative reinforcement is a question which can only be answered by studying actual performance data, or through constructing experiments. Until then, the question of whether a child tries to construct a sentence like (6) based on analogy with (5) only to receive negative reinforcement remains an open question. Chomsky claims that the reason that sentence (6) is unacceptable is because in Spanish there is a rule which bars two a-phrases from appearing together. He then sums up what he thinks we have learned so far from this brief analysis:

Summarizing, we have general principles, such as the principle for forming causative and other embedded constructions and the principle of barring successive a-phrases; principles that admit some variation in interpretation, such as the embedded clause property; and low-level rules differentiating very similar languages, such as the rule that requires insertion of a in Spanish before an animate object. Of course, these levels are not exhaustive. The interaction of such rules and principles determines the form and interpretation of the expressions of language. (1988b, 15)

Having given this brief outline of some simple rules of Spanish, Chomsky discusses how the child acquires these rules. He claims generally that there are three factors to consider when trying to understand how a child acquires the rules of language: (1) the genetically determined principles of the language faculty; (2) the genetically determined general learning mechanisms; (3) the linguistic experience of the child growing up in a speech community. In relation to the rules he discussed above, Chomsky speculates that the rule of a-insertion before animate objects is an idiosyncratic rule of Spanish which is learned from experience. Given that a-insertion before animate objects is not a feature of other closely related romance languages, he holds that it must be learned from the PLD through processes which we

do not as of yet understand. He speculates further that the rule which makes (6) unacceptable, has its source either entirely in the language faculty, or in a combination of the language faculty and experience. He claims that the embedded clause property must be a parameter which needs some experience to be learned because it does not occur in all languages. He speculates further that such embedded clausal complements which occur are not learned but result from general principles of the language faculty. At no point in his analysis, does he offer evidence in favour of this interpretation. He merely asserts a series of propositions which are presumably meant to be taken on faith until he justifies them later in the text.

He then goes on to consider further examples. He asks us to change (2) ‘Juan afeitado a Pedro’, by replacing ‘Pedro’ with a reflexive element. There are he claims two choices for a reflexive: *se* or *si mismo*. He asks us to consider here just the first of these, and to replace *Pedro* with *se*.

(7) Juan afeitado a *se*.
 Juan shaves to himself.

However, (7) is not a proper sentence. Chomsky notes that the element *se* is what is technically called a *clitic*, a form that cannot stand alone but must attach to some verb. According to Chomsky, there is a rule of Spanish that moves *se* from its ordinary position as direct object of *afeitar*, attaching it to the verb, yielding:

(8) Juan *se* afeitado.
 Juan self-shaves.
 ‘Juan shaves himself.’

So the reflexive form corresponding to (2), would then be (8) rather than (7). Note here that on a Quinean account of language acquisition, the child would probably try (7) on analogy with (2) receive negative reinforcement, and somehow have to work out that (8) is the correct rule. Chomsky then asks us to combine the causative and reflexive constructions, replacing *Pedro* in (4) by the clitic *se*,

Yielding:

(9) Juan hizo (afeitar a se).
 Juan made (shave to self).

Chomsky notes that since *se* is a clitic it must attach to a verb, and that there are two different ways that this could be done: *se* could attach to ‘shave’ or to ‘made’. He notes that in all dialects of Spanish, it is normal to attach it to ‘made’, though only in some is it allowable to attach it to ‘shave’. He sticks to discussing the more common case, where *se* attaches to ‘made’; this is obviously a simplifying assumption though nothing of much importance attaches to the assumption for our present purposes. He goes on to note that (10) will be the correct transformation of (9):

(10) Juan se hizo (afeitar).
 Juan self-made (shave).
 ‘Juan had someone shave him (Juan).’

He notes that in (10) the embedded complement of the causative verb is subjectless, as in (3) and (4). But of course the subject of the complement can be explicit, appearing as an a-phrase. However, he argues that if the subject of the complement is say, *los muchachos* (the boys), we would expect to derive:

(11) Juan se hizo (afeitar a los muchachos).^{*}
 Juan self-made (shave to the boys).
 ‘Juan had the boys shave him (Juan).’

Unfortunately, while (10) is an acceptable sentence, sentence (11) is not. So a child trying to derive (11) based on the analogy with (10) will construct an unacceptable sentence. Again, a child trying to learn the rules of language using analogy will end up using a deviant sentence such as (11).

So what conclusion does Chomsky draw from these facts about Spanish? He notes first that the examples give rise to Plato’s Problem. He also claims that such facts show the hopelessness of claiming that language is acquired using analogy. He

goes on to make the following empirical claim about the acquisition of these facts for a Spanish child:

The question then is how speakers come to know these facts. Surely it is not the result of some specific course of training or instruction; nothing of the sort occurs in the course of normal language acquisition. Nor does the child erroneously produce or interpret the sentences (11) or (12) 'by analogy' to (10) and (5), leading to correction of this error by the parent or other teacher; it is doubtful that anyone has undergone this experience and it is certain that not everyone that knows the facts has done so. (ibid., 21)

Here we have in Chomsky's own words an example of Plato's Problem. The question we need to ask ourselves is whether *this* version of Plato's Problem should be understood the way Pullum and Scholz claim or whether it is merely, as Collins argues, a challenge to the empiricist. It should be noted that when discussing Plato's Problem in *this* context, Chomsky makes three unsupported empirical claims: (1) children do not construct new sentences like (11) using analogy and induction; (2) children do not incorrectly utter sentences with the structure of (11); (3) children are not corrected by their peers for constructing such utterances. If there is evidence to support these claims, Chomsky does not produce any. He merely argues that such things never happen, and we are presumably meant to take him at his word. If he wanted to establish that sentences like (11) are never produced, he would need an extensive corpus analysis to justify such a claim. He has never provided such an analysis. Short of corpus analysis we have no justification for the claim that children never utter such constructions. Chomsky would probably argue that we know that children do not erroneously produce examples like (11) and receive negative reinforcement. To justify this claim, he could cite Brown and Hanlon (1970) who have shown that correction for bad grammar is rarely provided, and when it is provided it rarely has any effect. However, recent studies on implicit instruction

undermine this claim. So Chomsky has given us no reason to assume that this particular APS works.

It is impossible to read this particular APS as a challenge to the empiricist. On the contrary, it is more accurately read as a supposed refutation of empiricism. If one did want to view it as a challenge, it seems to be a very strange challenge. The challenge could be construed as follows: Chomsky makes arbitrary unsupported claims about the PLD. The challenge is that the empiricist has to find evidence to refute Chomsky's arbitrary unsupported claims, and until such evidence is provided, we are to assume that Chomsky is correct. Such a challenge is clearly absurd. The burden of proof is obviously on Chomsky to provide evidence to support his claims, not to merely point out that empiricists haven't yet refuted his unsupported claims.

Here it could be argued that I am being a bit unfair on Chomsky. He does after all offer some evidence to support his claim: for example, the evidence from Crain and Nakayama (1987) and Brown and Hanlon (1970). However, such a defence of Chomsky is historically inaccurate. In his 1968 paper "Linguistic Contributions: Present"⁴⁹ Chomsky discusses how auxiliary inversion illustrated a particular instance of the Universal Rule that all sentences are structure dependent. Here Chomsky makes the following claims:

There is no a priori reason why human language should make use exclusively of structure-dependent operations, such as English interrogation, instead of structure-independent operations, such as O1, O2, and O3. One can hardly argue that the latter are more 'complex' in some absolute sense; nor can they be shown to be more productive of ambiguity or more harmful to communicative efficiency. Yet no human language contains structure-independent operations among (or replacing) the structure-dependent grammatical transformations. The language-learner knows that the operation that gives 71 is a possible candidate for a grammar, whereas, O1, O2, and O3, and any operations like them, need not be considered as tentative hypotheses...Careful consideration of such problems as those sketched here indicates that to account for the normal use of language we must attribute to the speaker-hearer an intricate system of rules that

⁴⁹ This paper is in Chomsky's book *Language and Mind* pg 21-56.

involve mental operations of a very abstract nature, applying to representations that are quite remote from the physical signal. We observe, furthermore, that knowledge of language is acquired on the basis of degenerate and restricted data and that it is to a large extent independent of intelligence and of wide variations in individual experience. (1972a, 54-56)

Here Chomsky is making untested claims about the child's PLD; he is also making unsupported assertions about the structure of *all* human languages. Chomsky claims the child knows that only 71 is a possible grammar whereas O1, O2, and O3 are not. Here he is implicitly making a claim about the child's linguistic performance. The only possible evidence that a child does not consider O1, O2, and O3 is that a child never mouths sentences structured according to the rules of O1 etc. Characteristically Chomsky has not offered any evidence. It is important to note that he is making these claims two years before Brown and Hanlon published their paper, and nineteen years before Crain and Nakayama's paper was published. So here Chomsky is making claims for which he has provided absolutely no evidence. If such claims are interpreted as a challenge to the empiricist, they are a poor challenge indeed.

Chomsky's APS arguments typically rely on claims that the child does not have access to this or that datum. It is claimed that if the child were learning a particular construction by analogy with previously heard constructions they would produce barred sentences such as x or y . He then argues that children never produce sentences such as x or y , and that therefore negative or positive reinforcement cannot play any role in learning a particular construction.⁵⁰ However he does not offer any performance data to indicate how children actually speak in particular circumstances. So his claim that children do not offer certain deviant sentences cannot be substantiated until the relevant research is done.

⁵⁰ In Chomsky's *Knowledge of Language* he repeatedly makes APS claims whose structures are similar to those outlined by Pullum et al. In this book as well he offers no empirical evidence to support his claims. See *Knowledge of Language* pgs:55, 62, 78, 90, 145-149

Chomsky does sometimes argue from general considerations in the way Collins does. However, when doing his linguistics, Chomsky typically makes claims about lack of reinforcement, limited fragmentary data, and how analogy and induction are insufficient to learn certain constructions. Over-all, neither version of the APS casts much doubt on empiricist models of language learning. The APS which Pullum et al consider do not tell us either way whether nativism is true or not. What the research done by Pullum et al. shows is that much more empirical data is needed if we are to discover how children learn their first language. Furthermore, it shows that Chomsky's lack of interest in performance data cannot be justified. If we are to construct an accurate theory of language acquisition, we will need to consider actual linguistic behaviour, and the circumstances of such behaviour occurring. Collins's version of the APS really offers no compelling reason to accept nativism. So linguistic nativism has not been justified through any of the poverty of stimulus arguments I have seen so far.

Before leaving this topic I discuss a recent defence of the poverty of stimulus argument which Chomsky has mounted. Chomsky has co-authored a paper with Berwick, Pietroski, and Yankama called "Poverty of Stimulus Revisited" (2011). This paper does not address the primary criticisms which are raised against the APS in this thesis. In fact, the content of their paper would lead one to believe that Chouinard and Clark, Pullum and Scholz, and Geoffrey Sampson do not exist.⁵¹ However, since this is Chomsky's most up-to-date defence of the APS I will consider it in detail. I use it to demonstrate that Chomsky's most up-to-date defence does not in any way meet the concerns which I have raised in this chapter.

⁵¹ While they do mention Pullum and Scholz's paper, they do not consider its impact for Chomsky's particular APS.

PART 3: CHOMSKY'S LATEST DEFENCE OF THE APS

In his (2011) paper ‘‘Poverty of Stimulus Revisited’’ Chomsky, Berwick, Pietroski, and Yankama offer a defence of the APS against recent criticisms. The paper is divided into five sections:

- (1) An introduction
- (2) A discussion of empirical foundations
- (3) A discussion of their minimalist solution to the empirical issues
- (4) A consideration of three alternatives to their approach
- (5) A conclusion

In Section 2 of their paper, they label what they take to be the central theory neutral facts which need to be explained. They claim that facts about auxiliary inversion in polar interrogatives which reveal the structure dependence of linguistic rules generalises to other rules of natural language. They discuss facts such as **constrained ambiguity**: consider the following four sentences:

- (6) Darcy is easy to please.
- (7) Darcy is eager to please.
- (8) Bingley is ready to please.
- (9) The goose is ready to eat.

They claim that children intuitively know that (6) and (7) are unambiguous. In (6), ‘Darcy’ is the object of the sentence and the sentence means that it is easy for others to please Darcy. In contrast, in (7) ‘Darcy’ is the subject of the sentence, and the sentence means that Darcy is eager to please others. Sentences (8) and (9) are ambiguous. ‘Bingley’ can be taken as the subject of the sentence or the object; it can mean Bingley is ready to please someone else, or that Bingley is ready to be pleased. Likewise ‘the goose’ can be taken as the subject or the object of the sentence; thus (9) can mean that the goose is ready to eat something else, or that, the goose is ready to be eaten. Further examples of constrained ambiguity are sentences such as:

- (10) The boy saw the man with binoculars.
- (11) The senator from Texas called the donor.

These sentences are two-ways ambiguous instead of three-ways ambiguous. They argue that such examples reveal the structure dependence of language, in the same way as polar interrogatives do. Chomsky et al. also point to the fact that some sentences have zero readings but are not mere word-salad, while other sentences which are word-salad declaratives can be turned into word salad interrogatives using auxiliary inversion. Having discussed the various different cases of **constrained ambiguity**, they note that they are concerned (at this point) with the knowledge *that* people have acquired, not with *how* such knowledge is acquired.

It should be noted that there are problems with their claims that they are only concerned with the knowledge acquired. First, their belief in the fact that certain sentences have zero interpretation, one interpretation and or two interpretations is obviously derived from tests which are done on people's intuitions of grammaticality. However, in order for such tests to be considered an accurate sample of people's competence, we need statistics to support them. I agree with their interpretations of the facts; however, neither my intuitions nor the intuitions of a few linguists can be used on their own to form the foundation of a linguistic theory. Such intuitions need to be justified statistically. We need statistics which support the claim that people of different ages, and different socio-economic backgrounds have intuitions of the acceptability and unacceptability of the sentences which are used to support the belief that constrained ambiguity is a fact of natural language. Such statistics need to make explicit any gradience of acceptability/unacceptability which occurs in different socio-economic environments, and different age groups. With this statistical background in place, they are then in a position to say whether constrained ambiguity is something that all speakers accept. Until such time as this is done, their supposed theory neutral facts which they claim must be explained by any linguistic theory, have

not been shown to be an actual fact of natural language. Furthermore, appeal to people's intuitions of grammaticality needs to be tested against performance data. Corpus analysis of child-adult interaction, adult-adult interaction, and the linguistic interaction of people from different socio-economic backgrounds needs to be tested. If we are to say that people have intuitions that x is the case, we need to demonstrate that they perform as though they have such intuitions. And if performance data and competence data are at odds, then performance data clearly trumps people's intuitions of how they believe they perform.

This discussion demonstrates that Chomsky et al. are not merely pointing out facts about language that any theory must explain, rather, they are in fact making claims about what people know which they have not justified by appeal to empirical evidence. The fact that they do not provide statistical tests to determine whether people of different ages and socio-economic backgrounds have the same linguistic competence demonstrates that from the outset they are presuming that the intuitions of a few linguists are shared across the board. So far from giving a theory-neutral description of the facts of natural language, they are, in fact, from the outset presupposing a particular model of the nature of language.

They go on to discuss the following sentences:

- (21) hiker, lost, kept, walking, circles.
- (22) The hiker who was lost kept walking in circles.
- (23) The hiker who lost was kept walking in circles.

They note that given (21) we would expect (22) to be the declarative as opposed to

(23). However they also ask us to consider the following case:

- (24) Was the hiker who lost kept walking in circles?

They note that even if we focus on (21) we still read (24) as the interrogative version of (23). From this fact they conclude that one way or the other, the auxiliary verb *was*

is associated with the matrix verb *kept-* and not *lost*. They claim that considerations of coherence alone should lead one to construct a sentence like:

(25) Was ((the hiker who- lost) (kept walking in circles))

as opposed to

(24) Was ((the hiker who lost) (-kept walking in circles))

They note that this shows that the relevant constraint trumps considerations of coherence. However, here again they are not merely stating theory-neutral facts. It is true that I share their intuition that (24) is the interrogative form of (23); however, my intuitions are obviously going to be contaminated by my research into various different APS's. For Chomsky et al. to draw the conclusion they want, they need statistics to back up their claim that people of all ages and all socio-economic backgrounds share the intuition, however they provide no such statistical evidence.

The next phenomenon which they consider is **constrained homophony**. They discuss the following sentence.

(25) Can eagles that fly eat?

(26) (Can (eagles that – fly) eat))

(27) (Can ((eagles that fly) (-eat)))

They hold that (27) reveals the correct structure of (25), not (26). Since children cannot see the bracketing of (25) Chomsky et al. ask how children can know that (27) reveals the correct structure and not (26)? To further elaborate this point, they go on to consider *do* replacing the auxiliary verb *can*, since *do* bears morphological tense (*did*) but is otherwise semantically null. So they indicate the actual position of interpretation with **dv**, and the logically coherent but incorrect position by **dv***, now using this notion freely to indicate constraints on ambiguity/homophony.

(28) (do (eagles that **dv*** fly) **dv** eat)

They claim that this notation is entirely descriptive and reveals that (28) is unambiguous. This, then, raises a poverty of stimulus consideration: how do children know that **dv*** is a barred interpretation but that **dv** is not a barred interpretation? However, yet again Chomsky presents a certain supposed fact as a theory-neutral description without presenting any evidence to support this claim. They have not provided any statistical evidence of people's acceptability judgements across ages and socio-economic backgrounds; nor do they present any performance data. So the whole of Chomsky et al.'s claims about the theory neutral empirical facts which any theory must deal with stands on an extremely weak foundation.

They go on to claim that other languages such as German and Irish respect the same constraints. Yet again they provide no empirical evidence to support this claim. They then consider further examples of these constraints and claim again (unconvincingly) that they are merely producing theory neutral facts which any theory must explain. They cite four constraints which must be met by any theorist wanting to explain these supposedly theory-neutral facts:

- (1) Yield the correct pairings, for unboundedly many examples of the sort described.
- (2) Yield the correct structures, for the purposes of interpretation, for those examples
- (3) Yield the correct language-universal patterning of possible/impossible pairings
- (4) Distinguish v-pairings from w-pairings, in part, while also accounting for their shared constraints.

They argue that if one cannot meet the constraints of 1-4 then one has not got an accurate linguistic theory which can explain the relevant linguistic data.

They then proceed to outline their own account of how these linguistic facts are best explained. They explain these facts in terms of their minimalist programme. Once they outline their minimalist alternative they then discuss three contemporary attempts

to explain the above facts using domain-general knowledge. They outline the three rival theories and demonstrate weakness in all three theories. Having satisfied themselves that they have refuted their rivals they proclaim that their minimalist explanation is the best explanation of the above-mentioned facts. They conclude by arguing that after fifty years their poverty of stimulus argument still stands.

The three empiricist alternative theories which they evaluate are:

(1) STRING- SUBSTITUTION FOR ACQUISITION (CLARK AND EYRAUD)

In brief, Clark and Eyraud following Zellig Harris postulate ‘discovery procedures’ for grammars. Their inference algorithm when given examples like (37a) and (37b) will correctly generate examples like (37c) and exclude examples like (37d).

(37a) Men are happy.

(37b) Are men happy?

(37c) Are men who are tall happy?

(37d) *Are men who tall are happy?

The method works by weakening the standard definition of syntactic congruence, positing that if two items u and v can be substituted for each other in a *single* sentence context, then they can be substituted for each other in *all* sentence contexts. C and E call this notion *weak substitutability*.

Chomsky et al. claim that this method fails for two different reasons:

(A) It fails for English even when restricted to only *strings* that a language generates.

(B) It does not address the original APS which depends on which *structures* a language generates.

(2) BAYESIAN MODEL SELECTION OF GRAMMARS (PERFORS ET AL)

In a 2011 paper Perfors et al. (henceforth PTR) wrote a paper which aimed to address the question of domain-general versus domain-specific theories of how natural language grammar is acquired. In their paper they considered Chomsky’s famous

example of auxiliary inversion which had been used as a paradigm example of an APS since the sixties. PTR argued that using a Bayesian model selection of grammars they could demonstrate that the structure dependence of language which was revealed by auxiliary inversion could be learned through a Bayesian model. The Bayesian model is a domain-general theory of acquisition, so the fact that it can learn the structure dependence of grammar purports to show that Chomsky's APS does not work. It supposedly reveals that language acquisition does not require domain specific-knowledge.

PTR's model argues for a notion of a "Bayes learnable" grammar. Their model specifies a hypothesis space consisting of three different grammar types and uses a notion of Bayesian probability to decide amongst them on the basis of a sample from the corpus⁵². The three different grammar types they propose are: (1) Flat grammars which generate strings of a corpus directly from a single non-terminal symbol S, (2) Probabilistic (right) regular grammars (PRGs), (3) Probabilistic context free grammars ((1) and (2) are structure-dependent grammars while (3) is a structure-independent grammar)(*ibid.*, 43). They construct a grammar of each type to generate the sentences of the corpus, and score each grammar with a Bayesian probabilistic matrix. They use the CHILDES corpus as data for training and evaluating the grammars of their respective types (*ibid.*, 43). From their tests, they discovered that Probabilistic Context Free grammars are better able to predict the corpus with a smaller grammar than their rivals, and they are better at handling new constructions not contained in the corpus (*ibid.*, 50). The only learning prior which PTR use is a preference for a shorter, more compressed hypothesis. And Clark and Lappin correctly note that this learning prior is clearly domain-general. So given, that PTR's

⁵² Here I am following Clark and Lappin's description of PTR on page 43-45 of their *Linguistic Nativism and the Poverty of Stimulus*.

model prefers the structure-dependent hypothesis over the structure-independent one, we have evidence against Chomsky's original APS. Contrary to what Chomsky claimed, a learner using a domain-general procedure, can indeed, learn the structure-dependent rule for natural language.

Chomsky et al. reply to this argument as follows:

But even if a Bayesian learner can acquire grammars that generate structured expressions, given child-directed speech but no additional language-specific knowledge, this does not yet make it plausible that such a learner can acquire grammars that exhibit constrained ambiguity of the sort illustrated in Section 2. In particular children acquire grammars that generate expressions in accord with specific structure-dependent rules that govern interpretations...The question is whether learners can induce that expressions are generated in these *human* ways.(2011, 19)

Here Chomsky et al. are claiming the issue is not whether domain-general procedures can key in on structure-dependent rules, but rather whether the domain-general procedures used by PTR can capture more complicated phenomena such as constrained ambiguity. Chomsky et al. claim that PTR's model does not include or suggest any hypothesis about how expressions are generated according to the language-specific constraints which they discussed above. They argue that if one wants to address the real poverty of stimulus argument, then one needs to address the full range of examples which they discussed in Section 2 of their paper, not merely the simple examples of polar interrogatives.

It could be argued that Chomsky et al. are being unfair to PTR here. Throughout his career, Chomsky has claimed that a particular datum cannot be learned from experience, so must therefore be explained in terms of innate domain-specific knowledge. Then when PTR construct a model which can generate the expressions without domain-specific knowledge, Chomsky et al. argue that this fact is irrelevant because there are some further facts about language which the model cannot capture. Here, again, we are back to Cowie's criticism, that every time a nativist has

an APS refuted by an empiricist; the nativist can simply point out some other non-obvious fact which he claims cannot be explained in terms of domain-general learning, and when this claim is refuted another example is manufactured on the spot. The real difficulty with this approach is that it shifts the burden of proof onto the empiricist, any non-obvious fact of language is automatically assumed to illustrate an APS, and the empiricist must refute this claim. However, the burden of proof should not be shifted this way. The burden of proof is on both the nativist and the empiricist. It should not be assumed that some complicated fact of language can automatically be explained in either a nativist or an empiricist manner. Such issues are entirely empirical and should be judged based on the ability of either side to construct accurate models to explain the relevant data.

(3) Learning from bigrams, trigrams, and neural networks. (Reali and Christiansen)
 Realli and Christiansen in their 2005 paper “Uncovering the Richness of the Stimulus” constructed models which aimed to test whether yes/no questions could be learned by domain-general procedures. They used three models: (1) A bigram statistical model, (2) A trigram statistical model, and (3) a simple recurrent network model. They used a child-directed speech as training data for the three models.

BIGRAM STATISTICAL MODEL

Realli and Christiansen computed the frequency word bigrams and then the overall sentence likelihood for any word sequence, even for previously unseen word sequences (ibid, 25). This sentence likelihood was then used to select between opposing test sentence pairs similar to *Are men who are tall happy- Are men who tall are happy*, the idea being that sentences with the correct auxiliary fronting would have greater likelihood than those with incorrect auxiliary fronting. Chomsky et al. note that in RC’s experiment, which was done on 100 test pairs, the bigram likelihood

calculation successfully chose the correct grammatical form 96% of the time. However Chomsky et al. cite the work of Kam, Stoynezhka, Tornyova, Fodor and Sakas (2008), which shows that the strong success is a result of contingent facts of English and not with the bigram model itself. They note that the model exploits the fact that *who* and *that* are homographs, which are unclear as to whether they are pronouns/relativisers (ibid., 26) When Kam et. al correct this bias, the performance of the models decreased significantly.

RC'S TRIGRAM MODEL

The trigram model uses a test similar to the one used by the bigram model. It furthermore exhibits a similar level of success. However Chomsky et al. argue that, like the bigram model, the trigram model achieves its success because of contingent facts of English and not because of the model itself. They construct experiments themselves to test this claim. These experiments show that the performance of the model drops significantly once the English-specific bias is accounted for.

LEARNING FROM SIMPLE RECURRENT NETWORKS

Reali and Christiansen constructed a further experiment using a Serial Recurrent Network to try and learn a particular construction. These Serial Recurrent Networks contained a hidden context layer. Reali and Christiansen trained 10 different networks on the Bernstein corpus, and then tested whether they could discriminate between grammatical versus ungrammatical minimal pairs (2011, 28). So, for example, they tested whether the networks could correctly discriminate between:

- (1) Is the boy who is hungry nearby?
- (2) Is the boy who hungry is nearby?

Reali and Christiansen recoded the actual words into 1 of 14 possible parts of speech categories, e.g. Det (the), N (boy), PRON (who), V (is), ADJ (hungry), PREP (nearby) etc. (ibid, 28). Chomsky et al. note that Serial Recurrent Networks output a distribution over possible predicted outputs after processing each word. Reali and Christiansen tested their networks by providing the part of speech prefix of test sentences up to the point at which grammatical versus ungrammatical divergence would occur. They then checked the predicted output of trained networks over all word categories to see whether the network activation weight was assigned the grammatical continuation as opposed to the ungrammatical one. Reali and Christiansen confirmed that the grammatical continuation was an order of magnitude higher than the ungrammatical one. In other words, the Serial Recurrent Network was able to predict the correct (1) as opposed to the incorrect (2). Reali and Christiansen take this as evidence that their Serial Recurrent Network can learn the rule for structure dependence using this connectionist model.

However Chomsky et al. (following on from the work of Kam et al.) suggest that Reali and Christiansen's results may result from simple brute statistical facts:

In other words, one might wonder whether the success of the SRNs in selecting, for example, V as opposed to ADJ as above might also be attributable simply to "brute statistical facts". Kam et al's and our findings above suggest that bigram information alone could account for most of the statistical regularity that the networks extract. (ibid, 29)

Chomsky et al. tested this claim by analysing the Bernstein corpus to test whether there was a difference between the number of times a PRON- V occurs and to the number of times that a PRON-ADJ occurs. They found that the PRON-V occurs 2,504 times in the corpus, while the PRON- ADJ occurs 250 times in the corpus. So they claim that using a bigram statistical model they can easily predict the next occurrence of the grammatical sentence. They make the following point:

Since SRNs are explicitly designed to extract sequentially organised statistical patterns, and given that the is-is question types can be so easily modelled by sequential two-word patterns, this is not at all surprising. Indeed, it is difficult to see how SRNs could possibly fail in such a statistically circumscribed domain. (Ibid., 32)

They then go on to note that it remains to be seen how Serial Recurrent Networks will deal with a more complex interrogative such as:

(1) Is the boy who was holding his plate crying?

This example is of the sort explained by Crain and Nakayama. Here the matrix verb *is* differs from the relative clause auxiliary *was*. Chomsky et al. note that until such time as Reali and Christiansen can construct a model which can handle cases of this kind, we can conclude that the Serial Recurrent Network results are far from compelling. However this reply of Chomsky et al. is dubious, because, as I discuss now, Reali and Christiansen solve the APS which was raised by Chomsky in his 1975 book *Reflections on Language*.

In that book Chomsky writes:

We gain insight into UG, hence LT (H,L), whenever we find properties of language that can reasonably be supposed not to have been learned. (175, 30).

The example which Chomsky uses to illustrate this point is

(1) The man is tall – is the man tall?

(2) The book is on the table – is the book on the table?

Chomsky notes that a scientist will observe that children form questions in the ways indicated by (1) and (2). The scientist may form the following hypothesis to explain the above fact:

Hypothesis 1: The child processes the declarative sentence from its first word (i.e., from ‘left to right’), continuing until he reaches the first occurrence of the word ‘is’ (or others like it: ‘may’, ‘will’, etc.); he then proposes the occurrence of ‘is’, producing the corresponding question (with some concomitant modifications of the form which need not concern us) (ibid, 31).

He then argues that this procedure will lead the scientist into making incorrect predictions when it comes to more complicated sentences. He asks us to consider the following sentences:

(3) The man who is tall is in the room- is the man who is tall in the room?

(4) The man who is tall is in the room- is the man who tall is in the room?*

Obviously a scientist who was using hypothesis (1) would generate the incorrect sentence (4). Chomsky claims that a scientist will note that children never make mistakes like (4) and will therefore conclude that hypothesis (1) is incorrect. A reasonable scientist, he notes, will therefore try out a different hypothesis. Hypothesis (2), according to Chomsky, will be a structure-dependent hypothesis which analyses words into phrases. This hypothesis will differ from the structure-independent hypothesis which merely involves analysing words into the property of earliest defined word sequence.

He further argues that, by any reasonable standards, hypothesis (1) is simpler than hypothesis (2). Yet children supposedly unerringly use the structure-dependent hypothesis as opposed to the structure-independent hypothesis. Chomsky makes four points: (1) the child will never experience constructions which are relevant to helping him learn the correct rule; (2) the child never makes mistakes like sentence (4); (3) the child is not trained to learn the correct rule; (4) the correct rule is more complex than the incorrect one. Based on these considerations, Chomsky argues that the structure-dependence rule must be innate. Now I do not want to repeat the various arguments against this view that I have already voiced earlier in the thesis. However, in light of Chomsky' et al's new paper, there are some new points which need to be clarified. Reali and Christiansen managed to construct a Serial Recurrent Network which could learn the *is-is* construction through training. This demonstrates that Reali and

Christiansen's model has solved the original poverty of stimulus model which was raised by Chomsky in 1975. Chomsky et al. ask whether the model can handle the types of constructions discussed by Crain and Nakayama's 1987 paper. This is an interesting question. However, the fact that Reali and Christiansen have not answered it does not take away from their achievement. They have demonstrated it is possible for a domain-general learner to grasp a linguistic rule that Chomsky in 1975 claimed could only be acquired through domain specific learning priors. This fact is important. It illustrates Cowie's point of how easy it is for a nativist to manufacture a new APS in the face of a refutation of their original claim.

In this chapter, I have evaluated whether Chomsky's APS has refuted Quine's conception of language acquisition. By reviewing the best arguments which have been put forth by Chomsky in defence of the APS, I concluded that the APS has not refuted Quine's conception of language acquisition. However, the evidence which I have reviewed does not indicate which conception of language is the correct one. There is much research which needs to be done before we can decide whether the nativist or the empiricist conception of language acquisition is the correct one. I conclude by proposing future experimental research which may help to decide between the two different conceptions of language acquisition.

CHAPTER 5: WHERE DO WE GO FROM HERE?

PART 1: NATIVISM AND ANTI-NATIVISM

The debate between nativists about language acquisition (those who think it requires an innate domain-specific language faculty), and non-nativists (who accept innate priming but think it is domain general) tends to become extremely polarised. Thinkers on both sides tend to view the other as either intellectually dishonest or ignorant of certain obvious facts. This type of aggressive attitude has no place in an intellectual debate. Since the evidence does not as of yet indicate which conception of language acquisition is the correct one, an open-minded research programme should be advocated, not a tribal adherence to the dogmas of a particular group.

Anti-nativists claim that the following pattern is typical of nativist reasoning. A confident assertion is made that something cannot be learned either through induction or by trial and error. This assertion is then shown to be false by empirical evidence. The nativist replies to this refutation by claiming that the just refuted example was not vital for his theory and uses another example to illustrate the APS. A new example is then refuted and then the process begins again. Anti-nativists, for example, point to the fact that Chomsky claimed that it is mathematically impossible to learn certain rules of language statistically or by trial and error, only to have mathematicians and computer programmers show this to be false. They have also pointed out that Chomsky claimed that sentences which illustrate the rules for auxiliary inversion are not in the PLD, only to have empirical research show that he was incorrect about this as well. Chomsky and his followers typically reply to this refutation by claiming that this particular datum is not important because there are other constructions which children know which they have not encountered in the PLD. This new construction is then tested and shown to occur in the PLD as well. As

Cowie put it, Chomskians can claim that any non-obvious rule of language is underdetermined by the data of experience and as soon as this is refuted merely pick a different rule. Such a situation is obviously unsatisfactory.

Cowie's point is that it is just too easy for a nativist to simply claim that a particular construction is innate. On the other hand, nativists complain that anti-nativists simply do not try to construct an explicit theory of linguistic competence. While nativists are trying to construct a complex theory of linguistic competence, anti-nativists typically vaguely gesture in the direction of domain-general learning procedures. Such vague gestures typically do not attempt to account for the actual facts of linguistic competence; rather, they merely seek to show that this or that particular construction can either be learned mathematically or appears in the PLD. When they have shown that this or that construction can be learned by experience, they take this to have shown that the nativist is wrong on this point. However, the nativist correctly points out that, from it being shown that they are wrong that such and such a construction is innate, it does not follow that some crude empiricist model of language learning is tenable. If the anti-nativist wants to prove that some crude empiricist model of language learning is possible, they need to construct and test a model which is adequate to explain our current linguistic data. An empiricist will need to account for the actual rules of language as discovered by linguists, and demonstrate how such rules are learned from experience. Merely pointing out weakness in the nativist's theory and gesturing in the direction of learning by analogy, induction and reinforcement is not enough to establish an empiricist model of language learning.

The nativist and anti-nativist camps each make some legitimate points about each other's projects, although in some respects each goes too far and, in a manner, distorts the facts. In his "Linguistic Nativism: As We Were", Collins correctly points

out that it is a mischaracterisation of the project of generative grammarians to claim that they are postulating innate rule after rule in a desperate attempt to refute the empiricist. On the contrary, if one observes how, for example, Chomsky's linguistic theory has developed from his Aspects phase to his Minimalist Programme, one sees not innate rule after rule being postulated, but rather an effort to reduce rules to more general principles in order to achieve greater explanatory adequacy. So Cowie's criticism in some ways involves a caricature of the generative project. However she is correct that if one looks at the history of generative linguistics, one finds nativists pointing to supposed rules which cannot be learned from experience. One also finds that upon being shown that such rules are in fact learnable, the nativist shrugs his shoulders and claims that the example was not that central to his project. Furthermore, Cowie is surely correct that it is just too easy for nativists to claim that any particular complex construction cannot be learned from experience. Time and time again throughout his entire corpus of writing, Chomsky has claimed that a child does not produce this or that construction, or does not experience this or that construction in his PLD without providing any evidence to support such claims. It is surely a minimal requirement of intellectual responsibility that, if a theorist is making claims about people's linguistic behaviour or their PLD, they provide evidence to support such claims. Furthermore the stronger the claims being made, the stronger the evidence required. All of this should be truistic.

As I said above, nativists usually reply to these criticisms by claiming that empiricists have not provided a viable alternative to their model of language acquisition. However, whatever the fortunes of an empiricist model of language acquisition, Chomsky has no justification for making unsupported claims about people's linguistic behaviour and their PLD. If we have not yet sketched a satisfactory

empiricist account of language acquisition, it does not follow that we must accept an unsubstantiated nativist version instead. In this situation, the correct response is to perhaps try and construct a model of language acquisition which is neither nativist nor empiricist.

In the context of this thesis which aims to evaluate whether Quine or Chomsky's theory of language acquisition is correct, I will obviously not consider whether a third option offers a better account of language acquisition. At this point in time, the evidence is not strong enough to decide whether an empiricist alternative to nativism can be sketched. In his 1957 review of B.F. Skinner's book *Verbal Behaviour*, Chomsky is largely credited with destroying behaviourism. Quine offered a different behaviouristic account of language in his book *Word and Object*. Most philosophers believe that Chomsky has refuted this conception of language acquisition. This view is mistaken. Chomsky correctly showed that IDT cannot be viewed as fatal to linguistics because the IDT is merely a form of UD, and since UD is not considered fatal to physics, by parity of reasoning it should not be considered fatal to linguistics. Nonetheless, from this fact it cannot be inferred that Quine's behaviouristic account of language acquisition is false. Chomsky charges that Quine's behaviourist account of language acquisition is vague and does not spell out in any detail how certain linguistic constructions are learned. However, what Chomsky fails to note is that Skinner's *Verbal Behaviour* and Quine's *Word and Object* were intended to sketch out the way to study the phenomenon of language as objectively as possible. They were meant as guides to future research into language. But because Chomsky's rhetorical arguments won over so much support, most of this behaviourist research was not done. I suggest that in light of the fact that Chomsky's central argument against empiricist/behaviourist conceptions of language acquisition; his

poverty of stimulus argument, have not yet been justified by empirical research, we should seek to test the empiricist alternative to Chomskian linguistics. Future research should aim to give Quine's model a run for its money in the way Chomsky's has been over the last fifty-five years.

Some aspects of Quine's theory of language acquisition have been tested. In Chapter 3 I discussed an experiment done by Soja, Spelke and Carey which shows that Quine's claim that children only develop an ontology after they have mastered the syntax of quantification is false. The research done by Markman, Spelke et al, shows that a child's pre-theoretic ontology constrains the language learning process from the start, rather than it resulting from the process of language learning. So some aspects of Quine's views on language learning have been shown to be false. Nonetheless, there are other aspects of his views which still remain to be tested, including his views on the role reinforcement and trial and error play in the child learning his first language. We can no longer let Chomskian rhetoric decide the issue. The question needs to be tested empirically.

PART 2: SOME TESTS FOR THE FUTURE: QUINE

We saw in Chapter 2 that Quine defined a language as a set of dispositions to emit overt behaviour in certain determinate circumstances. Chomsky rejected this characterisation by claiming that it is impossible to construct a theory of the probability of certain sentences being spoken in certain determinate circumstances. The obvious way to reply to Chomsky on this point is to provide him with such a theory, but as we all know no such theory has been supplied. The question now arises how we can test Quine's claims about the probability of an utterance being spoken in certain determinate circumstances.

Analysis of linguistic corpuses has provided ample data to test Chomsky's claims about how people speak and about the data the child learner is exposed to. However, the above corpuses are not much use in evaluating Quine's claims about verbal behaviour. While such corpuses do in some special cases indicate a conversational exchange, they do not usually indicate the circumstances of the spoken utterance, nor in any cases do they reveal the history of such spoken utterances in various different circumstances. To analyse Quine's claims, what is needed is to record the history of the child's linguistic experiences up until a certain age. It is also necessary to record the history of the child's efforts to speak, and the history of the parents reactions to this speech, for example, whether they use implicit correction, repetition of words, etc.

The utility of such video recordings for studying whether Quine's behaviourist analysis of language is correct is obvious. Nonetheless there are severe limitations on this kind of approach. First, very few of these recordings have been done, so any data obtained will apply to the one or two children recorded. Second, unless done in various different cultures, such recordings will tell us little about whether they are valid for all human language learning. Third, even if we are to analyse the data in sufficient detail, for example, to check all the utterances of 'Mama' used by the child up to the age of four,⁵³ note the objective circumstances of the word being used, and the reaction (or lack of reaction) by other peers, we cannot do it in real time because such observation would take at least three years.⁵⁴ More than likely, it would take four or five times as long as the three years to analyse the data. So a researcher after twelve years of analysis would have a study which would be valid for one person only. There are ways around this difficulty. One could use computer software to pick out all and

⁵³ The age of four is picked because most children begin school at that age

⁵⁴ Three years because typical children do not begin to utter words until at least the age of 12 months.

only the circumstances where certain key words such as 'Mama' are used. Using this technology, we could analyse the circumstances of utterances much quicker than analysing the whole recording.

If such technology were available, then we would have an approach which would at least admit a test. We could use our recording and software to analyse all the times the child uses the word 'Mama' from aged one till three. After studying the circumstances under which the child used the term and received reinforcement, we could construct a theory about the probability of the child saying the word 'Mama'. We could then make our predictions and we would have the child's recorded utterances from three to four to test our prediction.

However, even if our predictions in this circumstance did turn out to be accurate, they would only apply to one person from a particular culture. The test would need to be replicated on different children across different cultures. One of the main difficulties would be receiving ethical approval to conduct such studies. A further difficulty is that while varying the children and cultures to test would be something which could be achieved, studying other age groups would be almost impossible. Prior to going to school, young children spend most of their time at home with their family, so recordings of their utterances in their house would be a good sample of their linguistic behaviour and reinforcement. However, once a child starts to attend school, recording of their utterances becomes virtually impossible, as children spend more and more of their time at school or out playing with friends, and attaining permission to record their utterances in circumstances which involved other people would be extremely difficult, not to mention the technological difficulties in recording children as they move about the world every day.

Nonetheless, if the experiment on children below the age of five could be successfully performed, it would be a step in the right direction in determining whether Quine's conception of language as a disposition to overt behaviour in certain determinate circumstances was, indeed, correct. It would still be an open question whether the explanation applied to adults who have achieved full linguistic competence, but it would at least be a starting point to studying linguistic behaviour.

Quine claimed that children formed different sentences using a process he called analogical synthesis, one form of which is analogical substitution. In analogical substitution a child learns to form the sentence 'My foot hurts' based on an analogy with 'My hand hurts' when the child learns that 'hand' can be substituted for 'foot'. Quine claimed that there are many different forms of analogical synthesis with analogical substitution being only one amongst them. How exactly these analogical syntheses are used by the child, Quine claims, is a matter of the child's long-forgotten history of acquisition. In 1982, when discussing analogical synthesis, the Quinean scholar Roger Gibson made the following claim:

Learning by analogic synthesis, on the other hand, is learning by analogies and requires further innate endowments. The innate mechanisms of such learning are little understood but nonetheless necessary for language learning. (1982, 195)

Thirty years later and these innate mechanisms are still little understood. In fact, they have not been specified in any greater detail in the thirty years since Gibson wrote. There are two reasons for this lack of progress. One is that Chomsky's polemics have more or less destroyed behaviourism as a research program. Even severe critics of Chomsky like Pullum, Sampson and Cowie are at pains to emphasize that they are not behaviourists. The other reason is that Quine and philosophers impressed by Quine have not spelt out the process of learning by analogy in any detail, and as a result there is no model to test as of yet.

If a Quinean model of language acquisition is to be taken seriously, then it needs to construct an explicit model of how analogy works in the production of new sentences. This model should be of the same explicit form as the generative grammars constructed by Chomsky et al. And it will obviously be constructed using data from developmental psychology showing at what ages children develop their different linguistic skills. It needs to account for how the child got from one developmental stage to the other using analogy, induction, etc. It will also be necessary to quantify which aspects of language are acquired through analogical reasoning, which through inductive abilities, and how much of a role reinforcement plays. Furthermore, such a model would need to account for the PLD which the child is typically exposed to. I have criticised Chomsky throughout this thesis for making claims about language acquisition which are unsubstantiated by empirical facts. However, when it comes to analogical synthesis, Quine is also making claims which are unsupported by the empirical facts until they are more explicitly constructed and tested.

It has been over fifty years since *Word and Object* was first published. Yet Quine's vague claims about sentences being constructed by the process of analogical synthesis have not yet been explicitly modelled nor have they been tested. If the Quinean model of language acquisition is to be taken seriously as a rival to Chomsky's, then the model needs to be made explicit and needs to be rigorously tested.

Quine speaks of a child being reinforced by his peers either positively or negatively when learning his first language. These behaviourist claims need to be tested, not merely asserted. In Chapter 4 I reviewed some experimental work which suggests that children do receive negative evidence which indicates that a construction is ungrammatical. A Quinean could claim that this implicit correction

which plays a role in a child learning his language is a type of negative reinforcement. However, the issue of how much negative or positive reinforcement children receive needs to be determined on the basis of detailed observations of children's actual speech in particular contexts. The type of experimental research I discussed in reference to the probability of certain utterances being used could help in determining exactly how much reinforcement children actually receive.

PART 3: SOME TESTS FOR THE FUTURE: CHOMSKY

Chomsky's research programme has been rigorously studied over the last fifty-five years. Recent experimental tests are revealing that some of the central claims of Chomsky about negative and positive evidence are actually false. Chomsky has always been critical of behaviourists for not realising that behaviour is merely the data we use to construct our explanation, not the explanation itself. Upon reading this, one would think that Chomsky uses mountains of behavioural data to construct his linguistic theory. Instead, all he uses are speaker's intuitions of grammaticality and ignores the actual linguistic behaviour of a typical language user. Clearly, if one aims to construct a theory of competence, one needs to consult actual linguistic performance, not just the intuitions of typical speakers. Furthermore, if Chomsky is constructing mathematical models to explain the intuitions of linguistic speakers, then the data he is working from needs to be published with these models. Chomsky needs to make explicit the age group of the people whose intuitions are being tapped, their nationality, the number of people interviewed, etc. Furthermore, he should provide a statistical analysis of the results of these tests. If there is gradience in people's

intuitions of grammaticality, then this gradience needs to be represented statistically. Such statistics should be represented on a scale which shows how (if at all) people's intuitions differ as they move along certain developmental stages. It is only after he has presented this data that Chomsky should begin to construct a mathematical model of the grammar people implicitly use. If generative linguistics is to become an actual science, it needs to be much more explicit about the sources of its data; otherwise, its findings cannot be taken to reveal actual facts about people's linguistic competence.

Even if Chomsky does provide statistics of people's linguistic intuitions, and constructs a mathematical model of the grammar that the intuitions imply, this model will need to be correlated with the facts of linguistic performance. Chomsky typically appeals to idealisation as a reason why performance is not considered in his theories. However, when it suits Chomsky, he has no problem making unsubstantiated claims about people's linguistic performance. Thus, for example, he states that children never try out structure independent rules when learning their first language. He argues that people do not usually receive or use negative feedback from peers for grammatical mistakes. Such claims need to be backed up with experimental research⁵⁵ and corpus analysis. It is only after Chomsky provides statistics to back up his claims about people's intuitions, and analyses in detail the PLD, and people's actual linguistic performance, that his claims for Universal Grammar can receive proper empirical support.

⁵⁵ Chomsky will sometimes use experimental research such as Crain and Nakayama's [1987] and Brown and Hanlon's [1970]. However he typically makes claims about performance independent of any evidence and will, after the fact cite experiments if they chime with his intuitions. What is needed is Chomsky to either have experimental research at hand to back up each claim about performance and the PLD or to conduct such research if he is going to make such a claim.

PART 4: NATURALISED EPISTEMOLOGY

In Chapter 1 we discussed Quine's and Chomsky's overall philosophical projects. Both thinkers are naturalists and tell a naturalistic epistemological story. They are concerned with explaining how people go from their meagre input (data of experience) to their torrential output (total theory of the world). Chomsky postulates various different faculties of the brain (science-forming faculty, language faculty, etc) which interact with each other and experience to form our total theory of the world. Quine claims that we use induction, analogy, and reinforcement as tools which help us learn language and learn to speak of the world.

Chomsky's science-forming faculty has no real evidence to support it. His language faculty is an interesting hypothesis which needs to be formulated more scientifically and tested more rigorously if it is to be accepted. Quine has a theory of language acquisition which is at present extremely vague and needs to be more explicitly characterised before it can even be fruitfully tested. At present, much more research is needed if we are to decide between the two competing pictures of how people begin to form theories and speak about the world. But much progress has certainly been made over the last fifty years.

In 1969 when Quine spoke about epistemology being a branch of empirical psychology, he set off a series of debates which continue to this day. Traditional epistemologists are concerned with whether psychology can account for things such as justification, while naturalistic epistemologists spend most of their time trying to reply to such critics. However, while debating with such critics is vitally important, naturalists who spend all their time engaged in such debates risk operating with an inadequate naturalistic conception of epistemology. It is important that our naturalistic

picture is an accurate representation of how exactly people go from stimulus to science, and not a mere caricature of this process. Quine and Chomsky are naturalised epistemologists but they hold different views on how people go from stimulus to science. If Naturalised Epistemology is to become anything more than a buzz word we need to decide between these competing pictures. We also need to see how these pictures compete with other naturalistic conceptions such as Piaget's Genetic Epistemology. The type of experimental research that I argued for earlier in this chapter will help us to resolve some of these debates. If Naturalised Epistemology is truly scientific, it needs to be constantly revised in light of new experiments, and to be constantly proposing new ways of testing its claims. Through this constant process of testing, naturalised epistemologists can form a much more accurate representation of how people's theories of the world evolve. While constantly improving their own naturalistic epistemology, they can also continue to engage in debates with traditional epistemologists.

BIBLIOGRAPHY

Works by Chomsky

- Chomsky, N. 1957. *Syntactic Structures*. The Hague: Mouton.
- Chomsky, N and Israel Scheffler. 1958 "What is Said to Be" *Proceedings of the Aristotelian Society* 59 pg 71-82
- Chomsky, N. 1959. *A review of B.F. Skinner's Verbal Behaviour*. *Language* 35. Pg 26-58
- Chomsky, N. 1964. *Current Issues in Linguistic Theory*. The Hague. Mouton.
- Chomsky, N. 1965. *Aspects of a theory of Syntax*. Cambridge. Mass. MIT Press.
- Chomsky, N. 1966 *Topics in the Theory of Generative Grammar*. The Hague. Mouton.
- Chomsky, N. 1969. 'Linguistics and Philosophy' In S.Hook ed *Language and Philosophy*. New York: New York University Press.
- Chomsky, N. 1971 *Problems of Knowledge and Freedom*. New York: Pantheon.
- Chomsky, N. 1972a *Language and Mind*. New York: Harcourt Brace Jovanovich
- Chomsky, N. 1972b *Studies on semantics in generative grammar*. The Hague. Mouton.
- Chomsky, N. 1975a. *The Logical Structure of Linguistic Theory*. New York. Plenum. Excerpted from 1956 revision of 1955ms. Harvard University and MIT. Chicago: University of Chicago Press 1985.
- Chomsky, N. 1975b *Reflections on language*. New York: Pantheon.
- Chomsky, N. 1977. *Essays on form and Interpretation*. Amsterdam: Elsevier North-Holland.
- Chomsky, N. 1980a. *On Binding*. *Linguistic Inquiry* 11, 1-46.
- Chomsky, N. 1979. *Language and Responsibility*. Sussex: Harvester Press.
- Chomsky, N. 1980a. *Rules and Representations*. New York: Columbia University Press.
- Chomsky, N. 1980b. *On Cognitive Structures and Their Development*. In M Piattelli-Palmarini ed *Language and Learning: The Debate between Jean Piaget and Noam Chomsky*. London: Routledge and Kegan Paul, pg 35-52
- Chomsky, N. 1980c. *Rules and Representations*. *Behavioural and Brain Sciences*, 3/1, pg1-15.
- Chomsky, N. 1981a. *Lectures on Government and Binding*. Dordrecht: Foris.
- Chomsky, N. 1981b. *Response to queries*. In Longuet-Higgins. Lyons, and Broadbent. 1981.
- Chomsky, N. 1982. *Some concepts and consequences of the theory of government and binding*. Cambridge, Mass: MIT Press.
- Chomsky, N. 1986a. *Barriers*. Cambridge, Mass: MIT Press.
- Chomsky, N. 1986b. *Knowledge of Language*. New York: Prager.
- Chomsky, N. 1987. *Response*. *Mind and Language* 2, 193-197.
- Chomsky, N. 1988a. *Generative grammar, studies in English linguistics and literature*. Kyoto University of Foreign Studies.
- Chomsky, N. 1988b. *Language and Problems of Knowledge: The Mangua Lectures*. Cambridge MA: MIT Press.
- Chomsky, N. 1990a. "On formalization and formal linguistics". *Natural Language and Linguistic Theory* 8, 143-147
- Chomsky, N 1990b. *Language and Mind*. In D.H. Mellor ed, *Ways of Communicating*. Cambridge: Cambridge University Press.
- Chomsky, N. 1991a *Prospects for the study of language and mind*. In kasher 1991a.

- Chomsky, N. 1991b. Linguistics and cognitive science: Problems and mysteries. In Kasher 1991b pg26-53
- Chomsky, N. 1991c. "Linguistics and adjacent fields: a personal view". In kasher ed *The Chomskian Turn*. Oxford: Blackwell pg 5-23
- Chomsky, N. 1991d. *Some Notes on economy of derivation and representation*. In R. Freidin, ed *Principles and Parameters in comparative grammar*. Cambridge, Mass. MIT Press.
- Chomsky, N. 1992a. *Explaining Language Use*. Philosophical Topics 20, 205-231.
- Chomsky, N. 1995. *The Minimalists Program*. Cambridge MA: MIT Press.
- Chomsky, N. 2000a. *The Architecture of Language*. New Delhi: Oxford University Press.
- Chomsky, N. 2000b. *New Horizons in the Study of Language and Mind*. Cambridge: Cambridge University Press.
- Chomsky, N. 2002 *On Nature and Language*. Cambridge: Cambridge University Press.
- Chomsky, N. 2005a "Three factors in Language design". *Linguistic Inquiry*. 36/1, 1-22
- Chomsky, N. 2005b "Language and Brain". In A. J. Saleemi, O-S. Bohn and A. Gjedde eds *In Search of a Language for the Mind Brain*. Aarhus: Aarhus University Press, pg 143-170
- Chomsky, N, Pietroski, P, Yakama, B, Berwick, R. 2011 "Poverty of Stimulus Revisited". *Cognitive Science* p 1-36
- Chomsky, N, Interviews with McGilvray 2012 *The Science of Language*. Cambridge University Press.
- Chomsky, N and Halle, M. 1968 *The Sound Pattern of English*. New York: Harper and Row.
- Chomsky, N and Lasnik, H. 1977 "Filters and Control". *Linguistic Inquiry* 8 pg 425-504

Works by Quine

- Quine, W.V 1950 *Methods of Logic*. Cambridge MA: Harvard University Press.
- Quine, W.V 1953 *From a Logical Point of View*. Cambridge MA: Harvard University Press.
- Quine, W.V 1960 *Word and Object*. Cambridge MA: Harvard University Press.
- Quine, W.V 1962 *Mathematical Logic*. New York: Harper and Row.
- Quine, W.V 1963 *Set Theory and its Logic*. Cambridge MA: Harvard University Press.
- Quine, W.V 1966 *Selected Logic Papers*. New York: Random House.
- Quine, W.V 1966 *Ways of Paradox*. New York: Random House.
- Quine, W.V 1969 *Ontological Relativity and Other Essays*. New York: Columbia University Press.
- Quine, W.V 1970 *Philosophy of Logic*. Engelwood Cliffs NJ: Prentice-Hall
- Quine, W.V 1974 *Roots of Reference*. La Salle IL: Open Court
- Quine, W.V 1981 *Theories and Things*. Cambridge MA: Harvard University Press.
- Quine, W.V 1992 *Pursuit of Truth*. Cambridge MA: Harvard University Press.
- Quine, W.V 1995 *From Stimulus to Science*. Cambridge MA: Harvard University Press.
- Quine, W.V and J.S. Ullian 1970 *The Web of Belief*. New York: Random House.
- Quine, W.V 1937 "New Foundations for Mathematical Logic". *American Mathematical Monthly* 44. Pg 70-80

- Quine, W.V and Nelson Goodman 1947 "Steps towards a Constructive Nominalism". *Journal of Symbolic Logic* 12. Pg 97-122
- Quine, W.V 1951 "On Carnaps View of Ontology". *Philosophical Studies* 2 pg 203-211
- Quine, W.V 1952 "Two Dogmas of Empiricism". *Philosophical Review* 60 pg 20-43
- Quine, W.V 1980 "Identity, Ostension, and Hypostasis". In *From a Logical Point of View*. Cambridge MA: Harvard University Press. Pg 65-79
- Quine, W.V 1956 "Quantifiers and Propositional Attitudes". *Journal of Philosophy* 53. P185-196
- Quine, W.V 1960 "Variables Explained Away". *Proceedings of the American Philosophical Society* 104. 347-
- Quine, W.V 1960 "Posits and Reality" in S. Uyeda, ed, *Basis of the Contemporary Philosophy*. Tokyo: Waseda Harvard University Press. Pg 246-254
- Quine, W.V 1966 "Carnap and Logical Truth" in *Ways of Paradox*. New York: Random House. Pg 107-132
- Quine, W.V 1966 "Immanence and Validity" in *Selected Logical Papers*. New York: Random House. Pg 242-250
- Quine, W.V 1966 "Mr Strawson on Logical Truth" in *Ways of Paradox*. New York: Random House. Pg 137-157
- Quine, W.V 1969 "Epistemology Naturalised" in *Ontological Relativity and Other Essays*. New York: Columbia University Press. Pg 69-90
- Quine, W.V 1970 "On the Reasons for the Indeterminacy of Translation". *The Journal of Philosophy* 67. Pg178-183
- Quine, W.V "Philosophical Progress in Language Theory", in *Language, Belief and Metaphysics*, Vol one of *Contemporary Philosophic Talk*. Albany NY: State University of New York Press. Pg 2-19
- Quine, W.V 1975 "Empirically Equivalent Systems of the World". *Erkenntnis* 9 pg313-28
- Quine, W.V 1975 "Mind and Verbal Dispositions". In *Mind and Language*, ed. S. Guttenplan. Oxford: Oxford University Press. Pg 83-95
- Quine, W.V 1975 "The Nature of Natural Knowledge", in *Mind and Language*, ed. S. Guttenplan. Oxford: Oxford University Press. Pg 67-81
- Quine, W.V 1976 "Wither Physical Objects?" In *Essays in Memory of Imre Lakatos*, eds R.Cohen, P Feyerabend and M. Wartofsky. Dordrecht: Reidel. Pg497-504
- Quine, W.V 1979 "Facts of the Matter", in *Essays on the Philosophy of W. V. Quine*, eds R.W. Shahan and C. Swoyer. Normon OK: University of Oklahoma Press. Pg155-169
- Quine, W.V 1980 "The Variable and its Place in Reference". In *Philosophical Subjects: Essays on the Work of P.F. Strawson*, ed Z. Van Straaten. Oxford: Oxford University Press. Pg 164-173
- Quine, W.V 1981 "Empirical Content" in *Theories and Things*. Cambridge MA: Harvard University Press. Pg 24-30
- Quine, W.V 1981 "What Price Bivalence?" *Journal of Philosophy* 78, no 2 pg 90-95
- Quine, W. V 1984, "Relativism and Absolutism". *Monist* 67 [1984] pg 293-296
- Quine, W.V 1991, "Two Dogmas in Retrospect". *Canadian Journal of Philosophy* 21 pg 265-274
- Quine, W.V 1992, "Structure and Nature". *Journal of Philosophy* 89. Pg 6-9
- Quine, W.V 1993, "In Praise of Observation Sentences". *Journal of Philosophy* 90, no 3, pg107-116
- Quine, W.V 1994, "Promoting Extensionality". *Synthese* 98, no 1 pg 143-151

Quine, W.V 1994, "Response to Hookway". *Inquiry* 37 pg502-504

Quine, W.V 1995, "Naturalism or Living within Ones Means", *Dialectica* 49 pg 251-61

Quine, W.V 1996 "Progress on Two Fronts". *Journal of Philosophy* 93 159-163

Anthologies:

Roger Gibson ed 2004 *Quintessence: Basic Readings from the Philosophy of Quine*. Cambridge MA: The Belknap Press of Harvard University Press.

Dagfinn Føllesdal and Douglas B. Quine eds 2008 *W.V. Quine: Confessions of a Confirmed Extensionalist and Other Essays*. Cambridge MA: Harvard University Press

Dagfinn Føllesdal and Douglas B. Quine eds 2008 *W.V. Quine: Quine in Dialogue*. Cambridge MA: Harvard University Press.

Works by other Authors

Aitchinson, J. 1976 *The Articulate Mammal*. Oxon: Routledge

Antony, L. M., and Hornstein, N. Eds 2003 *Chomsky and His Critics*. Oxford: Blackwell Publishing.

Ayer, A. J. 1959 *Language Truth and Logic*. New York: Free Press

Baillargeon, R., Spelke, E and Wasserman, S. 1985 *Object Permanence in Five Month Old Infants*. *Cognition* 20 pg 191-208.

Barrett, R., and Gibson, R. 1990 *Perspectives on Quine*. Oxford: Blackwell

Baron Cohen, S., 1995 *Mindblindness: An Essay on Autism and The Theory of Mind*. Cambridge MA: MIT Press.

Bloom, P. 2000 *How Children Learn the Meaning of Words*. Cambridge MA: MIT Press

Bloom, P. 2004 *Descartes Baby*. London: William Heinemann.

Boeckx, C. 2006 *Linguistic Minimalism*. Oxford: Oxford University Press.

Boeckx, C. 2010 *Language in Cognition*. London: Wiley-Blackwell

Boghossian, P. 2008 *Content and Justification*. Oxford: Clarendon Press.

Bohannon, J., and Stanowicz, L., 1988 "The Issue of Negative Evidence: Adult Responses to Children's Language Errors". *Developmental Psychology* 24 [5] pg 684-689

Bohannon J, N., Snow, C., and Mac Whinney, B., 1990 "No Negative Evidence Revisited: Beyond Learnability or Who Has to Prove What to Whom". *Developmental Psychology* 26 [2] pg 221-226.

Bonjour, L. 1998 *In Defense of Pure Reason*. Cambridge: Cambridge University Press.

Brandom, R. 1994 *Making it Explicit*. Cambridge MA: Harvard University Press.

Brandom, R. 2008 *Between Saying and Doing*. Oxford: Oxford University Press.

Brown, R., and Hanlon, C, 1970 "Derivational Complexity and Order of Acquisition in Child Speech". In R, Hayes ed *Cognition and Development in Language*. New York: John Wiley pg 11-53.

Susan, C. 2010. *The Origin of Concepts*. Oxford: Oxford University Press.

Carnap, R. 1950 *Meaning and Necessity*. Chicago: Chicago University Press

Chouinard, M., and Clark, E, 2003 "Adult Reformulations of Child Errors as Negative Evidence" *Journal of Child Language* 30. Pg 637- 669

Christiansen, M, and Chater, N. 1999 "Connectionist Natural Language Processing: The State of the Art". *Cognitive Science* 23 [4] pg 417-437.

Christiansen, M, and Chater, N. 1999 "Towards a Connectionist Model of Recursion in Human Linguistic Performance." *Cognitive Science* 23 [2] pg 157-205

- Churchland, P. 1981 "Eliminative Materialism and The Propositional Attitudes". Vol LXXVIII no 2 pg 67-90
- Cook, V, J., and Newson, M. 2007 *Chomsky's Universal Grammar: An Introduction*. Oxford: Blackwell Publishing
- Collins, J 2002 "On the Very Idea of a Science Forming Faculty". *Dialectica* Vol 56 pg 125-151
- Collins, J, 2003 "Cowie on the Poverty of Stimulus" *Synthese* 136 [2] pg 159-190
- Collins, J, 2007 "Meta-scientific Eliminativism: A Reconsideration of Chomsky's Review of Skinner's Verbal Behaviour". *British Journal of Philosophy of Science* 58. Pg 625-658
- Collins, J, 2008 *Chomsky: A Guide for the Perplexed*. London: Continuum.
- Cowie, F. 1999 *Whats Within: Nativism Reconsidered*. Oxford: Oxford University Press.
- Crain, S and Nakayama 1987 "Structure Dependence in Grammar Formation". *Language* 63 pg 522-543
- Crain, S, and Pietroski P 2002 "Why language acquisition is a snap". *Linguistic Review* 19 pg 163-183
- Davidson, D., and Hintikka, Eds, *Word and Objections*. Dordrecht: Reidel.
- Davidson, D. 1984, *Inquiries into Truth and Interpretation*. Oxford: Clarendon Press.
- Davidson, D. 2001, *Subjective, Intersubjective, Objective*. Oxford: Oxford University Press.
- Davidson, D. 2004, *Problems of Rationality*. Oxford: Clarendon Press.
- Davidson, D. 2004, *Truth, Language, and History*. Oxford: Clarendon Press.
- Davidson, D. 2005, *Truth, and Predication*. Cambridge MA: Harvard University Press.
- Devitt, M. 2006 *Ignorance of Language*. Oxford: University Press.
- Evans, G. 1975 "Identity and Predication". *The Journal of Philosophy* Vol lxxii, no 13 pg343-363
- Everett, D, 2012 *Language: The Cultural Tool*. London: Profile Books
- Evine, S. 1991 *Donald Davidson*. Oxford: Basil Blackwell.
- Fantcher, R. 1996 *Pioneers of Psychology*. New York: W.W. Norton and Company INC.
- Fitch, W.T., Hauser, M. D., and Chomsky N, 2005 "The evolution of the language faculty: Clarifications and implications". *Cognition* 97 pg 179-210
- Fodor, J. 1998 *Concepts: Where Cognitive Science Went Wrong*. Oxford: Oxford University Press.
- Fodor, J 2008 *LOT 2: The Language of Thought Revisited*. Oxford: Oxford University Press
- Follesdal, D. 1986 "Essentialism and Reference", In Louis Hahn ed *The Philosophy of W. V Quine*. LA Salle: Open Court. Pg 97-115
- Gernsbacher, M, A., Sauer, E., Geye, H., Schweigert, E., and Goldsmith, H. 2008 "Infant and toddler oral- and manual- motor skills predict later speech fluency in autism". *Journal of Child Psychology and Psychiatry* 49 [1] pg 43-50
- Gibson, R. 1982 *The Philosophy of Quine: An Expository Essay*. Florida: University Presses of Florida.
- Gibson, R. 2004 *The Cambridge Companion to Quine*. Cambridge: Cambridge University Press.
- Goodman, N. 1979 *Fact, Fiction and Forecast*. Cambridge MA: Harvard University Press

- Grice, H. P., and Strawson, P. F., 1956 'In defence of a Dogma' *The Philosophical Review* Vol lxv
- Harley, T. 2001 *The Psychology of Language*. Sussex: Psychology Press Ltd.
- Hauser, M. 2006 *Moral Minds*. USA: Harper Collins Publishers.
- Hauser, M., Fitch, W., and Chomsky N. "What is it, Who has it, and How Did it Evolve?" *Science* 298 pg 1569- 1579
- Hempel, C. 1945 "Studies in the Logic of Confirmation". *Mind* 213 pg 1-26
- Hylton, P. 2007 *Quine*. New York: Routledge
- Hinzen, W. 2007 *An Essay on Names and Truth*. Oxford: Oxford University Press
- Hornstein, N. 2009 *A Theory of Syntax*. New York. Cambridge University Press.
- Horwich, P. 1991 "Chomsky Versus Quine on the Analytic-Sythetic Distinction". *Proceedings of the Aristotelian Society* 92 pg 95-108
- Hume, D. 1988 *An Inquiry Concerning Human Understanding and Concerning Principles of Morals*. Oxford: Clarendon Press.
- Hurford, J., Heasley, B., and Smith, M. 1983 *Semantics: A coursebook*. Cambridge: Cambridge University Press.
- Jackendoff, R. 2002 *Foundations of Language*. Oxford: Oxford University Press.
- Jackendoff, R., and Pinker, S. 2005 "The nature of the language faculty and its implications for evolution of language. (Reply to Fitch, Hauser, and Chomsky)" *Cognition* 79 pg211-225
- Johnson, S., and Carey, S. 1998. "Knowledge Enrichment and Conceptual Change in Folkbiology: Evidence from Williams Syndrome". *Cognitive Psychology* 37. Pg 156-200.
- Juhl, C., and Loomis 2010 *Analyticity*. Oxon: Routledge
- Kasher, A. 1991 *The Chomskian Turn*. Oxford: Blackwell
- Katz, J and Postal 1991 "Realism vs Conceptualism in Linguistics". *Linguistics and Philosophy* 14 pg 515-554
- Katz, J., 1967 "Some Remarks on Quine on Analyticity". 64 pg 36-52
- Katz, J 1974 "Where Things Stand With The Analytic-Synthetic Distinction". *Synthese* 28 pg 283-319
- Kemp, G 2006 *Quine: A Guide for the Perplexed*. New York: Continuum
- Kripke, S. 1972 *Naming and Necessity*. Oxford: Blackwell Publishing
- Kuhl, P. 2007 "Is Speech learning 'gated' by the social brain?" *Developmental Science* 10 [1] pg 110-120.
- Lappin, S., and Clark, A. 2011 *Linguistic Nativism and the Poverty of the Stimulus*. London: Wiley-Blackwell
- Lasnik, H. 2000 *Syntactic Structures: Revisited*. Cambridge MA: MIT Press.
- Legete, J, and Yang, C, 2002 "Empirical re-assessment of stimulus poverty arguments" *Linguistic Review* 19 pg 151-162
- Lepore, E, and Ludwig, K. 2007 *Truth-Theoretic Semantics*. Oxford: Oxford University Press.
- Lyons, J. 1970 *Chomsky*. Suffolk; Fontana
- Marr, D. 1982 *Vision*. Cambridge MA: MIT Press.
- Martinich, A.P. 2001 *The Philosophy of Language*. Oxford: Oxford University Press.
- Marcus, G. 1993 "Negative Evidence in Language Acquisition" *Cognition* 46 pg 53-85
- Markman, E. 1989 *Categorization and Naming in Children*. Cambridge MA: MIT Press
- Morley-Bunker, N, "People's Intuitions of Analyticity" (unpublished undergraduate psychology thesis)

- O Grady, P. 1999 "Carnap and Two Dogmas of Empiricism." *Philosophy and Phenomenological Research* 4 pg 1015-1027
- Pagin, P. 2005 "Indeterminacy and the analytic/synthetic distinctions: a survey". *Synthese* 10
- Paul, R., Chawarska, K., Fowler, C., Cicchetti, D., and Volkmar, F., 2007 "Listen My Children and You Shall Hear": Auditory Preferences in Toddlers With Autism Spectrum Disorders." *Journal of Speech, Language, and Hearing Research* 50 pg 1350-1364
- Peregrin J. 2005 "Brandom and Davidson: What do we Need to Account for Thinking and Agency". *Philosophica* 75 pg43-59
- Piaget, J. 1926 *The Language and Thought of the Child*. London: Routledge
- Piattelli-Palmarini, M, 1989. "Evolution, selection, and cognition". *Cognition* 31 pg 1-44.
- Piattelli-Palmarini, M., Uriagereka, J., and Salaburu, P. Eds *Of Minds and Language: A dialogue with Noam Chomsky in the Basque Country*. Oxford: Oxford University Press.
- Pinker, S., and Bloom, P. 1990 "Natural Language and Natural Selection". *Behavioural and Brain Sciences* 13 pg 707-784.
- Pinker, S. 1994 *The Language Instinct*. London: Penguin Books.
- Pinker, S., and Jackendoff, R 2004 "The Faculty of Language: what's special about it?" *Cognition* 95 pg 201-236.
- Popper, K. 1992 *The Logic of Scientific Discovery*. New York: Routledge
- Pullum, G. 1996 "Learnability, hyperlearning, and the poverty of stimulus." In *Proceedings of the 22th annual meeting: General Session and Parasession on the role of learnability in Grammatical Theory*. Johnson, J., Juge, J., and Moxley, J., eds pg 498-513
- Pullum, G., and Scholz, B., 2002 "Empirical assessment of stimulus poverty arguments." *Linguistic Review* 19 pg 9-50
- Pullum, G., and Scholz, B., 2002 "Searching for arguments to support linguistic nativism" *Linguistic Review* 19 pg 185-223
- Pullum, G., and Scholz, B., 2006 "Irrational Nativist Exuberance". In Stainton, R ed *Contemporary Debates in Cognitive Science* Basil Blackwell. Pg 59-80.
- Pullum, G., and Scholz, B., 2010 "Recursion and the Infinitude claim" In Harry Van Der Hulst ed *Recursion in Human Language*. Berlin: Mouton de Gruyter pg 113-138
- Putnam, H. 1975 *Mind, Language and Reality*. Cambridge: Cambridge University Press.
- Ramberg, B. 1986 *Davidson's Philosophy of Language*. Oxford: Basil Blackwell.
- Rorty, R, 1972 "Indeterminacy of Translation and of Truth". *Synthese* 23 pg 443-462
- Rorty, R, 1979 *Philosophy and the Mirror of Nature*. Princeton: Princeton University Press
- Rorty, R, 2003 "The Brain as Hardware, Culture as Software". *Inquiry* 47 pg 219-235.
- Sampson, G, 1980 "Popperian Language-Acquisition Undefeated". *British Journal of Philosophy of Science* 31 pg 63-67.
- Sampson, G, 2002 "Exploring the richness of the stimulus". *Linguistic Review* 19 pg 73-104
- Sampson, G, 2005 *The Language Instinct Debate*. London: Continuum.
- Sampson, G, 2007 "There is No Language Instinct." *Ilha do Desterro* 52 pg 35-63.
- Searle, J 1984 "Indeterminacy, Empiricism and the First Person." *Journal of Philosophy* 81. Pg123-146

- Skinner, B, F. 1957 *Verbal Behaviour*. New York: Appleton-Century-Crofts.
- Soja, N., Carey, S., and Spelke, E., 1991 "Ontological categories guide young children's inductions of word meaning: Object terms and substance terms." *Cognition*, 38, pg179-211
- Stich, S. 1979 "Between Chomskian Rationalism and Popperian Empiricism". *British Journal of Philosophy of Science* 30 pg 329-347
- Tomasello, M. 1995 "Language is Not an Instinct". *Cognitive Development* 10. 1g 131-156
- Tomalin, M. 2003 "Goodman, Quine, and Chomsky: from a grammatical point of view." *Lingua* 113 pg 1233-1253
- Vygotsky, L. 1986 *Thought and Language*. Cambridge MA: MIT Press.
- Weir, A 2005 "Indeterminacy of Translation" In Leopore, E, and Smith, B eds *The Oxford Handbook of Philosophy of Language*. Oxford: Oxford University Press pg 233-249
- Williamson, T. 2007 *The Philosophy of Philosophy*. Oxford: Blackwell Publishing.
- Wittgenstein, L. 1922 *Tractatus Logico-Philosophicus*. London: Routledge and Kegan Paul.
- Wittgenstein, L. 1953 *Philosophical Investigations*. Oxford: Blackwell.