

Is Linguistics a Part of Psychology?

Gareth Fitzgerald
University College London

Submitted in partial fulfilment of the requirements for the degree of Doctor of
Philosophy.

I, Gareth Fitzgerald, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Abstract

Noam Chomsky, the founding father of generative grammar and the instigator of some of its core research programs, claims that linguistics is a part of psychology, concerned with a class of cognitive structures employed in speaking and understanding. In a recent book, *Ignorance of Language*, Michael Devitt has challenged certain core aspects of linguistics, as prominent practitioners of the science conceive of it. Among Devitt's major conclusions is that linguistics is *not* a part of psychology. In this thesis I defend Chomsky's psychological conception of grammatical theory. My case for the psychological conception involves defending a set of psychological goals for generative grammars, centring on conditions of descriptive and explanatory adequacy. I argue that generative grammar makes an explanatory commitment to a distinction between a psychological system of grammatical competence and the performance systems engaged in putting that competence to use. I then defend the view that this distinction can be investigated by probing speakers' linguistic intuitions. Building on the psychological goals of generative grammar and its explanatory commitment to a psychological theory of grammatical competence, I argue that generative grammar neither targets nor presupposes non-psychological grammatical properties. The latter non-psychological properties are dispensable to grammarians' explanations because their explanatory goals can be met by the theory of grammatical competence to which they are committed. So generative grammars have psychological properties as their subject matter and linguistics is a part of psychology.

Table of Contents

<i>Abstract</i>	3
<i>Acknowledgements</i>	7
1. Introduction	8
1.1 What is the subject matter of Linguistics?	8
1.2 The Psychological Conception of Linguistics	20
1.3 Non-psychological Conceptions of Linguistics	41
1.3.1 Social Conceptions	55
1.3.2 Physicalist Conceptions	76
1.3.3 Platonist Conceptions	90
1.4 Argument Summary	98
2. What does Generative Grammar Explain?	101
2.1 Goals and Interests	101
2.2 Descriptive Adequacy	108
2.2.1 Prescriptive Grammars	108
2.2.2 Observational Adequacy	110
2.2.3 Descriptive Adequacy and Productivity	113
2.3 Explanatory Adequacy	117
2.3.1 Further Conditions of Adequacy?	117
2.3.2 Grammar Acquisition and Universal Grammar	118
2.3.3 The Role of Explanatory Adequacy	120
2.4 Arguments against Explanatory Adequacy	126
3. The Competence-Performance Distinction	138
3.1 Competence and Use	141
3.2 Why Generative Grammar Needs a Competence-Performance Distinction	144
3.3 Grammatical Competence and Linguistic Processing	152
3.5 Katz's Case against "Competencism"	163
4. Linguistic Intuitions as Evidence for Generative Grammars	174
4.1 Intuitions as Evidence	174

4.2 The Orthodox Model: Linguistic Intuitions as Data for Psychological Theories	187
4.2.1 How do Intuitions bear on Competence Theories?	187
4.2.2 Linguistic Intuitions and Visual Impressions	190
4.2.3 Are Linguistic Intuitions the “Voice of Competence”?	197
4.2.4 Are Linguistic Judgements and Visual Reports Disanalogous?	200
4.2.5 The Orthodox Model and Introspection	201
4.3 Linguistic Intuitions as Theoretical Judgements	203
4.3.1 Devitt’s Model and Belief-Independence	203
4.3.2 Devitt’s Model and Folk theory	204
4.3.3 A modification to Devitt’s Model	207
4.3.4 Devitt’s alternative view of the Evidence	209
4.4 Linguistic Intuitions as Observational Judgements	215
4.4.1 The Observational Model	215
4.4.2 Error, Belief-Independence and Types	217
4.4.3 Grammatical Variation	219
5. <i>Explanation in Generative Grammar</i>	224
5.1 What sort of properties do Generative Grammars appeal to?	224
5.1.1 Explanation and Supervenience	225
5.1.2 Psychological Reality, Linguistic Reality and Grammatical Explanation	229
5.2 The No Violence Principle	234
5.3 Meeting Descriptive Adequacy	245
5.4 Meeting Explanatory Adequacy: Nativism	250
5.4.1 Theories of Grammar Acquisition: Nativist Theories and Learning Theories	251
5.4.2 Poverty of Stimulus Arguments for Grammatical Nativism	263
5.4.3 Conventions	287
5.5 Communication	292
6. <i>Conclusion</i>	296
<i>Appendix. Beyond Explanatory Adequacy</i>	298
<i>Bibliography</i>	317

Acknowledgements

This thesis was supervised by Dr. Mark Kalderon, Dr. Guy Longworth and Dr. Barry Smith. I owe them many thanks for their encouragement, ideas and generosity with their time. I thank my examiners Richard Breheny and Gabriel Segal for helpful suggestions and corrections. I would also like to thank the following people for reading and discussing parts of the thesis: Dr. John Collins, Craig French, Dr. Colin Johnston, Jonny McIntosh, Will McNeill, Simon Riches, and Lee Walters.

Parts of Chapter One and Five have been presented and then submitted as a paper “Is Linguistics a Part of Psychology?”. They have greatly benefitted from the questions of audiences in Warwick and London, and an anonymous reviewer for *Philosophical Psychology* for whom the paper is being revised and resubmitted. Part of Chapter One is also forthcoming in *Minds and Machines* as a review of Devitt’s *Ignorance of Language*. Chapter Two was presented at the Warwick Philosophy of Linguistics seminar where it profited from comments by Matt Soteriou. Chapter Four has been presented to the London Graduate Conference and London Philosophy of Linguistics seminar series. It has also benefitted from comments by two anonymous referees for the *British Journal for the Philosophy of Science*, for whom the paper version is being revised and resubmitted.

1. Introduction

1.1 What is the subject matter of Linguistics?

This is a thesis in the philosophy of linguistics. The thesis is about the subject matter of linguistic theories. In the broadest terms, linguistics is the scientific study of language. However, the scientific study of language is not an attempt to explain all linguistic phenomena, or everything that we pre-theoretically think of as a part of our language.

For example, the science of language does not directly address the kinds of questions literature students ask about Keats' language and what Keats meant by particular choices of words. They might consider the question of what Keats meant when, speaking of the nightingale, he said that he, Keats, was "too happy in thine happiness".¹ There are special subtleties to a poet's language and meanings, and discerning these subtleties is an art in itself. It may involve understanding the poet's particular style, historical circumstances, state of mind and literary intentions. The science of language, as currently practiced, does not afford us much insight into these questions. This is perhaps to be expected because such phenomena look to involve a wide range of potentially intractable factors; factors that may not be easy to make precise and objective. The scientific investigation of language looks to strip away our peculiar perspectives and interests in order to get some precise and objective explanation of central linguistic facts.

This is not to detract from the importance of other perspectives on language. It is a theme of Noam Chomsky's writings on language and mind that someone committed to scientific inquiry can consistently believe that there is much of human interest concerning what people say and do that falls well beyond the scope of such inquiry.² We might, for example, think that there is more to be learnt about the

¹ Keats' *Ode to a Nightingale* is available from the Every Poet archive at: http://www.everypoet.com/Archive/Poetry/John_Keats/keats_ode_to_a_nightingale.htm

² See Chomsky (1995a).

subtleties of Keats' meaning by reading his letters, studying Spenser, or consulting history books than from the whole of the scientific inquiry into language.

Linguists want to explain certain central facts about all human languages. To that extent, linguistics may reveal something about the fundamentals of even a poet's language; facts about linguistic form and meaning that do appear to be amenable to scientific inquiry. This thesis focuses on one part of the science of language. It focuses on that part of linguistics that is concerned with the *structure* of the sentences that we utter and hear. This part of linguistics is called *grammar*, or *syntax*. Grammarians are interested in languages as combinatorial systems that exhibit special, complex structures: the sentences that we utter and comprehend, read and write. Linguistic structures are built by combining *lexical items*. These lexical items are selected from a *lexicon*: a stock of the words and other items that populate the sentences of a language. Over the last fifty years there has been an intensive and extremely fruitful investigation into the structures of human languages, initiated by Chomsky, which has come to be known as the *Generative Enterprise*. It has this name because the central concern of linguists within this enterprise is to construct *generative grammars* for human languages.³ A generative grammar is a theory of a language that yields an *explicit*, structural description of all and only the sentences of that language.

Here is a simplified piece of theorising of the sort that generative grammarians are involved in. It concerns the analysis of sentences containing reflexives like *myself*, *yourself*, *herself*, *himself*, *itself*, *ourselves*, *yourselves*, and *themselves*. Though it is simplified, it is illustrative of the structural complexity which generative grammar has brought to light.⁴ Consider the following examples.

(1) I shaved myself

(2) *Myself shaved me.

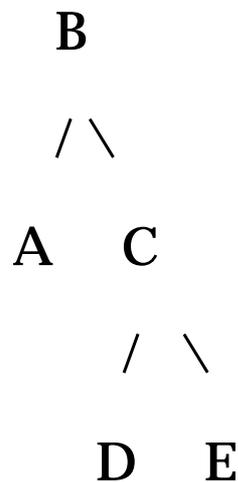
English-speakers recognise (1) as a perfectly good sentence. But those same speakers immediately recognise that there is something amiss with (2). As English-

³ Chomsky (1955/75, 1957, 1965) are founding texts of generative grammar.

⁴ This presentation is adapted from Adger (2003) pp.116-120.

speakers, we would replace (2) with (1). Generative grammarians want to explain the recognised difference between (1) and (2). These are constructions containing reflexives. So the grammarian tries to give a precise characterisation of the structures containing reflexives that are sentences of human languages, such as English, and those that are not.

Grammarians can explain a lot about the behaviour of reflexives, including the permissible structures in which they occur, in terms of a special sort of structural dependency between the items, or *constituents*, that make up sentences. This special structural dependency is called *c-command*, short for constituent-command, and is an extremely useful tool in theorising about linguistic structure.⁵ To see how c-command works, look at the tree diagram below.



Sitting at the top of the tree diagram is *node* B, which labels the whole structure. B *contains* the constituents at nodes A, C, D and E. Of those constituents, B *immediately contains* A and C which are directly below B in the tree. Similarly, C immediately contains D and E which are directly below C in the tree. A very intuitive way of thinking about these relations is in familial terms. B is the mother of A and C but the grandmother of D and E. C is the mother of D and E. A and C

⁵ C-command relations are part of the explanation of the behaviour of reflexives and pronouns (see Adger 2003 pp.118-20, 149-152) as well as negative polarity items such as *any* (pp.120-2), and relations of Agreement (Adger 2003 pp.167-9, Chomsky 2001).

are sisters, as are D and E. In addition to relations like motherhood and sisterhood, there are further structural relations that hold between the nodes in the tree. Some of these relations are important to linguistic structure, and c-command is one of these.

A node X c-commands a node Y if, and only if, X's sister either: (i) is Y, or (ii) contains Y. In our tree, A and C c-command one another as they are sisters. But A also c-commands D and E as its sister C contains them. B doesn't c-command anything as it stands in no relations of sisterhood. Neither is B c-commanded by anything since it contains all the other nodes. Nodes can c-command preceding as well as following nodes, running from left-to-right.

Reflexive constructions, like (1) and (2), are an example of c-command in action. In order for a reflexive to be part of an acceptable sentence like (1), it has to enter into special relationships with other constituents in the structure in which it occurs. For a reflexive, like *myself* in (1), to be a part of a good sentence it must be *bound*: it must be referentially dependent on, i.e. have its interpretation fixed by, an antecedent in that sentence. *Myself* in (1) is bound by *I*. We need to know how it is determined when a potential antecedent can bind a reflexive; for all seems to go well with (1) but not with (2). A popular hypothesis in grammatical theory is that a reflexive must be bound by an antecedent that c-commands it.⁶ In what follows I'll offer some linguistic argument for that hypothesis.

What could explain the contrasting judgements English-speakers make about (1) and (2)? It seems that in (1) the reflexive is co-referential (shares a referent) with another expression in the sentence, in this case *I*. So we might form the following generalisation:

The Co-reference Generalisation: A reflexive must be co-referential with another expression in the sentence.

However, the Co-reference Generalisation would not explain the difference between (1) and (2) because we can easily imagine circumstances in which the *myself* and *me* in (2) co-refer, and yet (2) would still be a recognisably poor sentence though (1) isn't.

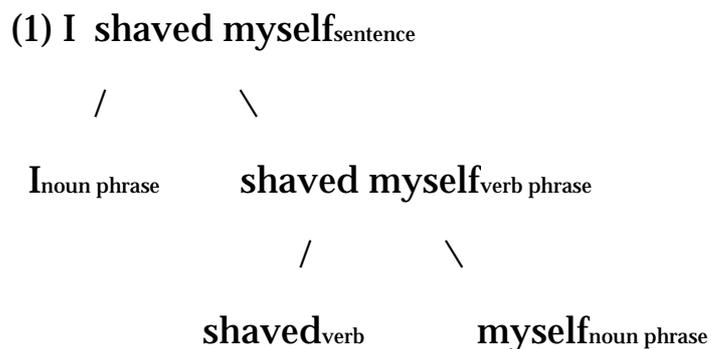
⁶ For a famous investigation of these issues see Chomsky (1981).

We might look at some of the properties the different lexical items in (1) and (2). Lexical features pertaining to number, person and gender are called ψ -features. We might hypothesise that a reflexive and its antecedent must bear the same ψ -features. We could then frame a new hypothesis about reflexives: that they must be co-referential with another expression in the sentence that shares the same ψ -feature specification. But again this wouldn't explain the difference between (1) and (2). In both (1) and (2) we have two expressions which share singular, first-person ψ -features, one of which is a reflexive. So this hypothesis falsely predicts that (1) and (2) should be equally good.

An aspect of (1) that we need to capture is that *myself* doesn't merely co-refer with *I* but is actually bound by *I*. It must be interpreted to refer to the same individual that *I* does. If, in addition to this dependency, we add the notion of c-command, we can drastically improve upon our previous generalisations:

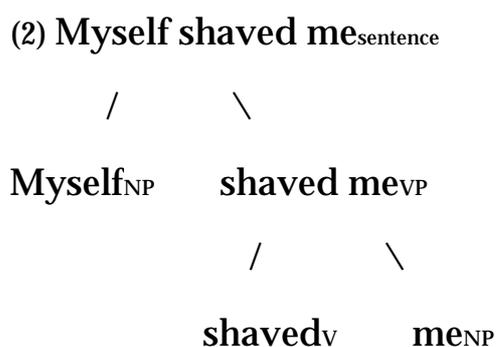
The Reflexive Generalisation: A reflexive must be bound by an antecedent that c-commands it.

As roughly depicted in the tree diagram below, the noun phrase (NP) *I*, c-commands the NP *myself* in (1). Generally, in simple sentences like (1), the object of a sentence will be c-commanded by the subject of a sentence, since the object is contained in the subject's sister and not vice-versa.⁷



⁷ I assume here that sentences are composed NP-VP, see Chomsky (1986) p.60 for discussion and Adger (2003) p.90-6 for relevant constituency tests.

In (2) the reflexive is not bound by a c-commanding antecedent because even if *myself* is interpreted as co-referential with *me* it is not c-commanded by *me* as we can see.



This Reflexive Generalisation making use of c-command relations thus explains the contrast between (1) and (2).⁸

We can test the Reflexive Generalisation that appeals to binding and c-command against competitors such as the principle that a reflexive must be bound by a *preceding* expression. This competing explanation does not extend to further data such as (3).

(3) *[The man I saw] shaved myself.

⁸ Our reflexive generalisation is roughly equivalent to Principle A of Binding Theory (see Chomsky 1981 and 1995) which states that a reflexive must be bound by a *local* antecedent. Reflexives cannot be bound by elements that are too “far away” structurally, whilst pronouns cannot be bound by elements that are too “close by”. The notion of locality is a theoretical one: there is a whole branch of grammatical theory that focuses on the lower and upper bounds on the distance that may separate two elements that enter into a grammatical relation, and on the movement of constituents. The principle can also be stated in terms of the notions of *governing category* or *domain* (see Chomsky 1987 p.185), where the domain of a constituent X is the phrase which immediately contains X, roughly “the smallest phrase containing it”. So X will c-command constituents in its domain. Principle A of classical Government and Binding Theory states that a reflexive must be bound by an antecedent in its domain. Principle B states that a pronoun may not be bound by an antecedent in its domain. Principle C states that a referring expression must be free (unbound) everywhere.

In (3) the pronoun *I*, which is a potential antecedent for *myself*, precedes the reflexive but (3) is not a good sentence. As indicated by the square brackets, *The man I saw* is a constituent of the sentence (3), according to standard tests for constituency⁹, and contains *I*.

Looking back to the schematic tree diagram, *The man I saw* occurs at node A. But there is further structure to *The man I saw*, so *I* is at a node below node A in the tree. We can see that *I* doesn't c-command *myself* as there is a node above *I* which does not contain the reflexive. The node which contains *myself* is C which is not a sister of *I*. Hence, *I* does not c-command *myself* though it does precede it. The notion we need is not precedence but c-command.

Keeping our simplified piece of grammatical theory in mind; how could it be an open question what linguistics is about? The question seems to have an easy answer: linguistics is about whatever linguists are studying. At least as far as grammar goes, linguistics is about the structural complexity of human languages. The subject matter of generative grammar seems reasonably clear: languages and the linguistic forms that they exhibit, such as the reflexive structures we just examined. At one level, the question of what linguistics is about seems to have an easy answer.

However, there is a deeper question about the subject matter of linguistics. Chomsky makes the following claim about the nature of linguistics:

Linguistics is simply that part of psychology that is concerned with one specific class of steady states, the cognitive structures employed in speaking and understanding.¹⁰

Chomsky thinks that linguistics is among the psychological sciences, a *part of* psychology, because its subject matter is the psychological structure of speaker-hearers.¹¹ Chomsky claims that the generative grammars investigated by linguists, including the lexicon and the principles that govern linguistic structures, describe

⁹ It has a coherent, independent semantic interpretation, can be replaced by constituents, can occur in object position and can be clefted as in "It was the man I saw that you wanted to meet".

¹⁰ Chomsky (1975) p.160

¹¹ I'll often use just "speaker" instead of "speaker-hearer".

the structure of the minds of those who speak and understand language. The Reflexive Generalisation, if correct, would therefore characterise psychological structure engaged in recognising the difference between (1) and (2). I'll call this conception of the subject matter of linguistics, the *psychological conception*.¹²

The psychological conception is widely adopted by linguists themselves. The following excerpt from Adger's introduction to syntax is representative:

[T]he idea of a sentence is more abstract than the idea of an utterance (which is something you can hear, record, feed into a computer as sound waves etc.). A sentence itself is something which can't be recorded, heard, or electronically manipulated, only **uses** of sentences can... The sentences themselves are defined by the knowledge of English that I put to use in writing them. Although it sounds counter-intuitive, what you see on this page are technically utterances, in that they have external, physical manifestation. *Sentences, on the other hand, are internal, mental entities, which have an abstract form.*¹³

As Adger conceives it, the grammatical structures that are the subject matter of generative grammar are "internal, mental entities". We make utterances, and these utterances are shaped by our underlying psychological structure. According to the psychological conception, what linguists study are these underlying psychological structures, as distinguished from the use of these structures in utterances.

Despite widespread endorsement amongst linguists, this psychological conception of generative grammar has been heavily criticised by philosophers.¹⁴ Michael Devitt's recent book, *Ignorance of Language*, constitutes a new attack on the psychological conception.¹⁵ Among Devitt's major conclusions is that linguistics is *not* a part of psychology. He believes that the subject matter of linguistics is not the psychological states of speaker-hearers. Devitt sets the issue up in stark terms that I'll devote a lot of attention to in this thesis. Ultimately, I'll argue

¹² This is Devitt's (2006) terminology. The psychological conception has also been labelled "the cognitive conception" (Smith 2006), "conceptualism" (Katz 1981), "mentalism" (Chomsky 1965, Smith 2007), and "The Right View" (Fodor 1985).

¹³ Adger (2003) p.2

¹⁴ See, for example, Devitt and Sterelny (1989, 1999), Dummett (1989), Katz (1981, 1985), Quine (1972), Searle (1974, 1992), Soames (1984, 1985)

¹⁵ Devitt (2006)

that he betrays fundamental confusions about the psychological conception of generative grammar and the reasoning that has been adduced in its favour.

The question at issue is not what it is for a true theory to be *linguistic*, which Fodor describes as a “boring” question.¹⁶ Chomsky and opponents to the psychological conception are not arguing about how the word “linguistic” should be used and which sorts of research are worthy of that label. The question at issue is rather the one that Fodor deems the “interesting” question, concerning what it is for theories of linguistic phenomena to be true. The interesting question is about the truth-makers of linguistic explanations: are linguistic explanations made true by psychological facts? If so, then linguistic theory is a psychological theory. If not, then linguistic theory is about something other than the psychological structures that are engaged in speaking and understanding language.

So there has been, and still is, controversy about the subject matter of linguistics. If Devitt is right, and linguistics is not a part of psychology, then it makes for an interesting situation. The orthodox psychological conception, of which Devitt takes the above statement from Chomsky to be representative, should be given up. The linguists in question are mistaken about the nature of their own enterprise in a fundamental respect. The theories they offer are not about the psychological states of those who speak and understand languages.

In place of the psychological conception, Devitt offers a *non-psychological conception* of generative grammar. According to Devitt, the subject matter of generative grammars is language and linguistic structures, but not psychological structures. Devitt calls the particular brand of non-psychological conception which he endorses, the *linguistic conception*. The linguistic conception has two components.

- L1. Linguistics is claimed to be about a non-psychological domain of linguistic facts. This component is common to all non-psychological conceptions.
- L2. Grammatical properties are properties of *physical tokens* such as acoustical signals and written marks. They are *relational* properties

¹⁶ Fodor (1985) p.146

of these physical occurrences: they are determined by the relations that bits of the sound stream and ink marks stand in to one another, and to social and psychological facts. They are also *high level* properties: realisable by heterogeneous physical tokens. This component is a particular claim about the nature of the non-psychological domain of grammatical properties not common to all non-psychological conceptions.

I'm going to reserve the label *linguistic conception* for Devitt's view that has these two components. And I'll call the collection of all the views that commit to L1 but not necessarily to L2, *non-psychological conceptions*.¹⁷

In §1.3, I'll consider Devitt's argument against the psychological conception. For now, I want to suppose that the subject matter of linguistics is an open question. What might help us to determine an answer to the question? We can look to three broad features of linguistic inquiry. The first is the linguist's *explananda*. We can examine the phenomena that the generative grammarian seeks to explain, and the goals of the theories. The second is the linguist's *evidence*. We can examine the sorts of evidence brought to bear in generative grammar, and what the evidence is employed as evidence for: i.e. what sorts of hypotheses the evidence supports. The

¹⁷ It is also possible to adopt L2, or some other non-psychological conception of grammatical properties, without adopting L1. For it is conceivable that someone could hold both that grammatical properties are non-psychological properties and that the scientific study of grammar is an investigation of the psychological states of speaker-hearers engaged in the cognition of grammar. Such *compatibilist* views have not received much attention. The idea that the science of grammar is directed at psychological properties of speaker-hearers is usually taken to foreclose on the possibility that there are non-psychological grammatical properties instantiated in our physical environment. Though he would put the view in terms of grammatical types rather than Devitt's tokens, Longworth has expressed some sympathy for such a view (Longworth 2007, forthcoming and unpublished ms.). Longworth thinks that the *scientific image* of grammatical properties according to which grammatical properties are properties of a psychological system and what he calls the *manifest image*, according to which grammatical properties are there in the environment for us to happen upon, are reconcilable. The concern of my thesis is to establish what the science of grammar is about; this is a prior question to whether any such compatibilist strategy is defensible.

third is the linguist's *explanans*: the sorts of hypothesised properties that are required to explain the explananda. With an account of generative grammar's explanatory goals in hand, and an account of the evidence grammarians draw upon, we can examine the nature of the posits of grammatical theory. We can ask what sort of theoretical properties are postulated by the grammarian to achieve his explanatory goals.

One point of clarification is required. Chomsky has, on occasion, argued that generative grammar is a psychological science as a matter of theoretical choice. His thought is that the subject matter of generative grammar is determined by which phenomena generative grammarians seek to explain. This, he claims, is settled by the *selection* of a domain of inquiry. Hence, he says:

The issue of mentalism versus antimentalism in linguistics apparently has to do only with goals and interests, and not with questions of truth or falsity, sense or nonsense.¹⁸

One might think that the question of whether generative grammar is a psychological science is thereby deflated. Generative grammar is a theoretical investigation, successful or unsuccessful, of a selected range of phenomena and attempts to answer questions thereof. Chomsky's thought is that generative grammarians have *chosen* a range of phenomena pertaining to our cognition of language as their target of explanation and so generative grammar is a psychological science. As Chomsky sees it, one can choose to be interested in the investigation of our cognition of language, as Chomsky claims generative grammarians are, or one can choose not to be. But there is no antecedent question of what linguistics is about as such.

However, there are substantive issues waiting in the wings. Opponents of the psychological question, such as Devitt, see themselves as asking for some justification for thinking that, in asking questions about central linguistic facts, we are thereby proposing a set of questions about our cognition of language. It may be that psychological phenomena are the stated targets of explanation. But there are important questions about whether this theoretical focus is well-motivated, whether linguistic inquiry has been in keeping with these explanatory goals, whether

¹⁸ Chomsky (1965) fn.1

linguists bring their evidence to bear on psychological hypotheses and whether the posits of their theories are best interpreted as psychological structures. And these are questions that Chomsky also addresses. So there is plenty of scope for defending the psychological conception beyond the appeal to an initial choice of explanatory goals.

I think Chomsky would agree that linguists are not infallible in their self-conception. It is, then, at least conceivable that opponents of the psychological conception could convince us that linguists are mistaken in taking themselves to be answering questions about our cognition of language, *if* the theories could be better interpreted as pertaining to some non-psychological domain.

Devitt is not the only philosopher who thinks that linguists should be convinced to give up the psychological conception of generative grammar. Katz is committed to L1 but not L2 because on Katz's view, linguistics is not a part of psychology but neither is it a part of any other empirical science. Katz thinks that the fundamental questions for grammarians concern a range of structural facts about languages that should be specified independently of hypotheses about the psychological states of speakers. He has sought to defend the view that grammatical theory is a part of mathematics rather than psychology.¹⁹ Katz thinks that mathematical propositions are about Platonic objects: atemporal, aspatial and acausal objects that are not located in the natural world, or indeed anywhere else. Hence, Katz thinks that mathematics, and therefore linguistics, are non-empirical investigations.

Whatever plausibility these challenges to the psychological conception have, resides in their ability to show how the inquiry focuses on non-psychological issues and how generative grammars appeal to non-psychological facts. Without such further support, challenges to the psychological conception would amount to little more than a prohibition to inquire into perfectly good questions. Chomsky would be justified in maintaining that generative grammar is an inquiry into psychological structures engaged in speaking and understanding language, insofar as he has offered a cogent account of that inquiry, unless there is a good argument for an

¹⁹ Katz (1981,1985, 1996)

alternative conception or for disputing that generative grammar is a promising avenue for investigating psychological structure.

Over the course of the twentieth-century, language has been conceived in various ways as a topic for systematic inquiry. Amongst these conceptions are conceptions of language as a collection of social facts, as something observable in the physical environment, as something abstract and immutable and as a structured psychological state. In the remainder of this chapter, I'll outline the psychological conception of generative grammar and Devitt's argument against that conception, before considering a range of non-psychological conceptions of language as a topic of inquiry.

1.2 The Psychological Conception of Linguistics

Chomsky claims that generative grammar is concerned with a class of psychological states engaged in speaking and understanding language. So, this is one way to conceive of language as a topic of inquiry: as a special cognitive proficiency that we have, like that we have for seeing, hearing or reasoning.

Barring special cognitive impairment or extremely hostile environmental conditions, all humans acquire at least one language. Moreover, a central component of this capacity for language seems to be species-specific. The nature of animal communication is a huge area of scientific research. Many other animals do make noises and gestures to communicate with varying degrees of sophistication. There are even a number of species that have very impressive abilities to discriminate between human speech sounds.²⁰ But the capacity to combine expressive elements, building hierarchical structure according to recursive rules, over an unbounded range seems to be unique to humans.

Considering the human *faculty for language* (FL), Hauser, Chomsky and Fitch (henceforth, HCF) point to this capacity as the key difference between human and non-human communication systems. They say:

²⁰ See HCF (2002) p.1574 and works cited there.

Animal communication systems lack the rich expressive and open-ended power of human language (based on humans' capacity for recursion).²¹

The recursive rules that characterise human languages generate a potential infinity of expressions and engender special hierarchical dependencies amongst the elements of these expressions. Other animals possess a very limited capacity to learn the long-distance dependencies that populate such utterances as *Jim, that bloke that's always down the pub, drinking beer and eating peanuts, shaves himself*. Other species apparently lack the ability to recombine meaningful units to form an unlimited variety of larger structures, each of which differs systematically in meaning. And studies of animal capacities for generating other discrete infinities, such as the numbers suggest that:

Animals can represent number approximately as a magnitude with scalar variability ... with greater discriminability among small numbers than among large (keeping distances between pairs constant) and between numbers that are farther apart (e.g. 7 versus 8 is harder than 7 versus 12). The approximate number sense is accompanied by a second precise mechanism that is limited to values less than 4 but accurately distinguishes 1 from 2, 2 from 3, and 3 from 4; this second system appears to be recruited in the context of object tracking and is limited by working memory constraints... The system apparently never takes on the open-ended generative property of human language.²²

Monkeys taught to count items up to three or four and to manipulate these numbers do not spontaneously progress in their understanding of number but have to be taught the whole process afresh for each new number.

Imagining a Martian scientist investigating life on earth, HCF surmise that:

[I]t might note that the faculty mediating human communication appears remarkably different from that of other living creatures; it might further note that the human faculty of language appears to be organized like the genetic code – hierarchical, generative, recursive, and virtually limitless with respect to its scope and expression.²³

²¹ HCF (2002) p.1570

²² HCF (2002) p.1577

²³ HCF (2002) p.1569

This infinite expressive power is responsible for the prolific and flexible nature of linguistic communication. It makes for a qualitative difference between human communication in language and the communicative efforts of species that lack language. Further, this discrete infinity is built out of a stock of more than 50,000 elements, far exceeding the number which can be stored by other creatures.²⁴ Apes that are trained intensively to use gestures or a keyboard come to grasp, at most, around 150 items and they do so laboriously. They can put them together in groups of two or three with little regard for order beyond the short sequences or repetitions that prompt reward, and they have no capacity for hierarchy. They lack that “spurt” in the growth of their combinatorial abilities that all human children experience, allowing children to go on building bigger and bigger structures.²⁵

This combinatorial proficiency, a cognitive capacity that only we have, naturally suggests itself as a locus of study if we want to understand language as something that we have but other species lack. As Barry Smith puts it:

We are unique both in our handling of recursive structure and our capacity to encode this in a limited range of speech sounds. And this ability to apprehend and integrate, so rapidly, the phonetic, syntactic and semantic information in virtue of which sound events are recognised as linguistically significant speech is a staggering achievement and requires explanation.²⁶

Proponents of the psychological conception are impressed and intrigued by these special properties of the human propensity for language. They conceive of linguistics as an investigation of this psychological capacity for building these structures distinctive of linguistic communication.²⁷ The central questions about this capacity concern our knowledge of linguistic structure, its acquisition and its use. Proponents of the psychological conception, like John Collins, think it plausible that

²⁴ The figure comes from Wallman (1992), Pinker and Jackendoff (2005) suggest that this is “more than 100 times the most extravagant claims for vocabulary in language trained apes or in natural primate call systems.”

²⁵ See Smith (2006) for discussion.

²⁶ Smith (2006)

²⁷ This interest in the productivity of language goes back to early work in generative grammar (see Chomsky 1965).

the subject matter of generative grammar is a reflection of these central questions about language. He says:

The psychological nature of linguistics flows from general considerations about the central questions we ask in linguistics: What do we know? How do we acquire the knowledge? How do we put the knowledge to use?²⁸

The psychological conception of generative grammar begins with the recognition that there is a state of the human brain that supports this psychological capacity for structuring sounds and marks linguistically. As this structure is conceived at a level of abstraction from the physical behaviour of the brain, we can call this psychological state, a state of the *mind/brain*.²⁹ As an idealisation, we can assume that we all start out with a FL in the same state prior to grammar acquisition. Grammar acquisition involves the development of this state of the mind/brain, FL, under the causal influence of the linguistic environments in which children find themselves.

This state of the human mind/brain, FL, has been referred to variously as a *grammar*, a *language faculty*, a system of *grammatical competence* and an *I-language*. The latter terminology, *I-language*, brings out the identification of language as a topic of inquiry with this special state of the mind/brain.³⁰ The *I* stands for three things. An *I-language* is an *internal* system of the mind/brain. It is *individual* to the speaker rather than a property of communities of speakers and can be specified independently of properties of the individual's environment. It is *intensionally* specified as a recursive procedure for pairing sounds with structured descriptions over an unbounded range, rather than a merely extensional specification of the set of paired sounds and sentences. *I-languages* are identified

²⁸ Collins (2006)

²⁹ Chomsky (1987 p.178) says that for the grammarians purposes we are to “understand talk about the mind to be talk about the brain undertaken at a certain level of abstraction from (as yet unknown) mechanisms, at which we can formulate an explanatory theory.” Jackendoff (2002) adopts the terminology *f-mind*, *f-* to suggest that we are talking about the structure of the brain at a functional level of description.

³⁰ See Chomsky (1986) for its introduction.

not with the utterances or written texts that speakers produce, but with an internal cognitive structure that shapes this production. I-language is part of our biological endowment, much like our visual system.

On this psychological conception of languages as I-languages, generative grammarians aim to characterise the initial state of the I-language and its development into its mature states. They also aim to understand the integration of the I-language with the other cognitive systems of the mind/brain that put this I-language to use. As Smith puts it, the psychological conception is concerned with the properties that make our use of language *language*; those psychological properties that give form and character to this distinctively human capacity to speak and comprehend.³¹

Chomsky adopted the term *I-language* in an attempt to disambiguate between two ideas.³² Chomsky had previously talked about the *grammar* of a speaker's language, explicitly using *grammar* to cover both the linguistic structure of a speaker's mind and the linguist's theory of that structure. I-language is not a new notion but rather a new term introduced "in the hope of overcoming persistent misunderstanding of the technical notion 'grammar'".³³ With I-language there was no opportunity for confusion as it was defined explicitly from the start as a state of the mind-brain:

Since the origins of work in generative grammar in the 1950s, it has been pointed out that the term "grammar" is being used with a systematic ambiguity: to refer to the internal states of Jones's faculty of language FL, and to the linguist's theory of that state. But that usage proved confusing. I therefore suggested that we restrict the term "grammar" ("particular" or "universal grammar") to the theories constructed by the linguist and refer to the internal state that grammars seek to describe as I-languages ("I" to suggest internal, individual, intensional).³⁴

³¹ Smith (2006)

³² See Chomsky (1986).

³³ Chomsky (2003) p.270

³⁴ *ibid*

In this thesis I'll often talk about the state of *grammatical competence* or the *grammatical competence system* which also serves to distinguish an aspect of the mind-brain from a grammarian's theory thereof.

When Chomsky coined the term *I-language* he also intended to draw a contrast with a rather different conception of language as a topic of inquiry, the *E-language* conception. The *E-* in E-language stands for external and extensional. On E-language conceptions, languages are external to the minds of speakers, extensionally specified as a set of pairs of sounds or sentences and meanings, which are not the property of an individual speaker-hearer. More specifically, Chomsky has used the term "E-language" to apply to conceptions of language as a social, public or otherwise non-individual phenomenon whose properties are fixed by convention or practical accomplishment. Chomsky's own evaluation of these E-language conceptions is that they are poorly defined, often incoherent and parasitic on the I-language conception for whatever theoretical interest they do have.

This is apt to sound very radical to philosophical ears. Influential philosophical figures, such as Quine and Wittgenstein, have argued that language is an inherently social and public phenomenon.³⁵ And furthermore, part of our commonsensical conception of language may be to think of languages as "out there", independent of their individual speakers.³⁶ In more reflective moments, we might think about languages as not only "out there" but also as having all sorts of interesting non-psychological properties such as aesthetic properties, as well as historical and cultural significance.³⁷ So the notion of a partially shared, public language of which speakers partake may be a feature of common thought and

³⁵ Quine (1960), Wittgenstein (2001)

³⁶ Caution is required here because investigating our ordinary conception of language is a scientific endeavour in itself: a part of ethnoscience. There are all kinds of disputes in psychology and philosophy about our commonsensical conceptions of physics, objects, psychology etc., so why not language too? Over 70% of the world's speakers are illiterate and it may be that cultures which lack written texts and other visual means to linguistic communication with durable environmental elements are less predisposed to think of languages as "out there", persisting objects of which they partake.

³⁷ Wiggins (1997) appeals to such properties of language as part of his attack on the psychological conception.

parlance. In speaking of the rules of *English* or *German*, it appears that linguists themselves happily appeal to such shared languages, but the appearance may be misleading.

In Chomsky's view, there is nothing peculiar or novel about the I-language perspective in scientific linguistics. He says:

It is not accurate to speak of "rejection of the notion of 'shared public language'" as peculiarly mine; it is a commonplace of the empirical study of language. In the technical literature one finds such terms as "Chinese" or "Italian", but these are understood to be mere conveniences, not notions with some place in the theoretical explanation of phenomena of language. The burden of proof falls on those who claim otherwise. To my knowledge, the burden has not been met, rarely even recognised.³⁸

Chomsky thinks that a lack of interest in E-language conceptions - conceptions of *Chinese* or *Italian* as external, public and partially shared - is just a feature of the way that scientific linguistics proceeds, absent any substantive proposals about the role such conceptions might play in grammatical theory. Chomsky's suspicion is that, to the extent that any sense has been made of E-languages, they are just interest-relative abstractions from the I-languages of sufficiently similar groups of speakers. They are assimilations of speakers whose I-languages are roughly similar according to linguistic principles and who are similar enough according to non-linguistic factors of geography, culture and history.

Hence, Chomsky surmises that E-languages are of little scientific interest. Though they may be of interest to the ethnoscientist who investigates our ordinary understanding of such concepts as *language*, our commonsensical conceptions have little bearing on how we should conceive of language as a topic for scientific inquiry. Chomsky thus views our ordinary conceptions of shared languages as much like our ordinarily thinking about the setting of the sun: such conceptions are not part of explanatory theory. There is no constraint on any scientific theory to stick to our commonsensical understandings; this extends to linguistics.

It's worth pausing to consider why Chomsky does not think that linguistic inquiry should be constrained by the way we ordinarily think about languages or by

³⁸ Chomsky (2003a) p.270

philosophical systematisations of ordinary thought. Chomsky is committed to *methodological naturalism*: a commitment to apply successful scientific methods to the study of linguistic and mental phenomena, in the hope of eventually integrating accounts of language and the mind with the core natural sciences.³⁹ Methodological naturalism does not preclude other ways of investigating language but rather provides a particular form of understanding. This sort of scientific understanding is supposed to give us some objective, precise and testable explanatory theory supported by empirical evidence. This is clearly a different sort of understanding to that which might be afforded by the study of our ordinary concepts or by arts such as history and literature. Chomsky claims that despite widespread acceptance of methodological naturalism, critics of the psychological conception typically critique the notions it employs and the theoretical framework in which they are embedded, on grounds other than their adherence to scientific methodology and the explanatory insight they offer.

To take an example, Chomsky sees John Searle's attack on the psychological conception of generative grammar as an instance of this tendency. Searle claims that the theories offered by generative grammarians cannot be explanations of speakers' minds because they fail to meet a general criterion for describing the mind. This criterion is Searle's "connection principle": the content of any *mental* states attributed to a subject must be consciously accessible in principle.⁴⁰ Chomsky argues that Searle's connection principle is at root *methodological dualism*: the supposition that the study of mind and language should be held to independent standards, beyond those that guide scientific inquiry in other domains. Without such dualistic strictures, what counts as psychological is just what our best theories of psychological phenomena discover. Chomsky thinks that much of the criticism of the I-language perspective and support for E-language conceptions is underpinned by such methodological dualism, holding the scientific inquiry into language to independent commonsensical or aprioristic standards, and that this is unjustified.

³⁹ See Chomsky (1995a, 2000) for discussion.

⁴⁰ Searle (1992) p.156

To recap, on Chomsky's psychological conception, when grammarians make claims about reflexives and their binding, they are making abstract claims about the psychological states of speakers. When they talk about *English*, this just serves to conveniently pick out the I-language states of a roughly similar population of speakers. If speakers sufficiently like us in the language they speak and understand find certain constructions that are in accordance with our hypothesised linguistic principles to be unacceptable then we might choose to revise those principles. Alternatively, we might take them thereby to be sufficiently dissimilar to us, to think that they speak a different language encoded in a differently organised I-language system. There are empirical issues about the similarities and differences between speakers but not about the notion of *English* which is informal and open-ended. The empirical issue is that there are speakers with differently configured I-languages, who began with the same biological endowment for I-language; and this requires some explanation. The similarity that loosely binds together English-speakers for the linguist is underwritten by a theoretical understanding of I-languages.⁴¹

In this thesis, the opposition I consider to the I-language perspective has a wider remit than Chomsky suggests for E-language conceptions. I consider non-psychological conceptions of language as something *external* to the mind, without any necessary commitment to languages being extensionally specified, social or shared, or conforming to ordinary conceptions.

It's noteworthy that what Chomsky means in locating languages internally, what he calls *internalism*, is different from what is usually meant by *internalism* in the philosophy of mind. Chomsky is explicit that by internalism he does not mean any doctrine that denies that "mental states are individuated by reference to features of the subject's environment or social context", or that holds that "subjects in the same internal (neural) state are in the same mental states".⁴² Chomsky's internalism is a distinctly methodological doctrine:

⁴¹ See Collins (2006) for discussion of these issues.

⁴² See Chomsky (2003). See Burge (1979, 1986) for classic arguments for the "externalist" or "anti-individualist" view that psychological states are individuated by reference to features of a subject's environment and/or social context. For an individualist view of psychology, see Fodor (1980). For opposition to Burge, see Segal (1989, 2000).

Internalism studies internal states, including those involved in what tradition (and common sense) often regard as mental aspects of the world. It may develop a concept of “mental state”, but if so, that will be a technical construct, finding its place within a particular context of theoretical explanation, denying or asserting nothing about other usages.⁴³

Chomsky sees studies of human and animal behaviour as dividing into two components. One component is an attempt to discover the internal organisation of the organism that equips it for the behaviour. The second component involves understanding how that internal organisation interacts with the environment. Chomsky’s internalism is his focus on the first of these components, and commits him to the fruitfulness of this approach.⁴⁴

The development of Chomsky’s conception of grammar is closely connected to his critique of behaviourist conceptions (on which I’ll say more in §1.3.1). In his famous review of Skinner’s *Verbal Behaviour*, Chomsky argued that behaviourist concepts of stimulus and response that had objective meaning in work on animal behaviour, became opaque, yielding trivialities or falsehoods, when applied to human language.⁴⁵ Chomsky argues that if we want to understand linguistic behaviour then the behaviour itself is the wrong target for a linguistic theory. For linguistic behaviour is itself the effect of a highly complex interaction amongst systems within humans of which only some are linguistic. For this reason Chomsky thought it unlikely that there could be an explanatory theory of language pitched at the level of behaviour, per se. Rather Chomsky suggested that it might be possible to characterise aspects of the knowledge of language that shapes linguistic behaviour through its interaction with systems for producing and perceiving language. He therefore proposed that the best object of study in linguistics is not verbal behaviour, which is itself the product of diverse and little understood systems, but rather our knowledge of linguistic form and meaning.

⁴³ Chomsky (2003) p.269

⁴⁴ See Williamson (2006) for an argument that the study of cognition cannot be factored into internal and external components.

⁴⁵ Chomsky (1959)

There is a corresponding thought about linguistic behaviour that seems to be shared amongst proponents of the psychological conception:

The cognitive perspective regards behaviour and its products not as the objects of inquiry, but as data that may provide evidence about the inner mechanisms of mind ... The properties and patterns that were the focus of structural linguistics find their place but as phenomena to be explained along with innumerable others, in terms of the inner mechanisms that generate expressions.⁴⁶

On the psychological conception, behaviour is not the primary locus of linguistic properties but is *linguistic* only derivatively of the role of special cognitive systems in shaping and apprehending it. It is the latter “inner mechanisms” which are supposed to explain the special properties of linguistic behaviour and which are, therefore, taken to be the primary target of linguistic inquiry.

One source of theoretical disagreement amongst linguists who endorse the psychological conception is over which phenomena form part of the core FL and which are part of the conceptual-intentional systems of thought and the other cognitive systems that interface with the grammatical system. By Chomsky’s recent definition, FL narrowly conceived (FLN) is that cognitive property distinctive of linguistic communication that equips us for recursive structure building. FLN is to be distinguished from language more broadly conceived (FLB) which involves the integration of recursive structures into systems for intention-conceptualisation and perception-articulation. Chomsky claims that, unlike the recursive structure building of FLN, the mechanisms that are part of FLB are not specific to humans.⁴⁷ One consequence of Chomsky’s recent view is to push some of what were once

⁴⁶ Chomsky 2000 p.5-6

⁴⁷ There is much controversy about which of the FLB mechanisms are replicated in other animals and to what degree. For example, our categorical perception of speech, which seems finely tuned to the details of human speech, may not constitute a uniquely human adaptation. There is evidence to suggest that macaques, chinchillas and birds have the capacity to discriminate amongst the human speech sounds. See HCF (2002) p.1572 and works cited there.

considered core principles of FL into the interface between language and other systems which are part of FLB.⁴⁸

There are also disagreements about how the I-language is to be understood amongst those who advocate the psychological conception. Some who adopt the psychological conception challenge the idea that the properties of I-languages are all properties of the individual speaker. Ludlow, for example, has suggested that I-languages have referential, semantic properties; properties which cannot be specified at the level of the individual in independence from their environment. In investigating the referential semantics of I-languages, Ludlow suggests that the linguist will be appealing to features of the world beyond the individual speaker.⁴⁹ The referential properties of I-languages are relations between the internal representations and aspects of the world beyond the speaker's mind. If Ludlow is right then I-languages would be internal and intensional but not individual. Chomsky himself is sceptical about a science of reference and about the idea that natural languages have referential semantics.⁵⁰ This interesting issue will not be addressed in this thesis.

And there are disagreements over the commitments of conceiving of I-languages as psychological states. Chomsky has a very thin notion of what it takes for a state to be "cognitive", "psychological" or "mental" and uses the terms interchangeably. On Chomsky's view, there are no substantial criteria that something has to meet in order to be "mental" or "psychological" beyond its postulation as part of our best theories of mental phenomena. For Chomsky, "mental" and "psychological" are rough and ready notions of commonsense, and there is no "mark of the mental" that serves to sort the mental from the non-mental independently of our developing psychological theories. He says that his approach is "'mentalistic' but in what should be an uncontroversial sense. It is concerned with

⁴⁸ See Safir (2004) for discussion of whether binding falls within FLN or at the semantic interface. For more general discussion of HCF's claims see Pinker and Jackendoff (2005), and HCF's replies in Fitch, Chomsky and Hauser (2005) and Chomsky, Hauser and Fitch (2005).

⁴⁹ Ludlow (2003), Chomsky (2003c)

⁵⁰ Chomsky (2000)

“mental aspects of the world”, which stand alongside its mechanical, chemical, optical and other aspects.”⁵¹

As Chomsky sees it, terms such as “mental” serve only to loosely carve out a domain of inquiry rather than to make a substantive claim about a phenomenon.

I will be using the terms “mind” and “mental” here with no metaphysical import. Thus I understand “mental” to be on a par with “chemical”, “optical”, or “electrical”. Certain phenomena, events, processes, and states are informally called “chemical” (etc.) but no metaphysical divide is suggested thereby. The terms are used to select certain aspects of the world as a focus of inquiry. We do not seek to determine the true ‘criterion of the chemical’, or the ‘mark of the electrical’, or the ‘boundaries of the optical’. I will use “mental” the same way, with something like the ordinary coverage, but no deeper implications.⁵²

The domain of psychological inquiry includes some phenomena for which our brains, and not other parts of our bodies, are chiefly responsible. Our knowledge of linguistic structure is one such phenomenon, though other functions the brain serves, such as regulating our body’s temperature, happen not to be included in our ordinary classification.

In keeping with this attitude to notions like *mental* and *psychological*, Chomsky is happy, unlike Searle, to posit mental states of which speakers have no conscious awareness and never will. Chomsky suggests that conscious states have no special place in the inquiry into language, forming a “scattered and probably uninteresting subpart of the full cognitive structure”.⁵³ Smith has pointed out the contrast here with philosophers like Dummett who afford conscious psychological states a special role in characterising a speaker’s language.⁵⁴ Dummett distinguishes:

[B]etween those regularities of which a speaker, acting as a rational agent engaged in conscious, voluntary action, makes use from those that may be hidden from him and

⁵¹ Chomsky (2000) pp.5-6

⁵² Chomsky (2000) p.106

⁵³ Chomsky (1975) p.163

⁵⁴ Smith (2006)

uncovered by a psychologist or a neurologist: only those regularities of which, in speaking, he makes use characterize the language as a language.⁵⁵

But most proponents of a psychological conception of language will insist for empirical reasons, as Smith does, that regularities of which a speaker has no conscious awareness do serve to characterise his language. Remember how the reflexive generalisation constrains our interpretation of (1) so that *myself* depends for its meaning on *I* and see how it shapes our understanding of sentences like (4) and (5) so that it is *he* that fixes the interpretation of *himself*.

(1) I shaved myself.

(4) He shaved himself.

(5) We knew that he shaved himself.

But our reflexive generalisation is not a generalisation of which speakers make “use” in Dummett’s sense, for they are not consciously guided by it or consciously aware of it. As Smith says:

Dummett is surely wrong, on empirical or evidential grounds alone, to insist that the only generalisations or regularities that feature as part of our language are those of which we have conscious apprehension. That would be a hopeless move in the face of the generalizations linguistic theory gives rise to, and for which there is evidence in the conscious intuitive judgements of speakers.⁵⁶

The empirical evidence does not support Dummett’s view because the generalisations in which linguistic theory traffics are, for the most part, beyond the conscious reach of speakers. It took careful examination of linguistic data and some theoretical inference to get to the Reflexive Generalisation. The facts about a speaker’s I-language are conceived as facts about the speaker’s mind that outstrip what is consciously available to him. The states of an I-language are unconscious

⁵⁵ Dummett (1993) p.104

⁵⁶ Smith (2006) p.968

states of the mind/brain that are *partially explanatory* of the speaker's conscious judgement and his use of language.

Chomsky's view is that the psychological is just whatever our best theories of "psychological" or "mental" phenomena, such as knowledge of language, tell us it is. He thinks that there is nothing more substantive to be said than that, independently of the results of empirical inquiry. But there are adherents to the psychological conception who make further substantive claims about what marks the domain of inquiry out as psychological and about what sort of psychological inquiry generative grammar is.

In particular, it has been suggested by some that I-languages are *intentional* or *representational* systems.⁵⁷ Advocates of such views think that the best way to understand the cognition of language involves states of the I-language *representing* linguistic properties or having linguistic properties as their *intentional content*, conceived as their truth- or correctness-condition.⁵⁸

Chomsky has sought to distance himself from the idea that I-languages are intentional or representational in the above sense. Chomsky's view is that the notion of representational or intentional content is obscure and unexplanatory. He says:

I do not know of any notion of "representational content" that is clear enough to be invoked in accounts of how internal computational systems enter into the life of the organism. And

⁵⁷ This component is adopted by Egan (1995, 2003), Rey (2000, 2003, 2005) and Smith (2006, 2006a, 2008) amongst others.

⁵⁸ This view is not ruled out by recent moves to a phase/derivational model, free of levels of representation as Hinzen suggests. Hinzen (2006 p.250) says: "If this line of thought is carried to the limit, the notion of representation ceases to play an explanatory role ... At least intuitively, a derivation does not represent anything, it just proceeds." But as Collins points out (unpublished ms. b p.6) intuitively a derivation does require representations, if this just means structures that can be built and transformed, so "doing without levels of representation is not to do without representations". Whether the notion of representation should be understood in terms of content or aboutness, or merely formal structure would still be an open question. Chomsky (2000), and Collins (2004, 2006, 2007, 2008a) favour a structural or categorical understanding of a representation (x-representations rather than representations of or about x) which eschews the notions of content and intentionality.

to the extent that I can grasp what is intended, it seems to me very questionable that it points to a profitable path to pursue.⁵⁹

Though I do not intend to enter into this issue in my thesis, it is noteworthy that there are those amongst the supporters of the psychological conception who would have a fourth “I” – that of *i*ntentional – added to the list of internal, individual and intensional. My question about whether the grammatical properties described by linguists are psychological properties is prior to the question of whether the psychological states that realise those properties are representational or intentional states.

A final noteworthy difference amongst supporters of the psychological conception is over the role that *knowledge* plays in the psychology of language. We saw that Chomsky denies E-languages have any explanatory role in linguistics. So, on his account, whatever our knowledge of language consists in, it cannot be a relation between internal states of speakers and the properties of an E-language. Chomsky eschews such independent objects with which the states of the I-language can align or fail to align. There is just the I-language: an internal mental state that effects a generative procedure for assigning structural descriptions to expressions. From Chomsky’s perspective, “knowledge of language” is just an informal locution that serves to pick out the relevant states of the speaker. Chomsky says “we should think of knowledge of language as a certain state of the mind/brain...furthermore as a state of some distinguished faculty of the mind – the language faculty – with its specific properties, structure and organisation.”⁶⁰

Though Chomsky thinks there is a perfectly reasonable sense in which generative grammar is an inquiry into a species of knowledge⁶¹ he is happy to give up the term and talk about our *cognition* of language, rather than knowledge, in the face of philosophical scruples about consciousness, justification and the alleged relational nature of knowledge. One’s “knowledge of language” so conceived is then the I-language state and its integration with other cognitive systems. Though

⁵⁹ Chomsky (2003) p.274. Jackendoff (2002) also eschews the notion of representational content.

⁶⁰ Chomsky (1987) p.178

⁶¹ See Chomsky (1980, 1986)

“knowledge *of* language” suggests a relation between the knower and that which is known, for Chomsky such a relational conception collapses without E-languages.⁶²

This Chomskian view that the correct theoretical conception of knowledge of language is as a state, the knowledge-of-language state, has been defended by Collins. Critiquing the relational conception, he says:

There is, as it were, nothing objective to *get right*. Languages are not external objects we can go right or wrong about... It turns out, as part a matter of discovery, part methodology, that we do not *know* languages.⁶³

Though Collins holds that there is a state we might think of as the knowledge-of-language-state, this is not a commitment to the relational conception.

But there are those, such as Smith, who envisage a more substantial role for a relational conception of knowledge.⁶⁴ The conception of knowledge of language as the I-language state eliminates the problem of specifying the relation between speaker’s mental states and the languages they know: knowledge of language, from this perspective, just is the I-language state. But Smith claims some such relation is required if we are to defend the view that people know languages - a view that he takes to have significant virtues - since knowledge “requires there to be a subject matter to be right or wrong about”.⁶⁵ This is Smith’s “Missing Object of Knowledge Problem”:

If languages are no longer seen as “out there” but are conceived as internal to speakers, what is the object of speakers’ knowledge of language? ... [W]e seem to have swapped one problem for another since there now appears to be nothing to relate them to: no *object* of knowledge.⁶⁶

⁶² Jackendoff’s (2002) view is similar: a speaker “knows a language” if their f-mind includes certain integrated structures.

⁶³ Collins (2004) pp.514-6

⁶⁴ See Smith (2006, 2006a, 2008) and Matthews (2006)

⁶⁵ Smith (2006) p.971

⁶⁶ Smith (2006) p.974

According to Smith, the trouble with the conception of knowledge of language as simply the I-language state, rather than a relation to linguistic fact, is that the state “fixes the facts of language rather than conforming to them”.⁶⁷

Smith thinks that although the facts about a speaker’s language are determined by their I-language, there is still room for their having knowledge of linguistic facts. On Smith’s account, our knowledge of language is conscious knowledge of particular linguistic facts, linguistic facts which are determined by the unconscious and general principles of our I-language. Smith’s account exploits the gap between the linguistic facts determined by our I-language and the verdicts of conscious judgements, which are higher level effects of a variety of cognitive factors, to get the distinction between something seeming to me to be a part of my language and its actually being so. In this sense, speakers’ judgements can succeed or fail to track the linguistic facts:

The speaker can be out of step with his own linguistic system...it is only when the effortless and groundless intuitive judgements speakers make flow from (in whatever way they do), and conform to, the assignment of structure made by the underlying I-language that the speakers’ intuitions count as knowledge.⁶⁸

When judgements are out of step with the I-language, speakers can lack knowledge of their own language. Although, phenomenologically, it might strike us that in making linguistic judgements we are reacting to linguistic facts, ‘out there’ and external to us, on Smith’s account, “we are reacting to something within our own breasts and not consciously accessible to us at all.”⁶⁹

Smith counts the cost of giving up on the relational conception more dearly than Collins. He wants to hold onto the idea that speakers have a large of number of items of specific knowledge about their language including authoritative first-personal knowledge of what their words mean, which arrangements of words are sentences of their language and how these sentences can and cannot be understood. The view that we know the meanings of our own words and their legitimate

⁶⁷ Smith (2006) p.975

⁶⁸ Smith (2006) pp.977-8

⁶⁹ Smith (2006) p.978

arrangements has its independent attractions. I know what I am saying and, on the whole, I know when you are saying something in a language I understand.⁷⁰ And Smith wants to integrate the linguist's account of the I-language state with an account of first-personal knowledge of language.

Smith is more sanguine than Collins about the prospects of an account of how the I-language states bear on personal-level states of conscious knowledge. Collins claims that our first-personal knowledge of language is about as complex an interaction effect as we could imagine. Consequently, Collins thinks we have little idea how to develop a theoretical framework in which to embed the view that first-personal states conform with states of the I-language. Smith's thought is that, given the absence of an external object of knowledge, any such framework for integrating the I-language and personal-level accounts will have to involve a constraint of tracking which holds between the first-personal states and the I-language states.

Smith's second motivation for building knowledge of language, relationally understood, into the psychological conception is his view that for speakers' linguistic judgements to be useful as evidence they must constitute knowledge of the language about which they are judgements.⁷¹ For example, most English-speakers consciously know that in (1) *I* means the same thing as *myself*. Smith thinks that if such judgements did not constitute knowledge of our language, as distinguished from mere responses to presented material, then they would not be good data for hypotheses about the language we have mastered.

On Collins' non-relational view, there is no requirement that the judgements constitute knowledge or bear any epistemic credentials at all. There is only a

⁷⁰ Compare Matthews (2006 p.215): "When it comes to justifying our reliance on language as a reliable means of communication...any explanation of the role of linguistic competence as the ground of this authority is going to have to characterise this competence in epistemic terms...[because] we have to *know* this pairing [that I-language effects] so that we can know what it is we and others are saying."

⁷¹ Matthews (2006) presses the same point: "Linguists need some justification for crediting these judgements as an accurate expression of the informant's linguistic competence. But there would seem to be no reason to credit these judgements unless informants actually *know* certain things about their language, e.g., the pairing of sounds and meanings that their language effects, and furthermore their judgements actually express this knowledge."

theoretical inference from the judgements to an ensemble of underlying systems of which the I-language is one. The knowledge-of-language state properly integrated into wider systems, yields linguistic judgements.

There is far more to be said on each side of this disagreement but I will not attempt to resolve the dispute here. Both sides agree that the grammatical properties studied by generative grammarians are properties of a psychological system. My question of whether linguistics is a part of psychology is therefore antecedent to this disagreement over the nature and role of knowledge within the psychological conception of language.

A common critical response to Chomsky's psychological conception of linguistics as an investigation of I-language is that it is a plain confusion: "*Language* is one thing, our *knowledge of language* or *linguistic competence* is another", so the response goes. When we think about the languages we know, we typically think of English, Russian and the like, of which we can come to have knowledge, and not of something encoded in our heads. Scientific linguists find these ordinary classifications of *English* and *Russian* unenlightening. But, so the response goes, linguists still appeal to *languages*, perhaps precisifications of *English* and *Russian* that individual speakers may have knowledge of or be competent in. The perceived confusion is that linguists are explicitly committed to the existence of such *languages* and that there is a deeply entrenched distinction between the study of language and the study of the psychology of language.

And perhaps we should expect such a distinction. Just as there is a distinction between the logical implications and even an ideal reasoner's competence with the implication relations, or a distinction between mathematical truths and our mathematical competence; so we should expect a distinction between the grammatical structure of a language and the system of grammatical competence. Just as there is a special human propensity for language so there is for logical deduction and also for mathematics. But in light of the distinctions, one would presumably not want to hold that logic and mathematics are a part of psychology.

Frege brought out the unpleasant consequences of thinking of the laws of logic as laws of empirical psychology. For imagine that they were and that there

were beings whose thoughts were governed by very different laws that contradicted ours. What would we then say about them?

The psychological logician could only acknowledge the fact and say simply: those laws hold for them, these laws hold for us. I should say: we have a hitherto unknown type of madness. Anyone who understands laws of logic to be laws that prescribe the way in which one ought to think – and not natural laws of human beings taking a thing to be true – will ask, who is right?... The psychological logician cannot ask this question: if he did he would be recognising laws of truth that are not laws of psychology.⁷²

Husserl, considering the laws of logic as they stand to the laws of thought, asks us to imagine an ideal thinker, in whom *all* thinking proceeds as the logical laws require:

Naturally the fact that this occurs must have its explanatory ground in certain psychological laws, which govern the course of the mental experience of this being... Would the natural laws and the logical laws in this assumed situation be one and the same? Obviously, the answer is 'No'. Causal laws, according to which thought must proceed in a manner which the ideal norms of logic might justify, are by no means identical with those norms.⁷³

Proponents of the psychological conception need to convince opponents that the study of language is different: that where there are good reasons to distinguish between the study of logic or mathematics and the study of our cognitive capacities in that direction, there are no such good reasons to distinguish between studies of linguistic structure and of the psychological structure of those that speak and understand language. Knowing all the truths about human psychology, we would still not have resolved all the mathematical questions. So why should we expect our knowledge of the truths about human psychology to reveal all the truths about the grammars of natural languages? As Katz puts the point, there seems to be:

⁷² Frege (1967) p.14

⁷³ Husserl (1970) p.103

[A] fundamental distinction between the *knowledge* speakers have of their language and the *languages* that speakers have knowledge of. This distinction is simply a special case of the general distinction between knowledge and its object.⁷⁴

In the next section I'll examine Devitt's recent development of this style of argument applied to the study of grammar. Then I'll look at the different conceptions of language as a topic of inquiry that are on offer if one endorses Devitt's argument against the psychological conception.

1.3 Non-psychological Conceptions of Linguistics

In his recent book *Ignorance of Language*, Devitt argues that we ought to take generative grammars to be true not of the psychological states of speakers, but instead of a *linguistic reality* distinct from speakers' psychological states.⁷⁵ Devitt has a particular conception of this linguistic reality which he calls the *linguistic conception*. His argument for the linguistic conception of generative grammars over the psychological conception involves drawing three quite general distinctions and

⁷⁴ Katz (1981) p.78 As far as Katz's own Platonist proposal for distinguishing our knowledge of language from the natural languages themselves goes, Chomsky (1986 p.33) takes a rather different view: "There is no initial plausibility to the idea that apart from the truths of grammar concerning the I-language and the truths of UG concerning [the initial state of FL] there is an additional domain of fact about P-language [a Platonic object], independent of any psychological states of individuals."

⁷⁵ Devitt calls this view the *Representationalist Thesis*: the view that speakers stand in an unconscious or tacit propositional attitude relation to the principles of their language which are represented in their language faculty. Devitt argues that, if knowledge of language is so understood, a competent speaker could be completely ignorant of their language. In the first main chapter of the book (pp.1-40) he offers a self-standing argument for the view that linguistics is not a part of psychology. Having satisfied himself that generative grammars are not characterisations of speaker's psychological states he considers a range of positions on the psychological reality of grammar, none of which he finds to offer any support for the Representationalist Thesis. It is the argument that linguistics is not a part of psychology, and not the issue about the Representationalist Thesis, which is my concern. Though Devitt's argument against linguistics being a part of psychology is intended to be self-standing, he takes it to be supported by a particular view of linguistic intuitions as they serve as evidence for grammars. I discuss this view of linguistic intuitions in Chapter Four.

then defending their applicability to the science of language. The distinctions are general in the sense that they are supposed to apply to any domain in which we find a competence and its products, and can be illustrated without appeal to the specific features of language. The first is:

- I. The distinction between a theory of competence and a theory of the products of that competence.

Devitt illustrates the distinction using the example of a blacksmith and the horseshoes he makes. The blacksmith has a competence, a competence in making horseshoes. The horseshoes themselves are the products, or outputs, of the blacksmith's competence. The nature of the produced horseshoes is one thing and the nature of the blacksmith's competence to produce such items is another. So, Devitt observes, one should distinguish between a theory of the blacksmith's competence to produce horseshoes and a theory of the horseshoes themselves.⁷⁶

Devitt's second distinction is between the sorts of rules that characterise competences and the sorts of rules that characterise the products of competences. The rules that characterise the operations of a competence, Devitt calls *processing rules*. The rules that characterise the system of products, Devitt calls *structure rules*. So his second distinction is:

- II. The distinction between the processing rules that characterise the operation of a competence and the structure rules that characterise the products of a competence.

Devitt illustrates this distinction using the example of bee dances. Bees returning from distant food sources produce a waggle dance. The positioning and pattern of this dance indicate the direction of the food source. Devitt notes that Karl von

⁷⁶ Devitt also discusses the examples of a theory of chess moves and a theory of chess competence, and a theory of well-formed formulas and a theory of a logic machine's "competence" to produce them.

Frisch won a Nobel Prize for his theory of the bee dance. Von Frisch hypothesised that:

To convey the direction of a food source, the bee varies the angle the wagging run makes with an imaginary line running straight up and down... If you draw a line connecting the beehive and the food source, and another line connecting the hive and the spot on the horizon just beneath the sun, the angle formed by the two lines is the same as the angle of the wagging run to the imaginary vertical line.⁷⁷

The dance is the product of a competence that the bees have for dancing. For Devitt, the dance is a representational system characterised by a set of structure rules and it is the task of a theory of the dance to discover those rules. But what a theory of the dance will not reveal is how the bees manage to produce the dance. The processing rules within the bee that enable it to perform this feat remain a mystery. For this, Devitt claims, one needs a distinct theory of the bee's competence, a specification of the processing rules that characterise this competence.⁷⁸

But Devitt allows that the theory of the bee dance will tell us something about the bee's competence. Whatever the processing rules that characterise the internal workings of the bee, they must produce dances that are characterised by the structure rules of the dance. Devitt allows for some deviations from the structure rules amongst the products, where there are errors in the processing and as a consequence the processing fails to produce outputs governed by the structure rules. But processing errors aside, Devitt thinks that the competence produces dances that are governed by the structure rules and he calls this relationship *respect*.⁷⁹

⁷⁷ Frank (1997) pp.181-2

⁷⁸ It is important to note at this juncture that Devitt here equates competence with processing. This is *not* the conception of competence that most proponents of the psychological conception of generative grammar appeal to. The standard view of grammatical competence is of a structured *state* of the mind/brain conceived as: a body of standing information that is integrated with independent processing systems for perceiving and producing speech. So it is usual to *distinguish* competence from processing. I develop this issue in Chapter Three.

⁷⁹ Devitt (2006, 2006a p.577). Devitt (2006 p.23) says "A theory of competence must posit processing rules that respect the structure rules of the outputs".

On the strength of a theory of the dance, Devitt claims we know that there is something “we-know-not-what” inside the bee that respects the structure rules of the dance, but we do not know the processing rules on this basis. We do not know *how* the processing rules respect the structure rules. In particular, we do not know whether any of the structure rules are also processing rules. This brings us to Devitt’s third and final distinction:

III. The distinction between processing rules respecting structure rules and the structure rules being included amongst the processing rules.

Whilst we know that the processing rules respect the structure rules, it would be a further hypothesis that those structure rules were amongst the processing rules. Devitt suggests that any further claims about the relationship between the structure rules and processing rules will require supplementary evidence or argumentation:

To move beyond the minimal claim [respect] and discover *the way in which* the bee’s competence respects the structure rules of the dance, we need evidence beyond anything discovered by von Frisch, evidence about the bee’s “psychology”.⁸⁰

Devitt suggests is that it is far more plausible to take the study of horseshoes or bee dances to involve theorising about horseshoes or bee dances themselves rather than the competences of the blacksmith or the bee to produce them. By analogy, a theory of language should be distinguished from a theory of linguistic competence, and we should take generative grammar to involve theorising about language. A language, as understood by Devitt, is the product of a competence in a language, rather than a competence to produce and understand linguistic material. According to Devitt, linguistics is not the study of a core system of linguistic cognition, I-language,

⁸⁰ Devitt (2006a) p.577. See also Devitt (2006) p.21-3. It is worth remarking that the relation between von Frisch’s theory and the bee’s cognitive capacity for dancing may be more intimate than Devitt suggests. As pointed out by Longworth (forthcoming p.13) it seems plausible that von Frisch was attempting to offer not just any old structural description of the bee’s dance, of which there may be many, but that set of structural relations to which the bees are sensitive in producing and recognising the dance patterns.

which he thinks Chomsky has conflated with a *language* or a *grammar*. Distinction I addresses this perceived conflation. Applying the general distinctions to the case of language, linguistics proper should be distinguished from the study of linguistic competence.

I. Distinguish between a theory of linguistic competence and a theory of the language which is a product of that competence.

Correspondingly, the structure rules that characterise the language speakers produce and understand should be distinguished from the rules that characterise the linguistic processing performed by the brain in production and consumption.

II. Distinguish between the rules that characterise the linguistic processing performed by the brain and the structure rules that characterise the language.

And finally, according to Devitt, the fact that the rules involved in processing language must respect the rules that characterise linguistic structure should not be confused with the hypothesis that the linguistic structure rules are amongst the linguistic processing rules.

III. Distinguish between the linguistic processing rules respecting the structure rules of the language and the structure rules being included amongst the processing rules.

One might wonder why Devitt needs all three distinctions. Devitt's desired conclusion - that a theory of language is distinct from a theory of linguistic competence - might appear to be arrived at by the application of distinction I. However, Devitt needs distinctions II and III to prevent the theory of language as a product, collapsing into the theory of competence. If linguistic competence and its products were characterised by one and the same set of rules then it might seem that the distinction between the two theories was merely notional. With that equation blocked, or at least shown to require supplementary argument, Devitt can, *prima facie* at least, draw a distinction between the theory of language and the theory of

competence in a language. The theory of language and the theory of linguistic competence will be mutually informing, on Devitt's model, but only insofar as the competence produces the language and thereby respects the rules described in linguistics proper.

How is the truth of Devitt's linguistic conception of generative grammar supported by his distinctions? What Devitt takes the distinctions to show is that "the truth of a grammar entails that its rules govern linguistic reality but does not entail that they govern psychological reality."⁸¹ On Devitt's model, the processing system respects linguistic structure; it produces outputs governed by the structure rules. It cannot but do so because it is a competence in the language that the structure rules describe. But this supports only a minimal proposal about the psychological reality of grammar (M):

(M) Something inside the speaker (we know not what) is responsible for producing outputs that respect a highly elaborate set of structure rules.⁸²

Devitt lays down the following gauntlet:

If the psychological conception is to be saved, there must be something wrong with the three distinctions or with their application to linguistics. It's as simple as that.⁸³

It is important in understanding Devitt's challenge to note how in drawing distinctions I – III, his chosen terminology deviates from the standard terminology used by grammarians. Devitt asks us to treat grammatical *competence* as a *processing* system and this is at odds with the orthodox understanding of competence established by Chomsky.⁸⁴ Chomsky makes a fundamental distinction between grammatical *competence*, as a system of underlying knowledge or information, and our linguistic *performances*, such as our perceptions and productions of utterances. Linguistic processing is a part of linguistic performance;

⁸¹ Devitt (unpublished ms. b) p.5

⁸² Devitt (2006) p.57, (2006a) p.578

⁸³ Devitt (2006a) p.578

⁸⁴ See Chomsky (1965) pp.4-62

we process speech in real-time as part of assigning structural interpretations to utterances and we build structures in real-time as part of producing utterances. But the underlying system of competence is not usually conceived as a system for the processing of linguistic material involved in our use of language. As standardly understood, competence is the underlying state of knowledge that encodes pairings of sounds and structures over an unbounded range: it is not a processor that once activated by a stimulus runs through a set of procedures for structurally describing speech and building linguistic structures for utterance.

Grammarians hypothesise that grammatical competence constrains or shapes our linguistic performances; that our use of language is informed by our knowledge of a grammar. So conceived, insofar as grammatical competence has products, they are not externalised performance events but inputs to systems internal to the mind. Most grammarians, following Chomsky, think that the explanation of our linguistic performance requires our understanding the factors involved in real-time language processing and that this is a distinct task from constructing the generative grammar that they take to be a theory of our grammatical competence.⁸⁵ Devitt's distinctions blur the competence-performance distinction, as standardly understood, by making competence a matter of processing. The complication this adds is that both Devitt and supporters of the psychological conception can agree that a grammatical theory should be distinguished from a processing theory.

Devitt says: "I do not take [a grammar] to be psychologically real in virtue of its meeting the Respect Constraint. But this difference may be merely verbal."⁸⁶ What Devitt denies is that generative grammars describe an aspect of our psychology in virtue of their being respected by a theory of competence, understood as a processing system. But Devitt's respect constraint, which he specifies as a constraint holding between the generative grammar of a speaker's language and the processing a speaker performs, could be understood as Chomsky's competence-performance distinction. Suppose we distinguish between the theory of the

⁸⁵ A richer understanding of linguistic performance would involve not only a theory of grammatical competence and the processing of linguistic structure but theories of other competences and levels of processing, semantic and pragmatic.

⁸⁶ Devitt (2006) p.67 fn.6

speaker's language, conceived of as a theory of the grammar he knows, and a theory of linguistic processing, with the linguistic processing meeting conditions imposed by the known grammar. Then we have a psychologically real generative grammar which stands in Devitt's respect relation to the processing theory. It is only Devitt's commitment to the non-psychological nature of what grammatical theories posit – i.e. to the external location of the linguistic structure rules – that makes his distinctions incompatible with the psychological conception. So the real disagreement between Devitt and the psychological conception he opposes is that Devitt will insist upon a distinction between the linguistic structure rules realised externally to the speaker and the rules that characterise the speaker's internal linguistic competence. Chomsky denies that generative grammar focuses on or requires such external structure. Devitt locates the structures that grammarians investigate external to the mind/brain, whereas Chomsky takes the investigated structures to be structures of the mind/brain.

Hence, Devitt's particular way of putting his challenge needs amendment because, on one construal, the distinctions that he presses his opponent to accommodate actually animate the conception he is opposing. Collins offers the following construal of Devitt's distinctions that is entirely in keeping with the psychological conception:⁸⁷

Revised Distinction I. A theory of the language faculty is distinct from a theory of its products, if such products are to be understood as performance events. This distinction is enshrined in the competence-performance distinction.

Revised Distinction II. The structure rules are the rules that generative theories posit and these are distinct from processing mechanisms. A theory of competence is not (directly) a theory of processing.

⁸⁷ See Collins (unpublished ms. a, 2007a, 2008, 2008a) but also Slezak (unpublished ms.) and Smith (2006a, unpublished ms.)

Revised Distinction III. Whatever the brain does in terms of processing, we take it to “respect”, i.e. meet, the conditions our competence theory specifies insofar as we take it to be a correct theory.

Devitt’s opponent can thereby distinguish a theory of linguistic structure from a theory of linguistic processing and accommodate the distinctions. Collins’ reconstruction shows how schematically similar Devitt’s views appear to Chomsky’s in distinguishing theories of grammar and of processing. The clear difference is that Chomsky and Devitt think of theories of language as theories of different domains. The point of divergence is that Devitt thinks the linguistic structure rules describe an aspect of our physical environment whereas Chomsky thinks they are best understood as describing an aspect of human psychology.

The challenge then is not whether Devitt’s opponent, who endorses the psychological conception, can show that there is “something wrong with the three distinctions or with their application to linguistics”. After all, on one widespread interpretation of the distinctions, the psychological conception accommodates them. The challenge is whether we ought to draw the distinctions as Devitt construes them; separating psychological and non-psychological theories in addition to drawing the relevant psychological distinctions.

In replies to critics, Devitt claims that no one defending the psychological conception has responded to his argument which is beguilingly simple: the three general distinctions apply to linguistics and thereby suggest that linguistics is not a part of psychology.⁸⁸ Collins is not the only proponent of the psychological conception to complain about Devitt’s conflating of competence and performance, and the way it obscures his argument.⁸⁹ A number of critics have claimed that, as they stand, Devitt’s distinctions provide no reason to deny that generative grammar is a part of psychology. Devitt’s distinctions do not specify what a language is or what the psychology of language involves, and so they do not make it perspicuous how the conclusion that linguistics is not a part of psychology follows from the distinction between a theory of linguistic structure and a theory of processing.

⁸⁸ Devitt (2006a, 2006b, 2007a, 2007b)

⁸⁹ Slezak (unpublished ms.), Smith (unpublished ms.)

Devitt's opponents, in particular Smith, by way of defending the psychological conception, maintain that Devitt only provides three imperatives to draw quite general distinctions and a claim that those distinctions apply to linguistics.⁹⁰ So the argument is elusive because it is unclear what is being claimed about the nature of psychology and the nature of language such that the study of the latter is not a part of the study of the former. Devitt tells us what his view of linguistic reality is, as described by his linguistic conception, but that conception was supposed to be the conclusion of his argument. The argument itself offers no specification of what language is. If the content of the distinctions is to be taken as a premise in an argument against linguistics being a psychological science, it has to be made clear what is being supposed about the nature of language. The argument can be elaborated so as to make clear what is supposed to justify the choice of Devitt's linguistic conception of generative grammar over Chomsky's psychological conception.

In a series of recent papers, Devitt's critics have determined the following elaboration of his argument which incorporates Devitt's view of linguistic reality:⁹¹

P1. There are three quite general distinctions I – III.

P2. Distinctions I – III apply to linguistics.⁹²

P1 and P2 belong to the original formulation of the argument but there are three further premises that Devitt is relying on in his argument.

P3. A theory of a language is a theory of the structure rules that govern outputs/products.

P4. The outputs/products are external, physical tokens (or properties thereof).

P5. From P3 and P4, A generative grammar is a theory of language dealing with the linguistic structure rules that are realised by external, physical outputs.

⁹⁰ Smith (unpublished ms.)

⁹¹ Presented in Collins (unpublished ms. a), Smith (unpublished ms.).

⁹² Devitt (2006) pp.23-30

From P5 Devitt draws the desired conclusion, C1. From P5 and his distinctions I - III, Devitt draws C2, his own position on the psychological reality of grammar.

C1. The linguistic conception is true and the psychological conception is false.

C2. A generative grammar is a theory of the nature of the external outputs that constitute language, not of the psychological reality of language in its competent speakers beyond adherence to (M).

This is where Devitt's original argument ends up⁹³, though the crucial P3 – P5 are now explicit. This brings out more clearly that everything turns on whether a generative grammar, as a theory of a language, is best interpreted as a theory of the rules that govern external outputs rather than the rules that govern grammatical competence. On this way of stating the argument, Devitt's key claim is that the grammatical properties which grammarians target are properties of external products. Devitt's argument now hangs on the nature of the grammatical reality that is being inquired into. He proposes a conception of language and grammatical reality that Chomsky rejects.

Devitt portrays Chomsky as thinking that grammar is an aspect of the human mind and offers the view that grammars are about language as a conflicting view. Devitt takes the latter view to exhibit a certain simple-minded quality and to be in tension with Chomsky's view. But there is no internal tension whatsoever in Chomsky holding both that grammars are about languages and an aspect of the human mind. There is no tension because Chomsky thinks that languages, in the relevant sense, are mental: they are I-languages, internal to the human mind/brain. So he can hold both that grammars are about language and about the mind, describing the linguistic structures licensed by our grammatical competence. On the psychological conception, a generative grammar does detail the linguistic structures of speakers' languages but these structures are cognitive. The question is whether

⁹³ Devitt (2006) pp.30-8

Devitt is right to think of the structures and properties appealed to in generative grammar as external to the human mind or whether Chomsky is right to think they are a part of the human mind/brain.

Devitt thinks that once his distinctions have been drawn and applied it is more plausible that linguistics is the study of language conceived as a non-psychological phenomenon. He says:

[The] actual and possible idealised outputs, governed by a system of rules and fitting into a structure are what we would normally call a language. Indeed, wherever there is a linguistic competence there has to be such a language, for the language is what the competence produces: the language is what the speaker is competent *in*; it is definitive of the nature of the competence.⁹⁴

On this basis, Devitt suggests that the psychological conception amounts to a blurring of the distinctions and requires special pleading in the form of auxiliary hypotheses or argumentation. We might suppose, for now, that Devitt is right on the issue of prima facie plausibility, since initial plausibility is not the important issue. It is unclear that any onus or burden of proof lies with Chomsky in the way Devitt suggests. Chomsky has presented a research program into core linguistic phenomena, guided by a certain theoretical conception of language. Even supposing Devitt is right about what we would *normally* call a language, that classification may not be the one that animates scientific research. Linguistics is an attempt to get some explanatory perspective on linguistic phenomena so it is not really to the point whether or not it matches up with our ordinary classifications. From Devitt's perspective, it is incumbent on advocates of the psychological conception to say more about how it is motivated and why it locates grammatical properties not in a physically-realised, extra-mental reality of produced noises and marks but in a psychological substratum that underlies those phenomena. But equally, it is incumbent on Devitt to say something about the fruits, potential or actual, of his non-psychological, physicalist conception of language; and how that conception meshes with the methodology and empirical results of generative grammar.

⁹⁴ Devitt (2006) p.31

If Devitt had simply said that generative grammar is about language, and said nothing about the nature of grammatical reality, he would then merely be challenging Chomsky to justify his claims about the psychological objectives of generative grammar, the role of psychological evidence in supporting psychological hypotheses, and his contention that generative grammars describe an aspect of human psychology. This would have highlighted the fact that to detail and defend Chomsky's psychological conception of grammatical reality is a substantial task. In contrasting the psychological conception with the view that generative grammars describe a linguistic reality, Devitt could then be understood as asking Chomsky to spell out the arguments and evidence in its favour. But without an alternative competing conception, this challenge might have seemed hollow. Hence, Devitt offers his own vision that *language*, the grammatical reality described in generative grammar, is physical and to be located outside the human mind. In doing so, he incurs just the same burden as there is on Chomsky to produce arguments and evidence for his view.

We might grant Devitt that, with the distinctions drawn, a non-psychological conception is the more, *prima facie*, plausible view. However, considerations of initial plausibility are clearly inconclusive for determining which sorts of properties or structures a scientific theory posits in achieving the explanatory goals it aims at. Devitt himself does not want to lean too hard on pre-theoretical ideas about language. As he remarks "Linguistics, like other sciences, largely determines its own domain": determining some explanatory goals and working out what sorts of theoretical constructs will be required to meet them.⁹⁵

Devitt's pro-psychological conception opponents say that, having selected a range of facts about linguistic form and meaning as their target phenomena, the linguist is involved in investigating the psychological structure that shapes the forms and meanings we find amongst streams of sound and marks. Devitt's conclusion is that his opponent is wrong on this point: linguistics is best interpreted as constructing non-psychological theories to explain non-psychological properties located in sounds and marks. Intuitively, Devitt thinks, linguistics is about languages but not psychological states.

⁹⁵ Devitt (2006) p.27

If we suppose with Devitt that his thought about prima facie plausibility is correct then, with the distinctions in place, proponents of the psychological conception would need to offer elucidation and arguments, theoretical or empirical, for why we should take the science of grammar to be an investigation of grammatical competence. Where might they look for such arguments? Devitt's view is that the explanatory successes of generative grammar do not tell in favour of the psychological conception widespread amongst linguists. He says of his critics that:

A theme...is that linguistics conceived psychologically is a piece of admirable "empirical science" whereas conceived linguistically it would be a piece of regrettable "metaphysics". This theme is spurious. The change in conception has no effect on the empirical nature of linguistics.⁹⁶

Devitt's suggestion is that the switch in conception is consistent with both the status of linguistics as an empirical science and the fruits of "the wonderfully successful research program in linguistics initiated and sustained by Chomsky".⁹⁷ His claim is not that linguists should give up linguistics as currently pursued and shift their attention to a new research program. He says: "Insofar as the research program is the one of constructing grammars in pretty much the way they are being constructed. I am all for it".⁹⁸ Rather his claim is that linguistics *as it has been pursued to date* is concerned with and animated by non-psychological grammatical properties.

To the extent that Devitt's objective is strictly interpretation of the science rather than revision of its methodology and results, I think he would hold himself to the following principle which I'll call the *No Violence Principle*.

No Violence: An account of what linguistics is about ought not to jeopardise its explanatory successes.

⁹⁶ Devitt (unpublished ms. a) p.15

⁹⁷ Devitt (2006) p.3

⁹⁸ Devitt (unpublished ms. b)

To see if non-psychological conceptions of generative grammar, of which Devitt's linguistic conception is one, meet this No Violence Principle and offer any further attractions beyond minimal consistency, we first need to know what non-psychological conceptions of language might be on offer. Devitt's suggestion is that generative grammars are best interpreted as describing properties of our physical environment, but I'd like to consider a wider range of options.

1.3.1 Social Conceptions

One such non-psychological conception of linguistic properties is a social conception. Broadly speaking, a social conception of linguistics would hold that linguistic properties are social properties: properties inherently of, or available to, humans living in groups, who interact and communicate with one another. The conception would be non-psychological insofar as social properties are not properties of our individual psychologies. I'll consider three ways of developing the view that languages are social phenomena which might support a conception of linguistics as concerned with social properties. The first is Quine's view of languages as socially acquired skills, whose properties are located in publicly observable behaviour. The second is McDowell's view that the properties of the languages we speak and understand should be understood in terms of our participation within communal practices. On both these views linguistic properties are manifest in social behaviour rather than properties of a psychological system like an I-language. The third is Wiggins' view of languages as historically determined, social objects.

Language and Behaviour

In 1959, Chomsky wrote a celebrated review of Skinner's *Verbal Behaviour* in which he critiqued Skinner's behaviourist view that human language could be explained in terms of conditioned responses to external stimuli.⁹⁹ Chomsky raised three main problems for the behaviourist approach. Chomsky argued that the

⁹⁹ Chomsky 1959

behaviourist's theoretical notions of stimulus and response lose all objectivity when we try to apply them to our verbal behaviour and I'll return to this issue below. He argued that the behaviourist framework fails to accommodate the freedom that is part of our use of language, the use of language generally being stimulus-free. He also pointed out that behaviourism seemed unable to accommodate the structural complexity of human languages. Despite the influence of Chomsky's critique, there are still those who think that an inquiry into linguistic facts, including grammatical facts, is an inquiry into social facts to be located in behaviour.

Quine famously claimed that "Language is a social art", arguing that the facts about languages are facts about socially observable behaviour.¹⁰⁰ Quine's social conception of language is based on his view about the evidence that is relevant to the study of language. Quine's view is that facts about a speaker's dispositions to overt behaviour are the only evidence relevant to a theory of their language. Quine had two arguments for this view, which I'll call *the acquisition argument* and *the manifestation argument*.

The acquisition argument runs as follows:

In acquiring [language] we have to depend entirely on intersubjectively available cues as to what to say and when. Hence, there is no justification for collating linguistic meanings, unless in terms of men's dispositions to respond overtly to socially available stimulations.¹⁰¹

Quine's view is that acquiring a language is a matter of picking up on cues available in social situations, cues that can be used to learn the form and significance of utterances. The major premise of the acquisition argument is that language acquisition is a matter of responding to socially available cues, such as an utterance of "shut the door" and the door-shutting behaviour that follows it, depending "entirely" on the observation of behaviour and the reinforcement of behaviours thereby learnt. This was a view that Quine maintained throughout his writings, claiming much later on that "Each of us learns his language by observing other people's behaviour and having his own faltering verbal behaviour observed and

¹⁰⁰ Quine (1960) p.ix

¹⁰¹ *ibid.*

reinforced or corrected by others.”¹⁰² For Quine then the facts about the languages we acquire are determined by socially available cues and by reinforcement and correction in social situations.

However, Chomsky rightly pointed out that it is an empirical assumption about language acquisition.¹⁰³ It is an *empirical* assumption, because it is not something we could know to be the case by mere reflection on the superficial aspects of acquiring a language. As I shall explain, Quine agreed that he was making an empirical assumption about language acquisition but thought his assumption rested on firm foundations.

The assumption that language acquisition is the development of conditioned responses to observable stimuli was an assumption that Chomsky had called into question in his review of Skinner:

It is not easy to find any basis (or, for that matter, to attach very much content) to the claim that reinforcing contingencies set up by the verbal community are the single factor responsible for maintaining the strength of verbal behaviour. The sources of the “strength” of this behaviour are almost a total mystery at present.¹⁰⁴

Chomsky was unconvinced by Quine’s account of language acquisition for similar reasons to those which left him unconvinced by Skinner’s. Quine’s account rests on the notion of *similarity*. The fundamental idea is that habits in our behaviour can be explained by our finding some stimuli more similar to one another than others and there being a correlative similarity in our responses. Accordingly, speakers of languages are conceived of as having a *similarity-space*: finding certain linguistic stimuli similar, so as to place them in this similarity-space which coordinates similar stimuli and similar responses. According to Chomsky, the behaviourist owes us an explanation of what special linguistic similarities consist in, how speaker-hearers are sensitive to the similarities and how they produce behaviour according to them. The behaviourist could offer the explanation that stimuli are linguistically similar when they provoke the same behavioural responses. But we then need to

¹⁰² Quine (1987) p.5

¹⁰³ Chomsky (1975a)

¹⁰⁴ Chomsky (1959)

know what characterises similarity of response in the linguistic case. And this needs to be specified without appealing to similarity of stimuli. If the behaviourist wants to define similarity of stimuli in terms of similarity of linguistic response then similarity of linguistic response will have to be defined in independent terms. Chomsky suggests that behaviourist theories, such as Skinner's and Quine's, tend to tacitly close this loop and lapse into vacuity.

To the extent that behaviourists avoid this problem, by giving the terms linguistic "stimulus" or "response" precise and independent meanings, Chomsky argued that they grossly oversimplify and misrepresent the linguistic phenomena. A pigeon's pecking or dog's salivating behaviours might be accommodated within the paradigm of stimulus and response. Animal psychologists train pigeons to peck on a pad or lever a certain number of times to receive food. The pigeon can vary the number of times that it pecks dependent on the number of pecks that will release the food. Pigeons have varying levels of success at such tasks depending on the size of the numbers and the accuracy required of them. The pigeon's pecking behaviour can be explained as a conditioned response. The behavioural response is conditioned by the number of pecks it takes to get the food, and the stimulus is the food. In such cases, the stimulus is exposure to a clearly specified variable, and we get clearly specifiable responses.

But it is less clear how we identify the "stimulus" for normal linguistic behaviours, such as commonplace utterances. As Chomsky argues:

A typical example of stimulus control for Skinner would be the response to a piece of music with the utterance Mozart or to a painting with the response Dutch. These responses are asserted to be "under the control of extremely subtle properties" of the physical object or event. Suppose instead of saying Dutch we had said Clashes with the wallpaper, I thought you liked abstract work, Never saw it before, Tilted, Hanging too low, Beautiful, Hideous, Remember our camping trip last summer?, or whatever else might come into our minds when looking at a picture (in Skinnerian translation, whatever other responses exist in sufficient strength). Skinner could only say that each of these responses is under the control of some other stimulus property of the physical object.¹⁰⁵

¹⁰⁵ *ibid.*

Further, the linguistic properties of verbal behaviour, such as the properties of reflexives, are more complex than the stimulus-response account is able to accommodate. There are no observable stimuli peculiarly germane to elicit linguistic responses structured at that fineness of grain; so it is unclear how to construct an informative description of the stimulus for my putting together a reflexive construction. Moreover, the preponderance of our verbal behaviour seems to be stimulus-free: a matter of choosing to express ourselves freely. If I were to utter “Shizzaam!” or “Zaapp!” or make some other spontaneous utterance, this seems to be a matter of my own free will. The behaviourist says that it is a conditioned response to an extremely subtle property of my environment that acts as a stimulus to determine my utterance.

Chomsky criticised Quine’s explanation of language acquisition in terms of the child placing linguistic stimuli (such as its parents’ utterances) on a pre-linguistic similarity space that measures properties of the sounds (such as length and pitch). He offered the following clarification of the view on Quine’s behalf. We can think of a similarity space as an inbuilt measure of qualitative distance, perhaps “restricted to dimensions with physical correlates such as brightness or loudness, and distance defined in terms of those physical correlates”.¹⁰⁶ Chomsky argues that we now have a “doctrine of innate spaces” which must face empirical testing along with other accounts of the organism’s contribution to language acquisition. By way of reply, Quine claimed the innate similarity spaces he appeals to must exist, as there can be no regularity or habit in our linguistic behaviour without dispositions to find stimuli similar. So, Quine claimed that it is the nature of these innate similarity spaces that must be experimentally determined but their existence is not in question.

Although Chomsky might agree with Quine that an innate similarity space is involved in language acquisition, he maintains that it is an empirical matter to determine the nature of such spaces and their development. Quine suggests that our sense of linguistic similarity is largely determined by the cumulative effect of linguistic stimuli on a space that maps physical magnitudes. But there are competing explanations of our sense of linguistic similarity, adopting different views about the child’s innate endowment. These competing explanations postulate

¹⁰⁶ Chomsky (1975a)

a rich, language-specific body of information that supports language acquisition. If Quine's view is simply that there is *some variety* of information pertaining to similarity deployed in language acquisition, *then* the hypothesis would not have to face the evidence or compete with alternative hypotheses; but only because it is trivial and so does not support any particular account of language acquisition. Only details and evidence could suffice to allay Chomsky's concern that Quine assumes an empirically vacuous notion of similarity space.

A further problem that Chomsky raises concerns the role that the notion of *analogical synthesis* plays within Quine's account. Analogical synthesis is a hypothesised process involving abstraction from sentences thus far encountered and the reassembly of the parts of those sentences to form new sentences. Consider the following example of analogical synthesis. A child hears the utterance "Simon is tall" tied to observations of Simon. The child abstracts the parts "Simon", "is" and "tall" from the sentence. The child then hears the utterance "James is short" tied to observation of James and abstracts the parts "James", "is" and "short". The child then synthesises these parts to form the new utterance "Simon is short" on analogy with the previously uttered sentences. Quine used the notion of analogical synthesis to explain how we can understand sentences that are not directly tied to observable stimuli. We first learn sentences that are directly tied to observables. Then we abstract the contributions of similar parts to those sentences. We then recombine those parts to form new sentences in the absence of directly observable stimuli. Quine assumes that the child can extract the words from the sounds by some kind of abstract categorisation. He also assumes that the child can recombine those words grammatically, i.e. the child doesn't utter "Simon James short". Quine owes us an explanation of how each of these steps is achieved.

Chomsky claims that the account is "empty until an account of 'analogy' is given, and none exists".¹⁰⁷ Equally, the claim that the child effects 'synthesis' labels a mystery until a basis for the abstraction and assembly is specified and empirically tested. Without a thorough investigation of the psychology of the child, a vast and uncomprehended contribution is attributed to steps labelled *analogy* and *synthesis*. According to Chomsky, it is an empirical problem, unaddressed by Quine's account,

¹⁰⁷ Chomsky (1975a) p.64

to assess what properties of the mind determine the nature of the language we acquire.

Chomsky highlights the magnitude of the empirical problem of language acquisition in his review of Skinner, when he writes:

Study of the actual observed ability of a speaker to distinguish sentences from nonsentences, detect ambiguities, etc., apparently forces us to the conclusion that this grammar is of an extremely complex and abstract character, and that the young child has succeeded in carrying out what from the formal point of view, at least, seems to be a remarkable type of theory construction. Furthermore, this task is accomplished in an astonishingly short time, to a large extent independently of intelligence, and in a comparable way by all children. Any theory of learning must cope with these facts.¹⁰⁸

On the basis of these facts, Chomsky began to think it very implausible that in acquiring a language we are learning a social art, for: “[W]hen we learn a language we are not ‘learning sentences’ or acquiring a ‘behavioural repertoire’ through training. Rather, we somehow develop certain principles (unconscious of course) that determine the form and meaning of indefinitely many sentences.”¹⁰⁹ Quine’s reply to the suggestion that language acquisition is a form of development involving heavy constraints from innate linguistic principles was to distinguish between our aptitude for language and our acquisition of language. He says:

Language aptitude is innate; language learning, on the other hand, in which that aptitude is put to work turns on intersubjectively observable features of human behaviour and its environing circumstances, there being no innate language and no telepathy.¹¹⁰

Since Quine is claiming that language acquisition is not a matter of innate linguistic constraint and there are no innate linguistic principles, this would appear to Chomsky to be merely a restatement of unjustified empirical assumptions, to which he accuses Quine of adhering.¹¹¹ According to Chomsky, whether language

¹⁰⁸ Chomsky (1959)

¹⁰⁹ Chomsky (1975a) p.64

¹¹⁰ Quine (1969) p.306

¹¹¹ See Chomsky (1975a)

acquisition is principally a matter of observing behaviour and circumstances is a matter for empirical investigation. Questioning Quine's assumptions, Chomsky claimed that no evidence had ever been presented for the view that children acquire their language by learning from verbal behaviour with the aid of correction.¹¹² Chomsky, along with many other linguists, has come to strongly favour a large role for the "innate language" that Quine dismisses. As work in generative grammar flourished, it revealed the astonishing complexity of the languages humans acquire. But it also revealed the depth and uniformity of their underlying principles. It became clear to Chomsky that:

We must deal with the crucial...fact that what a person knows [in knowing a grammar] is vastly underdetermined by available evidence, and that much of his knowledge is based on no direct evidence at all... There is little doubt that this problem of "poverty of stimulus" is in fact the norm rather than the exception.¹¹³

Following Chomsky, I argue in Chapter Five that Quine's assumption about language acquisition is a false empirical assumption and that there is a preponderant role for innate, language-specific principles that tightly constrain the course of grammar acquisition.

But suppose the chips fell the other way and Quine's view that language acquisition is learning from external stimuli with reinforcement and correction is correct. What follows? The second, suppressed, premise of the acquisition argument is that linguists can draw upon no more evidence than is available to the child in language acquisition. Quine recognised that he had provided the linguist with a "meagre basis" but he maintained that "the native speaker has had no other".¹¹⁴ He restricts the linguist to behavioural evidence because he thinks language acquisition takes place on the basis of that evidence alone. Chomsky takes issue with this claim about what is required of inquiry into speakers' verbal behaviour:

¹¹² Chomsky (1975a) p.53

¹¹³ Chomsky (1990) p.593

¹¹⁴ Quine (1987) p.5

Prediction of the behaviour of a complex organism would require, in addition to information about external stimulation, knowledge of the internal structure, the ways in which it processes information and organizes its own behaviour.¹¹⁵

Even if Quine had been right about language acquisition and the materials available to the native speaker, nothing follows about the evidence that *the linguist* might appeal to. The a priori restriction of the linguist's data to that which he shares with the child is unjustified, in Chapter Four we'll see the implausibility of the restriction in light of actual, scientific practice. Suffice to say for now that linguists, just like scientists in any other field, make use of whatever evidence is pertinent and available to them. There is no route from the nature of language acquisition to Quine's conclusion that the only evidence relevant to the study of language is socially observable behaviour.

The second of Quine's arguments for his behaviourist view of the evidence relevant to studying language turns on what is *manifest* in episodes of linguistic understanding. There is a point of contact here with the work of Michael Dummett and Crispin Wright on linguistic meaning.¹¹⁶ The shared thought is that there can be nothing more to the linguistic meanings grasped in episodes of linguistic understanding than what is manifest in linguistic behaviour. What a speaker is taken to mean by those that understand him is not generally informed by their knowledge of the speaker's internal organisation or other unobservable facts. Rather the speaker's meaning is grasped on the basis of what can be gleaned from his overt behaviour. Quine claims "There is nothing more to linguistic meaning than what is gleaned from overt behaviour in observable circumstances."¹¹⁷ And he takes this to support a behaviourist restriction on the linguist's evidence.

It does not, however, follow from the fact that we can typically grasp what is said by fellow speakers of our language on the basis of their verbal behaviour that there is *nothing more* to our understanding one another's meanings than our observation of overt behaviour. The fact that we typically grasp someone's meaning on the basis of their verbal behaviour is perfectly compatible with, and arguably it

¹¹⁵ Chomsky (1959) p.27

¹¹⁶ Wright (1993) pp.13-26

¹¹⁷ Quine (1987) p.5

presupposes, psychological determinants of meaning and understanding. A speaker's psychological systems shape his linguistic behaviour in special ways such that observers with sufficiently similar psychological systems recognise distinctive configurations in the noises he emits. Linguistic understanding is achieved by attending to verbal behaviour, but there are more determinants to the meanings expressed and the understandings achieved than the overt behaviour. Chomsky is right to warn us off the idea that properties of the mind amenable to empirical investigation "impose no conditions on language".¹¹⁸ We do typically grasp meanings by observing one another's behaviour, but the reason we hear the structured meanings in verbal behaviour is that we are appropriately configured to be sensitive to such properties. The nature of these sensitivities does not lie open to view in ordinary exchanges but is not thereby irrelevant to the linguist in theorising about the properties of the linguistic structures we produce and comprehend.

Behaviourists, like Quine, think that aspects of a speaker's mental organisation that have no reflection in our actual or potential dispositions to verbal behaviour are not pertinent to linguistic inquiry, so the linguist should appeal exclusively to evidence drawn from behaviour. But proponents of the psychological conception, like Chomsky, can agree that aspects of our mental organisation which have no reflection in our actual or potential dispositions to verbal behaviour are not the focus of linguistic inquiry. Chomsky may agree with this whilst maintaining that the real explanation of linguistic phenomena lies not in verbal behaviour but in cognitive states that shape that behaviour, of which the I-language is one. It is platitudinous that aspects of the speaker irrelevant to the kind of language they speak and understand will be irrelevant to the study of language. Drawing on remarks of the later Wittgenstein's, such as the famous "an inner process stands in need of outward criteria"¹¹⁹, philosophers like Dummett and Wright have been attracted to the thought that the inner workings of the mind do not explain the nature of meaning and understanding. But the internal structure that Chomsky

¹¹⁸ Chomsky (1975a) p.65

¹¹⁹ Wittgenstein (2001) §580

postulates to explain linguistic behaviour is not what Wittgenstein described as a “wheel that can be turned though nothing else moves with it”.¹²⁰

The cognitive states in question are not inner mechanisms, or “wheels”, that turn without shaping our verbal behaviours. Such idle wheels are irrelevant to behaviourist and Chomskian alike. Quine’s manifestation argument is no justification for his behaviourist scruples because it provides no reason to think that cognitive organisation is irrelevant to linguistic meaning and understanding. The verbal behaviour, to which Quine restricts the linguist, receives its explanation from the cognitive facts, such as the structure of our I-language, which he legislates against. The argument that linguists should consider only behavioural evidence because we, as cognitively equipped speaker-hearers, can understand our language without consulting further evidence is a non-sequitur.

There is another strand to Quine’s manifestation argument which concerns the basis on which one is counted as being a competent language user. Quine claims that we are counted as competent language users on the basis of our behaviour and not on the basis of our internal structure:

As long as our command of language fits all external checkpoints, where our utterances or our reaction to someone’s utterances can be appraised in light of some shared situation, so long all is well. Our mental life between checkpoints is indifferent as to our rating as a master of the language.¹²¹

Quine is very likely correct, at least as far as commonsensical classifications go, that our mental organisation is not the criterion by which we are usually counted or discounted as being a competent language user – or as Quine puts it “a master of the language”. We are normally counted a master of a language on the basis of our proficient use of a language. So, generally it is verbal behaviour and not mental organisation that indicates our understanding a language to fellow language users. But as Ryle notes, our ordinary conception of someone being competent in a language is vague and open-ended. Competence in French does not, for example, require that someone be competent to translate highly technical French treatises

¹²⁰ Wittgenstein (2001) §271

¹²¹ Quine (1987) p.5

though it can do in special circumstances. Usually, of someone competent in French “We expect no more than that he will ordinarily cope pretty well with the majority of ordinary French-using and French-following tasks. ‘Knows French’ is a vague expression and, for most purposes, none the less useful for being vague.”¹²²

Whatever the purposes or merits of attributing mastery of a language on this basis, there should be no a priori constraint on the science of language to count behaviourally similar language users as possessing the same language or equivalent grammatical competences. What requires explanation by the science of language is not our (commonsensical) criteria for attributing mastery but rather the properties of languages which speakers know – properties that equip speakers to exhibit such behaviours as fit, as Quine puts it, “all external checkpoints”.

On the psychological conception, the point of such attributions of mastery is tangential to the scientific conception of possessing a language, and language possession is not indifferent to internal organisation. A speaker’s possessing an I-language – his possession of a *competence* in the technical sense described in Chapter Three - *is* his possession of a special cognitive system. A speaker could retain this I-language whilst losing all his expressive powers and abilities to use his language, hence failing to meet “external checkpoints” and falling short of the ordinary criterion. The possession of an I-language might not be our everyday sense of linguistic competence, but that is no reason to think it is not the sense relevant in linguistics.

With the linguist’s evidence restricted to behaviour, Quine found that in ascribing linguistic forms and meanings to a speaker’s words the linguist always goes beyond the evidence so as to “impute our sense of linguistic analogy unverifiably to the native mind”.¹²³ The attributions of linguistic form and meaning are “unverifiable” because, Quine claimed, there will always be at least two alternative attributions of form and meaning to the speaker’s words consistent with the totality of behavioural evidence but mutually inconsistent. This was Quine’s thesis of the *indeterminacy of translation*: for any speaker of a language there are always at least two alternative translation manuals for the language consistent with

¹²² Ryle (1949) p.119

¹²³ Quine (1960) p.72

the totality of behavioural evidence but inconsistent with one another. Given Quine's view that there is no further non-behavioural evidence that could decide between the theories, there is no objective fact about the forms and meanings of the speaker's language. The thesis was thought by Quine to place severe restrictions on linguistics, limiting it to an investigation of behavioural surrogates for linguistic notions like meaning, and linguistic structure.

Chomsky is unmoved by Quine's indeterminacy thesis. He claims it is no more than the general and familiar underdetermination of empirical theory by evidence in the particular case of language.¹²⁴ According to underdetermination, there are always at least two theories in any given domain compatible with the available evidence. Hence, our choice of theory will be underdetermined by the evidence. There will always be alternative theories of the speaker's language inconsistent with one another but consistent with our evidence; such is empirical theory in Chomsky's view. The indeterminacy thesis holds that there is not merely underdetermination but no ultimate fact of the matter about language. Indeterminacy about language only follows from the underdetermination of linguistic theory by behavioural evidence if one assumes that there is no further fact of the matter about linguistic properties beyond the facts of actual and potential overt behaviour. Without the behaviourist scruples, there is no justification for the thesis. In reality, linguists look for other evidence about the properties of languages that might decide between competing theories. Linguistic theories progress by gathering more and varied evidence to develop competing explanations of linguistic phenomena. Chomsky suggests that linguists draw upon their own introspective judgements and those of native speakers. He also suggests that linguists appeal to theories of the native's mind and, in principle, brain to determine his language: these are empirical hypotheses about the languages humans can acquire and the rich internal organisation that is engaged in linguistic behaviour.

Quine's scepticism about linguistic form and meaning does not have many contemporary advocates. But the background conception according to which linguistic facts are social facts that must be open to view in social behaviour draws a lot of support in broadly Wittgensteinian philosophical quarters amongst those

¹²⁴ Chomsky (1975, 1980)

who think that language ought to be understood in terms of communal practices and publicly observed norms. John McDowell has been a highly influential source of ideas in this direction. He seeks to understand language in terms of linguistic practices and norms.

Language and Practice

McDowell argues that if we understand what is manifest in linguistic behaviour in terms of “bare behaviour”, physically characterised as Quine assumes, then linguistic understanding becomes impossible.¹²⁵ McDowell’s argument draws on the idea that linguistic meaning and understanding are *normative* phenomena: they depend upon the application of standards of correctness to episodes of verbal behaviour. For example, McDowell thinks that for a speaker to mean RED by an utterance “red” and for a hearer to understand RED by that utterance requires a common commitment to a standard of correctness for the application of the word “red”. On McDowell’s account this standard of correctness is determined by the two conversants’ participation in colour related linguistic practices.

These standards of correctness are required to secure a distinction between the correct application of linguistic expressions and what *merely seems* correct to participants in linguistic activities. McDowell thinks that a grasp of such a distinction is required for one to mean or understand something *determinate* by a linguistic expression for the following reason. The lack of a distinction between what is and what merely seems to be a correct use of an expression compromises any particular content being expressed by its use. For if anything counts as the correct use of an expression then an expression can mean anything at all. If whatever seems to be a correct use is correct, then my utterance of an expression would fail to mark out any distinctive content, the expression in question could apply to anything and any way of taking the expression would serve to have understood it.

McDowell argues that if we accept, with Quine, that verbal behaviour described in purely physical terms is all that is available to speaker-hearers in

¹²⁵ See in particular McDowell (1984) §10.

communication then we would never be able to reconstruct the normative dimensions of meaning and understanding. For Quine, behavioural dispositions are all there is to language; behavioural dispositions have replaced meanings. McDowell denies that behavioural dispositions should replace linguistic meanings. He believes that linguistic meanings exist but that we cannot ascertain such meanings on the basis of speakers' dispositions to verbal behaviour because meaning and understanding involve commitment to determinate standards. Patterns of physical behaviour can inductively confirm the attribution of behavioural dispositions. If a speaker utters the term "red" in the presence of certain physical objects we can inductively infer on this basis that he has a behavioural disposition to so apply the term. But, McDowell asks: "[H]ow can we extrapolate to a determinate conception of what the disposition is a disposition to do?"¹²⁶ In other words, how does the physical behaviour make manifest a commitment to a determinate pattern of use?

According to McDowell, such physical behaviour does not make manifest a commitment to the determinate patterns of use that are required of meaning and understanding. Behaviourism, McDowell argues, cannot explain speakers grasping a determinate pattern that extends indefinitely beyond finite instances of behaviour. So it cannot explain there being correct and incorrect ways of apprehending speakers' meanings. McDowell reasons as follows. If understanding is possible, then meaning must be manifest to speaker-hearers, but no determinate pattern is manifest in physically characterised behaviour – so *more than* physical behaviour must be manifest to speakers in episodes of understanding. Hence, a behaviourist conception of language is deficient as it cannot account for meaning and understanding.

What McDowell advocates is not that we give up on the requirement that linguistic properties be manifest in linguistic behaviour and adopt a psychological conception of language. McDowell, rather, suggests that we characterise the relevant linguistic behaviour in a different way to Quine. To make room for understanding, linguistic behaviour must be characterised in meaningful terms rather than physical terms; such that linguistic properties are there to be perceived

¹²⁶ McDowell (1984) p.68

on the surface of behaviour along with its other perceptible properties. These perceptible properties are not available to just any old observer. Our ability to speak a language and to perceive linguistic properties in the verbal behaviour of others is, on McDowell's conception, grounded in our participation in a *linguistic practice*. According to this view, speakers can traffic in the linguistic properties of their language because they have been inculcated into a practice, a practice shared with other members of a linguistic community. Without the relevant *training* in a set of norm-governed behaviours and inculcation into a group practice, observers are blind to these properties and lack linguistic understanding. For McDowell the central issue about language is how "drilling in a behavioural repertoire" can enable one to perceive linguistic properties of which "one would not otherwise be aware."¹²⁷

Hence, though he has a very different view of the nature of linguistic properties, McDowell agrees with Quine that language is to be understood as a social art rather than in terms of our possession of a special psychological faculty.

McDowell offers us an account of normative aspects of meaning and understanding. Whatever the merits of McDowell's conception of language it is unlikely that grammatical properties fall within its scope. If the grammatical properties of speakers' languages were amongst those that become available to speakers in virtue of their inculcation into communal, normative practices, then we should expect these properties to vary with the social behaviours that are involved in training and in the achieved communal practices. But empirical investigation of grammar acquisition suggests, firstly, that grammatical properties do not vary in this way, and secondly, that training plays a very limited role, if any, in our acquiring a grammar (see my §5.4).

Sapir endorsed the view that grammatical properties vary from community to community. He said:

Speech is a human activity that varies without assignable limit as we pass from social group to social group, because it is a purely historical heritage of the group, the product of long-

¹²⁷ McDowell (1998a) p.333 See Smith (2006) for discussion of McDowell's account and its limitations.

continued social usage. It varies as all creative effort varies – not as consciously, perhaps, but nonetheless as truly as do the religions, the beliefs, the customs, and the arts of different peoples. Walking is an organic, an instinctive function...speech is a non-instinctive, acquired, ‘cultural’ function.¹²⁸

But it is false that the grammatical properties of speech vary “without assignable limit as we pass from social group to social group”. And there are no known explanations of grammatical principles, such as binding principles, in terms of “historical heritage”. Unless, that is, we pervert the meaning of “historical heritage” to include Universal Grammar (UG): the rich grammatical principles which assign unobservable structures to speech in the absence of “creative effort”. Katz rightly points out, that this underlying grammatical uniformity greatly reduces “the significance for the nature of language of the surface aspects of sentences which most reflect cultural and historical accretions.”¹²⁹ By “surface features” of sentences Katz means more superficial features affected by culture, like stylistic features.

While grammatical properties do vary across human languages, the variance is within strict limits. They do not vary as “all creative effort varies”. The postulation of training and adherence to community norms in our acquisition of language, central to McDowell’s conception, does not fit the facts about our acquisition of *grammar*. Grammatical principles, like those of binding theory, are not practical accomplishments: perceptual skills that children develop as they find their way in the social world. I argue (§5.4) that explicit instruction and implicit guidance play a very limited role in their acquisition. We don’t observe the variation in the structural properties of language that one would expect if the grammatical knowledge were “the product of long-continued social usage...a non-instinctive, acquired, ‘cultural function’”.

So whatever the strengths of McDowell’s conception of language it is important to recognise that it is limited when we consider grammar. Collins expresses the point at which we have arrived in very clear terms:

¹²⁸ Sapir (1921) p.4

¹²⁹ Katz (1981) p.8

Why do competent speakers recognise that himself in Bob expects to wash himself is Bob, while in Bill wonders who Bob expects to wash himself, the reflexive is neither Bill nor Bob? What convention might govern this? How might one be induced into this practice? How did you recognise that the stated relations hold, even though the sentences were new to you? Such questions arise with every construction and philosophy has not been forthcoming with answers. The moral here is not that the study of language should just amount to the investigation of the cognitive structure behind reflexives, wh-movement, etc. but that an account of language that does not factor in such structure will be demonstrably inadequate.¹³⁰

We can allow that there are normative aspects to the public use of language, whilst insisting that this norm-based conception falls well short of accounting for the grammar that we know. Knowledge of a grammar is required by individuals to successfully structure the expressions which the community uses. Smith makes the sensible point that whatever our view of public language use, we need to know how individuals combine the words of their language into grammatical structures.¹³¹ A much better explanation of the grammatical properties that speakers recognise than training, involvement in a practice or the recognition of community norms, is his knowledge of an underlying grammar with recursive procedures for building a discrete infinity of structures out of finite parts.

Language and History

David Wiggins attacks the psychological conception of language and defends a social conception of language different to that of Quine or McDowell. His view is that linguistic properties are properties of *social objects*. Wiggins thinks that languages like *English* and *Polish* are such objects. These are public objects, whose properties are irreducible to the psychologies or behaviours, of speakers. Wiggins thinks we can identify these social objects as follows:

English is a language arising under the influence of Norman French from the West Germanic language, Anglo Saxon; it is the language possessed of many forms and dialects

¹³⁰ Collins (unpublished ms.) pp.3-4

¹³¹ Smith (2006) p.944

spoken in the British Isles, North America, etc; it is the language in which you can say ‘Fear no more the heat of the sun’... If you disjoin enough true statements, each having the form ‘at time t sentence s means in English that p’, then you will easily manufacture a complex sentence that you or I could not possibly deny of English without putting in jeopardy our claim to know what the English language is.¹³²

Wiggins finds little value in the psychological conception. He does not find it “at all credible” that in virtue of our differing I-languages we may not share a language – namely, *English* – with Dr. Johnson “the first lexicographer of that very language”.¹³³ We can make sense of there being one and the same language we share with Dr. Johnson if we think of that language as a malleable, social object.

Wiggins thinks that it is via these social objects that we express our meanings in speech:

A language is an instrument by which speakers may frame their thoughts, something not concrete, yet ‘out there’ and wide open to be encountered in the human world.¹³⁴

These social objects are enduring, having “an origin, a past, a present and a future”.¹³⁵ But, interestingly, on Wiggins account they are not essentially defined by sets of lexical items and grammatical and semantic rules. So the scope of Wiggins’ attack actually extends beyond psychological conceptions of language, to all conceptions that characterise languages by constructing a grammar or semantic theory. On Wiggins’ view no specific grammar that generates the sentences of a language is an essential property of that language.¹³⁶ Rather Wiggins’ *languages* are “historically given” objects “changing or changeable and possessed of sentences that have, as a matter of history, such and such meaning at this, that or other time

¹³² Wiggins (1997) p.500

¹³³ *ibid.*

¹³⁴ Wiggins (1997) p.499. Wiggins owes us an explanation of how we make use of this independent, “instrument” in “framing our thoughts” in language.

¹³⁵ *ibid.*

¹³⁶ Wiggins’ claim is that no specific grammar is an essential property of a language, but it may be that it is essential that at any point in time a language has a grammar.

and place.”¹³⁷ Although Wiggins presents his conception of a language as a piece of refined commonsense and not a theoretical construct, it is his intention that his social object conception be considered as in competition with the psychological conception.

But it is not clear that there is friction between Wiggins’ conception and the psychological conception. As HCF are happy to concede:

The word “language” has highly divergent meanings in different contexts and disciplines. In informal usage, a language is understood as a culturally specific communication system (English, Navajo, etc.) [But] in the varieties of modern linguistics that concern us here, the term “language” is used quite differently to refer to an internal component of the mind/brain (sometimes called “internal language” or “I-language”).¹³⁸

HCF point out that different investigations of language adopt restricted senses of “language” suggesting different demarcations. And they think this is a good thing, with different conceptions of “language” to be judged on their fruits.

The dispute over language acquisition and our knowledge of grammar, on which Quine and McDowell’s conceptions come into conflict with Chomsky’s, does not surface on Wiggins’ conception. Most adherents to the psychological conception *are* committed to a view of grammatical properties that conflicts with them being properties sentences have as a matter of history. But Wiggins is not conceiving of languages in terms of these grammatical properties that individual speakers recognise. On his view there can be different grammatical properties recognised by speakers of the very same language at different points in time because he does not conceive languages as essentially sets of grammatical properties. Moreover, proponents of the psychological conception are not committed to denying the existence of any of the special properties of languages to which Wiggins appeals, or even their determination by historical fact. So the disagreement is not over the alleged existence of the historical and cultural properties that Wiggins says that languages have.

¹³⁷ *ibid.*

¹³⁸ HCF (2002) p.1570

One might speculate that Wiggins finds something to disagree with in the methodology behind the psychological conception, which aims to sharply distinguish the grammatical properties of languages from their historical and cultural properties. However, suppose we do separate out the phenomena along the lines grammarians suggest. It is then consistent with that fractionation of the linguistic phenomena that there are other properties of languages beyond their grammatical properties. These properties might well include aesthetic and normative properties which Wiggins mentions. If these properties are susceptible to systematic explanation, then perhaps their explanation will appeal to such social objects as Wiggins calls *languages*. Generative grammar would then be that part of a package for explaining “language”, exploring a range of core properties amongst “the past and present attributes of languages”.

The phenomena that interest Wiggins concern language as a historical accretion with “achievements and latent resources...something that influences normatively, by its palpable presence in the social world, the linguistic strivings of children, adults, foreigners, poets, writers, politicians and the rest.”¹³⁹ Wiggins takes it that proponents of the psychological conception, indeed all those generative grammarians who consider languages as grammatical systems, disagree with him. But there is no disagreement because they are not committed to denying the phenomena he highlights. The disagreement then must be over methodology.

To explain, Wiggins finds it “strange” to exclude phenomena including “culture, mores and social institutions” from the theoretical picture, and focus instead on the grammatical properties of individual speaker’s languages (properties of I-language on the psychological conception). But linguists think you have to divide and conquer the linguistic phenomena. For an all inclusive conception of language would be so overwhelmingly complex that the demands placed on delivering explanatory theory would be impossible to meet. Generative grammar is a particularly fruitful avenue of inquiry because the phenomena do look susceptible to some isolation. Here, the scientific linguist can effect some separation amongst the motley of contributing factors that we pre-theoretically class as *language*. An aim of generative grammar is further fractionation amongst the phenomena

¹³⁹ Wiggins (1997) p.500-1

intuitively classed as grammatical, not reintegration with little understood, perhaps scientifically intractable, social and normative forces. In my view, Wiggins is confusing the methodological limitations of generative grammar for philosophical shortcomings.

1.3.2 Physicalist Conceptions

On a physicalist conception of language, grammatical properties are properties located in our physical environment. On this conception, linguistics is the study of such properties, external to the mind. Though the physicalist and social conceptions are not necessarily mutually exclusive – for example, Quine’s conception could be classified as falling into both categories – there are cases where the distinction is clear. For example, McDowell denies that linguistic properties are physical, holding that they are normative properties of social practices. One can hold a physicalist conception without any commitment to the social character of linguistic properties.

Devitt’s linguistic conception is a species of physicalism. For on the linguistic conception, grammatical reality is a non-psychological reality consisting of *expressions* understood to be “physical entities forming representational systems” with phonetic, grammatical and semantic properties.¹⁴⁰ Devitt’s explanation of his technical term “expressions” is that they are *physical* entities, like acoustical signals and written marks, and this is intended to contrast with Chomsky’s use of the term “expression”, where Chomsky intends a *mental* representation pairing information on sound and linguistic form. Devitt says the grammatical properties of expressions are “high-level relational properties” of the physical entities.¹⁴¹

Devitt describes grammatical properties like c-command as “high-level” because they have multiple physical realisations: “Although formally so different, a written and spoken token...might share all their syntactic properties. Sentences that

¹⁴⁰ Devitt (2006) p.27

¹⁴¹ *ibid.*

‘look different’ might be syntactically alike.”¹⁴² C-command relations can be realised by pretty much any physical medium, as is apparent if one considers all the conceivable orthographies, acoustical signals, hand gestures and - as I’ll discuss in Chapter Five - the apparent amodality of language. Devitt thinks that grammatical properties are “relational” properties of expressions because properties like c-command are properties that hold between constituents, or “symbols” in Devitt’s terminology, in virtue of relations amongst elements in a structure: “Syntactic properties are ones that reflect a token’s relations to other tokens in the language”.¹⁴³

Devitt’s conception of grammatical properties is similar in some respects to views prominent immediately prior to Chomsky’s conceptual innovations. It is worth considering the views of Leonard Bloomfield and Chomsky’s teacher Zellig Harris as they provide an important background to both Chomsky’s psychological conception and Devitt’s revival of physicalism.¹⁴⁴ On their view, as on Devitt’s, grammatical properties are properties of the physical sounds and marks we find in our environment, properties that the physical entities have in virtue of certain relations they stand in to one another.

Bloomfield was committed to a view of scientific theory according to which any scientifically meaningful statement reports a physical movement in space and time. This was closely related to a view about scientific method, called *operationalism*, according to which all science must be framed in terms of statements that describe basic operations, recorded in reports about measurements of time and distance. Bloomfield took it that these doctrines had been established by the philosophers of the Vienna Circle such as Rudolf Carnap and Otto Neurath.

From these views, Bloomfield inferred that statements about linguistic ‘ideas’ – where his notion of ‘idea’ covered mental organisation and content – should be translated into statements about *speech-forms*. Bloomfield’s speech-forms were physical categorisations of acoustical disturbances. This programme in

¹⁴² Devitt (2006) p.154

¹⁴³ Devitt (2006) p.154

¹⁴⁴ Bloomfield (1936), Harris (1985)

linguistics was part of a larger vision of science and, in particular, the human sciences:

Linguistics as actually practiced employs only such terms as are translatable into the language of physical and biological science; in this linguistics differs from nearly all other discussion of human affairs. Within the next generation mankind will learn that only such terms are usable in any science. The terminology in which at present we try to speak of human affairs – the terminology of ‘consciousness’, ‘mind’, ‘perception’, ‘ideas’, and so on – in sum, the terminology of mentalism and animism – will be replaced in minor part by physiological terms and in major part by terms of linguistics.¹⁴⁵

Bloomfield believed that this course of action was not only desirable but also possible. He believed that Carnap and Neurath had already laid the foundations for this project by discovering that all scientifically meaningful statements could be translated into physicalist statements about observable movements and space-time co-ordinates. This estimation of what Carnap and Neurath had shown will seem too optimistic to those familiar with Quine’s attack on reductionism in epistemology and the philosophy of science, and Carnap’s own later moderation of these views.¹⁴⁶ But as Bloomfield saw it, statements in linguistics or any other special science that could not be so reduced were either scientifically meaningless or were statements about the language employed in the theory.

Bloomfield’s statements about the language employed in theories included statements like “In the English language the word *redness* is a noun” and the explication of these statements would draw upon the resources of linguistics. According to Bloomfield, the term “noun” must be defined as must wordhood in general. Bloomfield said of nouns that:

A noun is a word which enters centrally into endocentric phrases with preceding adjective modifiers, serves as an actor with a finite verb, as the goal of a verb or preposition, and as a predicate complement, appears always in one of two sub-classes, singular and plural.¹⁴⁷

¹⁴⁵ Bloomfield (1985) p.19

¹⁴⁶ Quine (1953, 1969)

¹⁴⁷ Bloomfield (1985) p.20

For Bloomfield a word was the smallest linguistically significant unit that could be uttered alone.¹⁴⁸ According to Bloomfield's strategy, wordhood had to be characterised in terms of the fundamental notions and definitions of linguistics. These fundamental linguistic terms were thought to admit, ultimately, of reduction to physicalist vocabulary. They were not taken to involve any appeal to meanings or "metaphysical" notions.

Bloomfield foresaw that his physicalist hypothesis about such linguistic categories would be tested by the generation of linguists that followed him:

If the description so revised [in a physicalist vocabulary rid of 'ideas'] is better than the old – simpler and fruitful of sounder and easier prediction – then the hypothesis will have been confirmed and mankind will accept it as we accept the Copernican astronomy.¹⁴⁹

But he was sanguine about the prospects of his hypothesis and warned against linguists who, unless they are physicalists, "constantly forget that a speaker is making noise, and credit him, instead with the possession of impalpable 'ideas'".¹⁵⁰ Bloomfield felt that linguists could play an important role in science, and in particular psychology, by showing that speakers do not have 'ideas' guiding their linguistic performance; that the noise is sufficient for linguists' purposes. He argued that the noise was sufficient to account for communication because "the speaker's words act with a trigger-effect upon the nervous systems of his speech-fellows."¹⁵¹ But he did not consider the study of the nervous system receptive to such triggers, and the mental structures it supports, to be a part of the explanation of language. The focal point was the noise itself.

Like Bloomfield, Harris thought that linguistic theorising should begin with attention to speech sounds. Both supposed that linguists must start with a set of observed utterance as their data. These utterances form a stream of speech on which linguists then perform basic operations: they can *segment* an utterance into

¹⁴⁸ Quine (1985) argues that the inclusion of linguistic *significance* in the definition of wordhood makes it hard to see how linguistics will be purged of ideas and meanings.

¹⁴⁹ Bloomfield (1985) p.23

¹⁵⁰ *ibid.*

¹⁵¹ *ibid.*

individual sounds, or *phones*, and they can *classify* phones into *phonemes*. Phonemes are the smallest units of sound in a language that are capable of conveying a distinction in significance, such as the *b* sound in *bat* and the *m* sound in *mat*. These operations can then be repeated at higher levels of analysis. Strings of phonemes can be segmented and classified into *morphophonemes*, a level of classification in between phonemes and morphemes. Morphophonemes are made up of the *allomorphs* – the phonetic variants - of a single *morpheme*. A morpheme is the smallest grammatically significant unit of the language, constituting a word or meaningful part of a word, that cannot be split into a smaller grammatically significant unit. So, for example, the phonetic (s) of *cats* pronounced (kats), the phonetic (z) of *pigs* pronounced (pigz) and the phonetic (iz) of *horses* pronounced (horsiz) are all allomorphs of the English plural morpheme.

Harris suggested that investigation of fragments of natural language reveals it to have a *distributional structure*.¹⁵² Harris used the term *structure* such that a set of linguistic elements *e* is structured with respect to some feature *f* to the extent that we can form an organised system of statements describing *e* and the interrelations of its members in terms of *f*. It can then be decided by inspection of a corpus of utterances in respect of which features language is structured – whether historical features or social features or semantic features and so on. Harris thought that languages are structured in terms of their *distributional* features: “in terms of the occurrence of parts (ultimately sounds) relative to other parts”.¹⁵³

The distribution of a particular linguistic element was understood as the sum of all the environments in which it occurs. A linguistic *environment*, within Harris’ approach, is just an array of co-occurring elements each in a particular position and which taken together make up an utterance. The first step in a distributional analysis is to divide the fragment of speech into segments to determine the regularity of the occurrence of one of those segments relative to the others. This must be done on grounds that are independent of the distributional analysis so Harris looked for the divisions that were “phonetically more obvious”.¹⁵⁴

¹⁵² Harris (1985)

¹⁵³ Harris (1985) p.26

¹⁵⁴ Harris (1985) p.41

The segments are then compared and statistically similar segments are grouped together. These similarity groupings themselves occur in partially similar and partially different environments so that we can in turn compare the distributions of the initial groupings. By such processes, Harris built up his linguistic elements. As the analysis developed, Harris could alter the original segmentations to the less phonetically obvious so as to achieve more regular distributions. The analysis proceeded by tweaking the underlying similarity groupings and framing a variety of more complex relations between the linguistic elements.

Harris felt confident in finding a distributional structure to all natural languages because he thought that: “All elements in a language can be grouped into classes whose relative occurrence can be stated exactly.”¹⁵⁵ He was confident in this because the elements of a language do not occur arbitrarily with respect to one another but only in certain positions relative to other elements. He saw these facts about occurrence as shaping our powers of expression:

The perennial man in the street believes that when he speaks he freely puts together whatever elements have the meanings he intends; but he does so only by choosing members of those classes that regularly occur together, and in the order in which these classes occur.¹⁵⁶

Moreover, these grammatical restrictions on distribution hold fairly uniformly. They are rarely, if ever, disregarded for reasons such as semantic needs. Harris thought that these statements about distributional regularities could descriptively cover the whole of a language without drawing on further types of information such as “normative rules” or “historical derivation”.

Harris’ view was that the structures he described genuinely exist in the languages in question “in as much as any scientific structure really obtains in the data which it describes”.¹⁵⁷ But though he believed that distributional facts somehow guide “the perennial man on the street”, he thought that the question of whether the structure of a language, conceived as a distributional structure amongst

¹⁵⁵ Harris (1985) p.27

¹⁵⁶ *ibid.*

¹⁵⁷ Harris (1985) p.30

physical elements, is realised in speakers' minds is a further question that linguistics does not address. Speakers' outward behaviour suggested to Harris that they perceive something like the distributional structure. But he was ambivalent about the idea that the structure is psychologically realised by those who are sensitive to it:

A reasonable expectation is that the distributional structure should exist in the speakers in the sense of reflecting their speaking habits...Concerning any habit, i.e. any predisposition to form new combinations along particular distributional lines rather than others, we know about its existence in speakers only if we have some outside evidence (such as association tests)¹⁵⁸

Notice how close this pre-Chomskian view is to Devitt's position, according to which whatever exists in the speaker should reflect the language they produce, in the sense that their competence *respects* without realising the structures of the language they produce. Harris' use of "reflect" here seems to parallel Devitt's use of the term "respect". Both Harris and Devitt, conceiving of grammatical properties as properties of sounds, suggest that no more than the speaker's mind reflecting or respecting the produced language can be reasonably inferred. It is a further question, by their lights, how the speaker is cognitively organised.

Historically, physicalist conceptions of linguistic properties, such as those of Bloomfield and Quine, have been coupled with behaviourist theories of language acquisition which involve learning language by generalising from the properties of observed verbal behaviour. This is also true of Harris:

[T]he position of the speakers is after all similar to that of the linguist. They have heard (and used) a great many utterances among which they perceive partial similarities: parts which occur in various combinations with each other. They produce new combinations of these along the lines of the ones they have heard. The formation of new utterances in the language is therefore based on the distributional relations – as changeably perceived by the speakers – among the parts of the previously heard utterances.¹⁵⁹

¹⁵⁸ Harris (1985) pp.31-2

¹⁵⁹ Harris (1985) p.32

This historical connection between physicalism and a behaviourist view of grammar acquisition may not be a necessary connection. There seems no principled reason why one could not hold a physicalist view of grammatical properties together with the nativist view that our acquisition of grammatical competence is an instance of growth under heavy innate constraint.

Chomsky thinks that Harris was more concerned with the description and classification of fragments of language, and the usefulness of such classification for other tasks such as machine translation, than with the *explanation* of the facts about speakers' languages.¹⁶⁰ However, there is a well-known sense in which describing and classifying languages so as to bring out distributional regularities, and then drawing inferences from these regularities, might be a form of explanation. On Hempel's deductive-nomological model, systems of regularities and entailments *can* be explanations.¹⁶¹ Hempel characterised scientific explanations as deductive arguments with at least one natural law statement amongst the premises. On this model, we have a statement that describes the phenomenon to be explained E and a set of statements S that "explain" E. At least one member of S must be a universal, or law-like, statement of the form *All Fs are Gs*. If the members of S are all true and the argument from S to E is deductively valid (as with *All Fs are Gs, x is an F, therefore x is a G*) then we have an explanation in the proposed sense. On this model, Harris' theories would count as forms of explanation because they involve statements of law-like regularities and deductions from them.

So why does Chomsky think that Harris was not so interested in explanation? Chomsky has a more substantial sense of explanation in mind than that supported by Hempel's model. Hempel's model does not appeal to notions of best explanations, causation or simplicity. But Chomsky is interested in *why* the sentences of speaker's languages take the special forms that they do, so he will therefore be interested in producing the best explanation of why the linguistic facts are as they are. In constructing a theory of language, Harris sought a "simple set of ordered statements that would *express* the empirical facts" but not explain them at

¹⁶⁰ See my Chapter Two.

¹⁶¹ Hempel (1965)

any deeper level.¹⁶² Chomsky treats linguistics as an attempt to muster the most explanatory theory of the properties of speaker's languages. It is to this end that Chomsky argues for the integration of linguistic descriptions with theories of psychological structure and language acquisition (see my Chapter Two).

The focus on operations and physically characterised evidence by Bloomfield and Harris suggests that linguistics is about observable matters, rather than about the underlying principles and unobservable entities of nature that other sciences discover. Such a focus on the observable is now thought to place overly narrow limits on the sorts of theories that can be entertained by scientists, and incapable of capturing the content of successful scientific theories which make inferences to unobservable entities and principles. Chomsky argues that this focus on observables in linguistics made manifest a commitment to unmotivated a priori restrictions on scientific inquiry. This serves to restrict the options for evaluating how successful theories are at capturing significant generalisations about language.¹⁶³

Linguists like Bloomfield and Harris took the operationalist approach further than it had been taken in other fields. They made it an aim of linguistic theory to devise *discovery procedures* that would automatically apply the techniques of segmentation and classification to any corpus and produce a correct analysis. Chomsky argued that no other science, not even the most developed of the hard sciences, has developed such procedures. Discovery procedures would, in effect, be an algorithm for automatically arriving at the correct theory in a particular domain drawing upon only the restricted data that physicalist ideology allows. Chomsky begins with the weaker assumption that a linguistic theory, like other scientific theories, is an attempt to explain some phenomena. What can be expected from linguistic theory is not a full account, or justification, for how its theories are produced but some explanatory perspective from which one can evaluate how good a linguistic theory is. Rather than discovery procedures, Chomsky aimed for competing grammatical theories, explicit explanations of the grammatical properties

¹⁶² Harris (1985) p.28

¹⁶³ See Chomsky (1955/75, 1957, 1965)

of languages which could be tested by whatever relevant data could be found so as to determine the best theory.

So for physicalists in the first half of the 20th century, grammatical properties were to be reduced to physical properties of uttered sounds and marks. Grammars were conceived as theories of this reality: theories of disturbances in the air or marks on surfaces. Chomsky began work on linguistics within this physicalist tradition before proposing a new way of exploring grammatical properties and a new conception of grammatical reality. Before we subject Chomsky's own conception of linguistic theory to scrutiny, it is worth noting an apparent limitation of reductive physicalism. Chomsky was struck early on by the abstractness of the categories and principles that grammatical theories seemed to be committed to.¹⁶⁴ Even fairly crude categorisations of sentence form are more abstract than physical descriptions in that they employ general categories like Noun, Verb, Sentence and the like. Moreover, a single physical string such as "Flying planes can be dangerous" can be more than one sentence form. Initially Chomsky accepted that these abstract categories could be explained away as the physicalist machinery for linguistic analysis is developed and sharpened. Just as Bloomfield had conceived of the order of application of his linguistic rules as a convenient fiction, so the categories might be construed as a way of speaking and not as categorisations that must match reality in a true linguistic theory.

But Chomsky felt compelled to recognise the importance of grammatical categories at a level abstracted from the physical disturbances in the air for theories of grammatical structure. Prior attempts to derive the phonological and grammatical categories from the physical properties of a corpus according to explicit procedures achieved little success. Chomsky found the inductive inferences from the corpus to phonological and grammatical categories to be unclear. The theoretical inferences relied on abstract categories and intuitions about the grammatical categories instanced in the corpus material. As Chomsky struggled to frame generalisations on the basis of the physical properties in the corpus, he continued to define grammatical categories at this more abstract level and then used these abstract

¹⁶⁴ See Katz (1981) pp.33-34 for discussion of Chomsky's pre-generative grammar view.

categories to examine languages, thereby giving up on the physicalist agenda that preceded him.¹⁶⁵

When all the interesting constructs of the developing theories involved abstract grammatical categories, and no explanatory appeal was being made to physical categories, Chomsky began to rethink the prevalent conception of grammatical reality. The abstract grammatical categories seemed so indispensable to these theories that he thought it disingenuous to hold that they could be explained away as convenient fictions. Chomsky inferred that the special structures that natural languages instantiate are abstract categories at the level of linguistic form rather than the level of physical occurrences. Although more abstract than the physical utterances and marks, Chomsky argued that grammars could be empirically tested according to whether their predictions about grammatical structure could explain the judgements of native speakers. He then developed a methodology within which a wide range of empirical evidence might bear upon hypotheses about the grammar of a speaker's language.¹⁶⁶

The view that theories of the grammatical properties of languages are "theories of speech-produced, physical disturbances in the communication medium" or that a theory of language is "a theory of the common distributional patterns in such physical disturbances" does not currently have many supporters amongst grammarians. Successes in generative grammar have been perceived to discredit the view that grammatical properties can be *reduced* to physical properties, because the categories that generative grammars appeal to are so much more abstract.

The early transformational grammars reflected the concern with more abstract categories because they investigated underlying levels of sentence structure, unobservable in the merely physical properties of utterances. Transformational grammars were able to overcome many of the objections to the idea that there were universal grammatical principles at work across languages. These objections were based on superficial, physical properties of structures in different languages such as the different sound patterns and orthographies. Transformational grammars were also able to capture relations in meaning between such structures as active and

¹⁶⁵ See Chomsky (1955/75) pp.30-33

¹⁶⁶ See Chomsky (1965) and my Chapter Two.

passives at a level of *deep structure* determined prior to the application of transformational rules that determine a sentence's *surface structure*. The transformations served to explain these differences in appearance evidenced by sentences of similar structural interpretation.

Once it was appreciated that the transformational grammars could capture a deeper level of similarity between linguistic structures, in terms of more abstract properties, it began to seem compelling that there might be such abstract uniformities across the languages people speak. For example, there are agreement relations in natural languages, but if this agreement relation is common between languages then it can't be reducible to the physical properties of the words because languages exhibit different physical patterns of sounds and written marks. The agreement relation is more abstract; it can be instanced in all these different sound patterns and written systems. The more abstract properties needed are *morphosyntactic* features that relate word shapes to grammatical properties. These abstract grammatical features permit the formulation of general hypotheses about grammatical structure. As Katz puts it:

Taxonomic constraints on the admissibility of constructs – imposed to ensure that everything at higher grammatical levels can be reduced back down to the physical events at the lowest level – precluded grammatical categories that are required to satisfy even minimal standards of grammatical explanation.¹⁶⁷

Chomsky's problem with the, once prevalent, reductive physicalism as a conception of grammatical properties is that it has no resources to characterise the special grammatical structures of speakers' languages. Neither does it explain why the amorphous noises, marks and signs exhibit the very special grammatical structures that cross-classify physical properties.

In summary, Chomsky abandoned three features of the physicalist conception: (1) that there is a specific, set data base for linguistics antecedent to theory construction, namely, corpuses of utterances consisting of emissions of noise, (2) that the scientific content of grammatical theories consists solely in what they say about this data base, and (3) that a linguistic theory was a theory about the

¹⁶⁷ Katz (1985) p.194

noises. To be clear, Chomsky's reasons presented thus far do not *demonstrate* the failure of even the reductive physicalist conceptions of grammatical properties associated with Bloomfield and Harris. At most, they suggest that Bloomfield and Harris offered unconvincing arguments for physicalism and highlight that early attempts to explain grammatical properties in reductive physical terms were not successful. One should not interpret Chomsky as offering a knock-down argument against such conceptions. The real issue for Chomsky is about explanation, and which conception of language promises richer and deeper explanations of the properties of speakers' languages.

Devitt attempts to revive a physicalist conception of grammar. On his linguistic conception, the properties described by generative grammars are *high-level* properties of physical entities in our environment. So, Devitt is not committed to the *reduction* of grammatical properties to physical properties that characterised pre-Chomskian physicalism. Yet Devitt agrees with Bloomfield and Harris that, ultimately, linguistic reality is to be located amongst the physical sounds and marks. Devitt's linguistic reality consists of "certain sounds in the air, inscriptions on paper"¹⁶⁸, "like the very words on this page".¹⁶⁹ Yet, Devitt is a *non-reductive* physicalist who thinks grammatical properties supervene on physical properties (inter alia). On Devitt's view, linguistics is a special science with its own proprietary domain of laws and generalisations, notwithstanding its supervenience on the physical. Devitt says that the outputs of a linguistic competence are "physical *sentence* tokens" and being a sentence for Devitt is a property that his *symbols* have.¹⁷⁰ As sounds and marks are only symbols and sentences insofar as they form parts of representational systems, Devitt's conception works at a higher level of generalisation than a reductive conception limited to physical disturbances.

A noteworthy feature of Devitt's view is that though he thinks that grammatical properties are themselves non-psychological properties, he is clear that he believes they have psychological determinants. He says:

¹⁶⁸ Devitt (2006) p. v

¹⁶⁹ Devitt (2006) p. 31

¹⁷⁰ *ibid.*

(T)he grammar is describing the syntactic properties of (idealised) linguistic expressions, certain sounds in the air, inscriptions on paper, and the like. These items are produced by minds, of course, and presumably get many of their properties somehow from minds, but they are not themselves mental.¹⁷¹

So, Devitt's non-psychological conception of grammatical properties is not a commitment to the mind-*independence* of grammatical properties.

One immediate question for Devitt's proposal is how a study that focuses on actual "physical sentence tokens" can account for the infinite number of sentences never uttered or written down: non-actual but possible sentences. To this Devitt responds:

The truth behind the talk of the nonactual can be simply that the grammar is lawlike. And the truth behind the talk of the infinite can be simply that there is no limit to the number of different sentence tokens that might be governed by the rules the grammar describes.¹⁷²

But Devitt doesn't want to merely repackage Chomsky's commitment to the infinite generative capacity of the language faculty. Devitt's idea of the grammar being lawlike and infinite in its generative capacity refers to the properties of a set of recursive rules that are realised in the physical environment and govern physical entities.

In my view, Devitt creates a difficulty for his account by saying that generative grammars are about tokens in the physical environment rather than types. If grammars are really just about physical tokens rather than more general types then it is more difficult to explain how the same structured expression could crop up in many places at the same time, or even multiple times. Devitt thinks that all talk of grammatical types can ultimately be cashed out in terms of tokens. But as Smith points out, if the same grammatical properties can be assigned to a wide variety of sounds, written marks and instances of signing, then it is difficult to see the sense in which linguists are studying physical tokens rather than the grammatical types that

¹⁷¹ Devitt (2006) p. v

¹⁷² Devitt (2006) p.27 fn.13

these otherwise disparate tokens instance.¹⁷³ However, this aspect of Devitt's view need not be adopted by all those who look to locate grammatical properties in the environment. A physicalist might instead hold that there are grammatical *types* located in our physical environment and that it is these environmental types that grammars are about.

1.3.3 Platonist Conceptions

Platonists about language conceive of languages as abstract rather than concrete objects.¹⁷⁴ According to Katz, a foremost proponent of a Platonist conception of grammatical theory: "Grammars are theories of the structure of sentences, conceived of as abstract objects in the way that Platonists in the philosophy of mathematics conceive of numbers."¹⁷⁵ Katz conceives the grammar of a language as an objective, eternal, and immutable structure which is independent not only of speakers' minds but of the whole spatio-temporal causal order.¹⁷⁶ Conceived in this way, particular grammars are theories about the sentences of natural languages, which are abstract objects. General linguistic theory states invariances over all such abstract objects. Katz thereby interprets UG as a theory of what is common to the structures of all such abstract objects and not as a theory of the uniform biological endowment that serves as a basis for language acquisition.

Katz claims that despite our *feeling* that "our language is more a part of us than it would seem to be on a Platonist conception" this is a mistake. What we have this feeling about, Katz claims, is not really our language which is mind-independent and abstract but rather "something closely related to language":

Whatever it is that we have these feelings about is what we acquire in the process called "language learning". Since in this process we do not acquire English itself but rather

¹⁷³ Smith (2006) p.439

¹⁷⁴ Katz (1981, 1985, 1996).

¹⁷⁵ Katz (1985) p.173

¹⁷⁶ See Katz (1981) for a clear statement of his Platonism.

knowledge of English...there is a distinction between a speaker's knowledge of a language and the language itself.¹⁷⁷

For Katz, processes like language acquisition do not actually involve any changes in a speaker's language but only changes in their knowledge of language and their relationship to an infinite range of abstract grammatical structures.

Katz was himself once a proponent of the psychological conception.¹⁷⁸ But he noticed that even if we give up on the reductive physicalist conception that preceded Chomsky, there is nothing to necessitate the choice of the psychological conception. Katz thinks that Platonism is "a real, if undeveloped, alternative".

Although I've classified the views of Bloomfield and Harris as varieties of physicalism, Katz describes them as *nominalist* because they hold that grammatical properties are token dateable, placeable parts of the physical world rather than abstracta. Katz's idea in describing these views as nominalist is that, just as nominalists deny the existence of abstract objects, these views deny the existence of abstract grammatical categories, treating grammatical descriptions as naming the physical occurrences. *Conceptualism* is Katz's name for the psychological conception. In explaining why Platonism constitutes a genuine alternative to the psychological conception, in light of the rejection of Bloomfield and Harris' views, Katz says:

Whatever Platonism's defects, they are surely not those of nominalism. Nominalism's defects stem from the insufficient abstractness of the interpretation of grammars as theories of sound waves and orthographic marks, while Platonism provides interpretations that accommodate the highest degree of abstraction. Since Platonism cannot be rejected for the same reasons as nominalism, new and different reasons are required to justify conceptualism over this third ontological position.¹⁷⁹

Katz infers from the demise of reductive physicalism that abstraction is a theoretical virtue in grammatical theory. It is a cornerstone of his position that grammatical theory ought to aim for the *maximum* level of abstraction. Katz argues that

¹⁷⁷ Katz (1981) pp.8-9

¹⁷⁸ Katz (1964), Chomsky and Katz (1974)

¹⁷⁹ Katz (1981) p.45

grammars ought to abstract away from the psychological states of speaker-hearers, as well as from physical occurrences, so that there are no substantive, non-grammatical constraints on grammars imposed by psychology. Though proponents of the psychological conception conceive of grammatical reality at a level of abstraction from the brain and from processing mechanisms, they still conceive of grammars as theories of speakers. So although the psychological conception allows for more abstract categorisations than reductive physicalism, Katz thinks it still constrains grammatical theory beyond the maximum level of abstraction. Katz claims that further abstraction from psychology is necessary to meet “intrinsic constraints concerning the successful description and explanation of grammatical structure.”¹⁸⁰ On the basis of this maximum abstraction principle, Katz argues that no interpretation of grammars other than the purely mathematical one will do. Platonism about grammatical theory follows if, like Katz, one thinks that mathematics should be understood as about Platonic objects.

So Katz denies that linguistics is an empirical science. He also claims that choosing between the psychological conception and the Platonist conception is an *a priori* issue about what a theory of language is. He argues that it would make no sense to construe their competing claims as empirical since Platonists deny that empirical matters are relevant to linguistic theory at all.¹⁸¹ So, any empirical evidence that was offered in favour of a psychological conception would beg the question against Platonism.

Katz’s argument here is problematic because one could equally well say that it made no sense to construe the competing claims as *a priori* claims because one side to the disagreement denies that non-empirical matters are relevant to linguistics. As Katz’s opponents think linguistic theory is an empirical theory, they might deny that *a priori* considerations are relevant and so hold that any *a priori* claims would beg the question. Even if it is true that on the Platonist conception no empirical evidence is relevant to the investigation of grammatical properties, this places no restriction on the considerations we might appeal to in deciding between conceptions of grammatical theory.

¹⁸⁰ Katz (1985) p.195. I address Katz’s arguments for this claim in §3.4.

¹⁸¹ Katz (1981)

It is anyway unclear that Platonist views cannot be motivated by empirical evidence. If we consider the case of mathematics, a Platonist conception of mathematics might be motivated because it is the best explanation of our scientific practices.¹⁸² Here the inference to Platonism would be going via empirical evidence from examining scientific theories. Even though mathematical knowledge is gained by a priori means that doesn't entail that one's theory of the ontology of mathematics must be a priori.

The important point is that it doesn't follow that it's nonsense to think of the competing claims of the Platonist and the psychological conception as empirically decidable, simply because one side says that grammatical theory is a priori. Suppose I thought that whether God exists can only be known empirically and I argue with a theist who claims it can be known only by a priori reasoning. Would we be begging the question against one another? It seems like we can still have a disagreement, and that it is not nonsense for me to think that the matter is empirically decidable though I may turn out to be wrong. Even assuming that the answer to the question we disagree over can be known only a priori or empirically and not both, we can disagree about which way it can be known without that making nonsense of our disagreement.

But Katz is happy to couch the issues in terms of explanatory coverage so long as the explanatory issues are put in terms that are neutral between conceptions of linguistics as an empirical or a mathematical science. His view is that Platonism, unlike the psychological conception, has explanatory benefits for the linguist because it places no restrictions on the degree of abstraction that grammatical theory can work at. This seems to commit him to a claim about the relation between grammars and speakers' knowledge of a grammar. For the benefit to accrue to Platonism, the facts about the grammars of natural languages and the facts about speakers' knowledge of a grammar must come apart. The speakers' knowledge of his language must be partial or distorted with the abstract grammatical structures extending beyond the subset of structures generated by the speaker's grammatical competence. For this is the only obvious way in which the psychological conception

¹⁸² Putnam (1975) makes the case for abstract objects on the grounds of their indispensability to empirical science.

could place unwarranted *restrictions* on the theory of language: limiting generative grammar to a psychological reality that grammatical reality outstrips.

However, there are infinitely many languages conceived as mathematical objects. If linguists are concerned to study the natural languages acquired in first language acquisition and spoken by humans then the Platonist needs to answer to two questions. Firstly, he needs to know how, on a Platonist conception, to determine in general which structures, amongst this infinite range, are structures of natural languages of the sort that concern the linguist. Secondly, he needs to know how, in any particular case, to determine which one of the infinitely many grammars is the grammar of a speaker's language.¹⁸³ If grammatical theory is concerned with infinite sets of abstract objects then to produce a grammar for a particular speaker's language we need to work out which of the sets is the one that characterises his language, and which is the grammar that generates it.

The psychological conception has ready answers to both these questions. According to the psychological conception, the linguist is concerned with the facts about the natural languages that speaker's acquire competence with and put to use in speaking and understanding. The linguist determines the grammar of a particular speaker's language by investigating the speaker's grammatical competence to see which grammar he has realised internally. Without answers to these questions the Platonist will be unable to motivate his conception of grammatical theory as an alternative to the psychological conception because he will be unable to constrain his "maximally abstract" inquiry to account for the grammatical properties of the natural languages of speakers, generally or in particular cases. The Platonist's maximum abstraction is a vice rather than a virtue if he cannot fix the target domain of inquiry. If the Platonist uses the psychological facts about speakers' acquired grammatical competences to answer these questions then it will be unclear that the conception he defends poses a genuine alternative to the psychological conception rather than presupposing the explanatory commitments of the psychological conception. There is no obvious reason why the "maximally abstract" theory, abstracted from facts about human cognition, should target all and only the

¹⁸³ This is the problem that Schiffer (2006) and Smith (2006) describe as "defining the actual language relation".

grammatical properties of the languages that humans acquire and speak, i.e. the properties of the natural languages.

Fodor puts this challenge to the Platonist in characteristically sharp terms:

In principle, he might just as well attend to the construction of grammars that predict only intuitions about sentences with more than seven vowels, or sentences whose twelfth word is 'grandmother', or sentences that happen to be uttered on Tuesday. Once you start to stipulate it's Liberty Hall.¹⁸⁴

If the Platonist view is that grammatical theory is a theory of the common properties of sentences then it is underspecified. Grammatical theory focuses on only a highly constrained range of the properties of sentences that are part of natural languages. Abstracted from the facts about what humans immediately cognise in speaking and understanding, there are infinitely many languages, and an infinite number of structural properties that we might find amongst the expressions of those languages. Amongst the infinite set of conceivable languages, including the natural languages but infinitely more languages besides, expressions will have a vast range of properties. There are properties like the spurious properties that Fodor highlights. We could also imagine extra, invented relations of structural command amongst constituents. Grammarians working within the Generative Enterprise are clearly not interested in these properties. The Platonist needs to explain what motivates their particular interest in just this subset, restricted in such highly-specific ways. The Platonist is stipulating because he is not committed to explaining any particular range of linguistic facts such as, for example, the *facts about the linguistic structures that speakers actually know*.

What Platonists have to provide is motivation for thinking that the relevant grammatical facts outstrip the facts about what is licensed by speakers' grammatical competences (I examine Katz's attempt to provide such motivation in §3.4). In the case of arithmetic there is a great deal of initial plausibility to the Platonist view because:

¹⁸⁴ Fodor (1985) p.177

[T]ruths of arithmetic are what they are, independent of any facts of individual psychology, and we seem to discover these truths in the way we discover facts about the physical world [*i.e. by examining a mind-independent reality*] ... knowing everything about the mind/brain, a Platonist would argue, we still have no basis for determining the truths of arithmetic.¹⁸⁵

Chomsky's view is that, by contrast, we have not been given "the slightest reason" to suppose that if we knew all the truths about the human mind/brain there would still be truths about human language that escape our grasp.

It is important to note that the position that Katz endorses is that grammatical structures are abstract, rather than concrete, *objects*. This is what motivates his choice of the Platonism/nominalism terminology. But there is a second sense of *abstract*, having to do not with whether the objects being described are abstract or concrete in the sense outlined, but with the *level of abstraction* involved in a description.

The point against reductive physicalism was that grammatical descriptions require a higher degree of abstraction than the physical descriptions, often cross-classifying physical descriptions so that the former properties cannot be reduced to the latter. In this sense of *abstract*, grammatical descriptions are *more* abstract since they have a greater level of generality. One can, however, abandon the reductive physicalism without conceiving of grammatical structures as abstract *objects*. The point about the insufficient abstraction involved in physical descriptions and the requirement for more abstract grammatical categories does not in itself suggest Platonism. To say that grammatical categories are more abstract could merely be to say that they are non-reducible *types* because they cross-classify the physical properties, with fundamental physical differences between instantiations of the same grammatical property and fundamental grammatical differences between instantiations of the same physical properties. In order to meet such requirements of abstraction one might employ the notion of a *grammatical type* rather than Katz's *Platonic objects*. A non-reductive physicalist, who locates grammatical properties in our physical environment, could appeal to grammatical types that are not reducible to the physical properties.

¹⁸⁵ Chomsky (1986) p.33

Katz fails to distinguish these two senses of “abstract” when he suggests that the requirement of grammatical types is support for the view that grammatical structures are Platonic objects. Katz’s claim is that grammatical structures are a special sort of entity. But the issue about generality and types of things is distinct from the issue about the space-time location and causality of a set of objects. Though Katz defines his position in terms of the aspatiality, atemporality and acausality of grammatical structures¹⁸⁶, he also frequently appeals to a type-token distinction to support his Platonism:

As C.S. Peirce drew the type/token distinction, and as everyone since understands it, types are abstract objects.¹⁸⁷

As Peirce understood the type/token distinction:

There will ordinarily be about twenty ‘the’s on a page, and of course they count as twenty words. In another sense of the word ‘word’, however, there is but one ‘the’ in the English language...it is impossible that this word should lie visibly on a page or be heard in any voice.¹⁸⁸

Peirce, like Katz, suggests here that types are not located in the physical world to be “visible on a page” or “heard in any voice”. Katz takes the type-token distinction to entail the existence of Platonic objects because he fails to distinguish these two senses of “abstract”: for him “abstract” and Platonic are one and the same. But, as Strawson points out, this doesn’t follow, for one can quite plausibly endorse abstract types without the Platonist ontology:

Let it be granted that spatio-temporal particulars – or spatio-temporal particulars of certain sorts – are model cases of what really exists or occurs. The fear is that a theoretical commitment to the existence of universals amounts to a confused half-assimilation of the general to the particular, accompanied perhaps by a confused analogical picture of the

¹⁸⁶ Katz (1981)

¹⁸⁷ Katz (1996) p.274

¹⁸⁸ Peirce (1958) p.423

relations of these spurious quasi-particulars, the universals, to actual objects to be found in space and time.¹⁸⁹

Katz's reasoning seems to be that if types are not physical objects then they are not in the physical world. But this doesn't necessarily follow. Supervenience physicalism, according to which everything supervenes on the physical, would be one position that admits types into the physical world without claiming that they are physical objects.

The view that grammatical categories are more abstract than particular physical tokens then might appeal only to the idea that there are grammatical categories or *types* which cross-classify the token, physical phenomena. One needn't arrive at the kind of two-world model (physical and Platonic) that Katz's talk of aspatial, atemporal and acausal objects suggests.¹⁹⁰

1.4 Argument Summary

My thesis meets Devitt's challenge to the psychological conception of generative grammar by building a positive case for that conception over its non-psychological competitors. My argument for the psychological conception consists of four parts. First, I defend a set of psychological goals for generative grammars (Chapter Two). Second, I argue that generative grammar makes an explanatory commitment to a psychological distinction between grammatical competence and linguistic performance (Chapter Three). Third, I argue that evidence from speakers' linguistic

¹⁸⁹ Strawson (1979) p.3

¹⁹⁰ Amongst those who believe in Platonic objects not everyone accepts each of the aspatiality, atemporality and acausality. Frege (1956) believed that Thoughts were aspatial and atemporal. He also held that though we cannot causally affect Thoughts, they can have a causal impact on the material world by being grasped and taken to be true. He compared Thoughts to the stars that we can apprehend in the night sky though we have no reciprocal powers to affect them. Frege held that Thoughts, though abstracta, could be a part of our mental life, the objects of our attitudes and that with reference to which our behaviour is to be explained.

intuitions is used to investigate this psychological distinction (Chapter Four). In the final part of my argument, I defend the view that the explanatory goals of generative grammar can be met by the theory of grammatical competence to which grammarians are committed (Chapter Five). Consequently, non-psychological grammatical properties are explanatorily dispensable to generative grammar, and grammatical theory is best interpreted as a theory of the psychological properties of grammatical competence.

In Chapter Two, I clarify Chomsky's methodological framework for generative grammar (§2.1), before defending the explanatory goals of descriptive adequacy (§2.2) and explanatory adequacy (§2.3). I then reject arguments against explanatory adequacy from Devitt and Katz (§2.4). In the Appendix I discuss recent developments in grammatical theory in which grammarians seek to justify generative grammars at a level *beyond explanatory adequacy*.

In Chapter Three, I first distinguish between the theoretical notion of competence and the commonsensical notion of competence (§3.1). I then argue that linguistic theory requires a theory of grammatical competence in order to determine the generative grammar of a speaker's language (§3.2). I then clarify the distinction between theories of grammatical competence and theories of processing which Devitt conflates (§3.3). At the end of the chapter, I defend the grammarian's use of a theory of grammatical competence to determine the generative grammars of natural languages against two arguments from Katz (§3.4).

In Chapter Four, I explain the role of evidence from speakers' linguistic intuitions in grammatical theory (§4.1). I outline and defend an orthodox model of the intuitions evidence according to which linguistic intuitions are psychological data, used to investigate the grammatical competence system (§4.2). I then reject Devitt's model of linguistic intuitions as speakers' theoretical beliefs about their language (§4.3). In the final section, I describe an alternative model of linguistic intuitions on which they constitute evidence for hypotheses about non-psychological grammatical properties located in speakers' environments (§4.4). This alternative model raises the question of how best to interpret the hypotheses on which the intuitions evidence is brought to bear, to which I turn in the final chapter.

In Chapter Five, I first consider arguments that non-psychological interpretations of generative grammar are inconsistent with the structural complexities that grammars detail and find them inconclusive (§5.2). I argue that the theory of the grammatical competence system to which generative grammar is explanatorily committed serves to meet the conditions of adequacy on a grammatical theory, and develop a parsimony argument against non-psychological interpretations of generative grammars (§5.3, §5.4). Finally, I consider and reject Devitt's claim that non-psychological grammatical properties are indispensable to theories of communication (§5.5).

2. What does Generative Grammar Explain?

2.1 Goals and Interests

Chomsky (1965) set out a new methodological framework for the study of language, which supplanted the frameworks of Bloomfield and Harris described in §1.3.2.¹⁹¹ Work on the properties of languages that was carried out within this new framework came to be known as *Generative Grammar*, a part of the *Generative Enterprise*.

Chomsky's new methodological framework was continuous with its predecessors in one broad sense. Like his predecessors, Chomsky was concerned with the structure of natural languages and to specify rules that operate on the minimally functioning linguistic elements so as to determine the linguistic structures. Though Chomsky conceived of theories of grammatical structure in a very different way, broadly speaking, his predecessors shared this aim.

But in a number of crucial respects the new methodological framework was a radical break with the assumptions that had gone before it. Perhaps most strikingly, generative grammars, as Chomsky conceives of them, are theories of the *speaker-hearers* who speak and understand languages. Chomsky says:

Linguistic theory is concerned primarily with an ideal speaker-listener, in a completely homogeneous speech-community, who knows its language perfectly and is unaffected by such grammatically irrelevant conditions as memory limitations, distractions, shifts of attention and interest, and errors (random or characteristic) in applying his knowledge of language in actual performance.¹⁹²

The methodology of his predecessors had focused on the study of utterances, conceived of as physical occurrences, and how the uttered sounds could be classified and compounded so as to reconstruct linguistic structures. What Chomsky suggests in the above passage is that the linguist re-orientates his attention away from the uttered speech sounds and towards the speaker-hearers whose knowledge

¹⁹¹ Chomsky (1965) pp.3-62

¹⁹² Chomsky (1965) p.3

of a language shapes those utterances and allows them to recognise structure in the utterances of others.

Part of the rationale for this suggestion is that utterances are complex *interaction effects*, the results of our putting the language we know to use in verbal behaviour. They are a reflection not just of the languages that speakers know but of the interaction of a wide variety of factors; Chomsky mentions memory limitations, distractions, shifts of attention and interest, and errors, but there are more besides. Amongst these factors, the language that a speaker knows is only one. What Chomsky proposes is that the study of complex linguistic phenomena, such as utterances, requires a divide and conquer strategy; just like the scientific study of other complex phenomena. The variety of factors that enter into verbal behaviours, like utterances, motivates Chomsky to draw a fundamental distinction between the language a speaker knows and the rest of the factors that enter into the etiology of linguistic behaviour. This is the provenance of the distinction between *competence* and *performance* (the topic of my Chapter Three).

A central fact motivating Chomsky is that by targeting the language a speaker knows, rather than the utterances and marks they produce, linguists get a clearer reflection of the grammar of their language. For as Chomsky notes, a record of what is produced will include false starts, deviations from rules, changes of plan in mid-course, failings of short-term memory and so on. These performance factors are not properties of the speaker's language; they might be quite general cognitive conditions or reflections of other cognitive demands on the speaker. So Chomsky suggests that generative grammars target the language that speakers know (or in which they are *competent*). This knowledge encompasses a system of rules underlying the observed linguistic behaviour (their linguistic *performances*). The system of knowledge is considered independently of extraneous limitations on a speaker's use of their language. Chomsky's suggestion is that generative grammars can use data from the linguistic performances, along with data from speakers' judgements about their language, as evidence to construct theories about the systems of rules that characterise the known languages. So as Chomsky originally conceived of generative grammar it is the study of rules of language that are realised by the mental states of speakers.

At the time this conception was taking shape, Chomsky claimed that “the issue of mentalism versus antimentalism in linguistics apparently has to do only with goals and interests, and not with questions of truth or falsity, sense or nonsense.”¹⁹³ Chomsky thought that the dispute over the mentalistic framework he was proposing amounted to a “rather idle controversy”.¹⁹⁴ He thought the dispute came down to three non-issues.

The first issue involved Bloomfield’s contrast between *mentalism* and *mechanism*. In drawing this contrast, Bloomfield suggested that mentalism is dualistic because it assumes a mental medium or substance distinct from physical mechanisms. But Chomsky was quick to point out that he is not committed to the system of knowledge being realised in a non-physical substance. So Chomsky claimed that Bloomfield’s distinction is misleading and irrelevant to the issue at hand. The theories of mental systems that Chomsky has in mind are framed at a level abstracted from physical mechanisms but this involves a commitment to a level of psychological explanation rather than to non-physical substance. The same holds true of the more recent notions of FL or I-language, which are theories of the mind/brain. More recently, Chomsky has argued that the special problems associated with locating the mind in a mechanistic world have effectively collapsed since Newton’s forces were accepted over the mechanistic philosophy, “exorcising the machine but leaving the ghost intact”.¹⁹⁵

The second issue involved his opponents’ commitment to behaviourism; which Chomsky characterised as the view that the data of linguistic performances exhaust the linguist’s domain of interest. Chomsky’s generative grammars are a partial explanation of those performance facts in terms of “deeper systems that underlie behaviour.”¹⁹⁶ As such Chomsky thought that the disagreement was not an arguable matter but “simply an expression of lack of interest in theory and explanation... [For a behaviourist] the enterprise has no point because all that interests him is the behaviour itself.”¹⁹⁷

¹⁹³ Chomsky (1965) fn.1

¹⁹⁴ *ibid.*

¹⁹⁵ See (Chomsky 2000, 2002)

¹⁹⁶ Chomsky (1965) fn.1

¹⁹⁷ *ibid.*

Chomsky thought it characteristic of this “lack of interest in theory” that behaviourist linguistics is limited to a summary and systematisation of the data on verbal behaviours, like utterances. By Chomsky’s lights these systematisations of the data do not constitute *explanatory* theory and he claims that it is unclear what sense of “theory” those behaviourists who proposed to limit themselves in this way had in mind.¹⁹⁸

The third issue concerned the use of data from our intuitive judgements to ascertain the structure of the known language. Generative grammars draw on evidence from native speaker’s intuitive consideration of linguistic material and the reports they would make of what was acceptable and unacceptable to them, as well as the ways the material could be interpreted or construed. Chomsky classed these judgements as a form of introspection. He thought that if we reject the behaviourist strictures then there is no good reason why linguists shouldn’t make use of evidence from introspection of their own linguistic intuition, as well as from other speakers’ reports of their intuitive responses to linguistic material. To disregard the judgements of native speakers would serve to limit the linguist to the performance data. Given the limitations of performance data for determining what constitutes a part of a speaker’s language and what is an effect of extraneous factors, Chomsky claims this would lead to “utter sterility” in grammatical theory.

The view that choosing between psychological and non-psychological conceptions of generative grammar, between “mentalism” and “antimentalism”, is a matter of “goals and interests” is a recurrent theme of Chomsky’s writings. In defending his view that generative grammar is a part of psychology, Chomsky often returns to the theme that generative grammar, psychologically conceived, is “a topic which one may or may not choose to study.”¹⁹⁹ Collins, echoing Chomsky, says:

As I understand it [the generative grammarian] has presented a research program guided by questions about our knowledge and development of language. By the terms of this program,

¹⁹⁸ As discussed in §1.3.2, there may be weaker notions of explanation according to which identifying regularities and making deductions from these regularities does count as a form of explanation. Such explanations do not give deeper answers to “why” questions or appeal to best theory criteria as Chomsky requires of explanations.

¹⁹⁹ Chomsky (1986) p. 4

linguistics is a branch of psychology. There is no fancy a priori argument here, just an invitation to look at language as a cognitive capacity. This might turn out to be the wrong way of approaching things, but it is, pro tem, a coherent way of proceeding.²⁰⁰

Chomsky notes that the decision to study a system of grammatical competence does stake out positions on matters of fact. It commits Chomsky to the existence of mental structure that shapes our production and comprehension of language. It is a condition on sustaining the psychological conception that one remains committed to such broad hypotheses about the psychology of language, although the content of those hypotheses may change drastically, according to developing theories of the language faculty and its role in the cognition of language.

However, Chomsky's claims about the psychological conception of generative grammar being a direct consequence of "goals and interests" need some careful handling. Unlike Bloomfield, who challenged mentalism across the board, not all opponents to Chomsky's psychological conception deny the existence of a mental reality underlying linguistic behaviour. In broad terms, to deny that there is some special cognitive organisation responsible for the patterns that we observe in speakers' linguistic behaviour is implausible. But opponents need not deny the existence of such cognitive organisation or that it constitutes a very good subject matter for *a* science. So they need not disagree with Chomsky over the coherence of the goal of understanding this mental reality or the interest in doing so. Rather what such opponents deny is that *generative grammars* themselves have this mental reality as their subject matter.

Opponents argue that Chomsky is mischaracterising the explanatory aims of generative grammarians. They agree that the "goals and interests" Chomsky describes are good ones, but think that attention to grammatical work and the theoretical framework in which it is carried out reveals that generative grammarians do not have the psychological aims Chomsky describes. Opponents claim that generative grammar does not have as its goal a theory of the speaker's knowledge of language, or grammatical competence. According to Devitt, this lack of self-understanding on the part of Chomsky, and other linguists, reflects a failure to

²⁰⁰ Collins (2006) p.4

recognise some general distinctions separating psychological theories from theories of the outputs of psychological processes (§1.3). Once we recognise these distinctions, Devitt thinks it clear that generative grammars are theories about properties located in our physical environment. For Katz, Chomsky's psychological conception is a kind of *bad faith*: a failure on Chomsky's part to exercise the full freedom offered by a Platonist conception of grammars as Platonic objects, instead choosing to labour under the stultifying constraints imposed by psychological theories about speakers (§1.3.3).

Chomsky claims that there are a number of levels of adequacy that grammatical theories should aim to achieve. Grammatical theories should be *extensionally* or *observationally* adequate, meaning that they generate the grammatical strings of a language, the right words in the right order. But Chomsky claims that they should also aim for *descriptive* adequacy, meaning that they assign the correct structural description to each sentence of a language indicating how the sentence is understood. Descriptive adequacy requires that the theory contains sufficient means to correctly describe all the different structures of the language under investigation and the interactions between them, thus providing analyses of all the possible sentences of the language. This is a more stringent condition than observational adequacy which only requires that the grammar generate the right set of strings and not that it assign the correct structure to the sentences in question. Descriptive adequacy applied to grammatical theory as a whole requires that this be done for all natural languages.

Chomsky also claims that there is a further level of adequacy that grammatical theories should meet, "a much deeper and hence much more rarely attainable level" of justification - *explanatory* adequacy.²⁰¹ This further condition of adequacy serves to determine one grammar, amongst the possible descriptively adequate grammars (DAGs), as *the* grammar of a speaker's language. To do so, an explanatorily adequate grammatical theory (EAG) will explain how, given a particular course of experience, a speaker acquires the grammar that they do. An EAG serves to capture the commonalities that all languages share and incorporates only those grammars that are humanly acquirable. Hence, it is "an explanatory

²⁰¹ Chomsky (1965) p.27

hypothesis about the form of language’; explaining what the grammar of a speaker’s language is in a *principled* way. General grammatical theory then embodies an empirical theory about our psychological endowment for language and the way it interacts with experience. These commonalities to all natural languages are called Universal Grammar (UG).

Generative grammarians have developed a theoretical framework for meeting explanatory adequacy called *Principle and Parameters* (P&P). Within P&P there are a set of commonalities across all acquirable languages (the *principles* of UG) and a number of ways in which individual grammars may vary whilst according with these common principles (the *parameters* of variation within UG). An important feature of P&P is that interaction between only a few parameters can yield the observed divergent structures in the sentences which individual grammars generate. Since the P&P framework has become relatively stable Chomsky has gone on to propose that there might be further conditions of adequacy on grammatical theory, *beyond explanatory adequacy*. This further condition of adequacy involves providing some explanation for why UG takes the form that it does. This would constitute an even deeper explanation of the properties of natural languages.

In this chapter, I aim to defend Chomsky’s claim about the explanatory goals and interests of the grammatical theories he has been at the forefront of developing. The central argument of this chapter is that generative grammars should meet conditions of descriptive adequacy (§2.2) and explanatory adequacy (§2.3, §2.4) and that these conditions impose psychological goals on grammatical theories. In my Appendix, I describe recent attempts to integrate grammatical theory with theories of the evolution of language, as grammatical theory moves *beyond explanatory adequacy*. In the Appendix, I aim only to explain the proposals that have been developed in this direction rather than attempt any assessment.

2.2 Descriptive Adequacy

2.2.1 Prescriptive Grammars

One way to approach the explanatory goals of generative grammarians is to consider how they differ from those of *prescriptive* grammarians. The notion of *the grammar of a language* is often associated with the idea that there is a set of correct prescriptions for speaking *good* French and English, and so forth. One might wonder how these prescriptive rules targeted by institutions like the French Academy differ from the rules targeted by generative linguists.

The rules that the French Academy decide upon are supposed to make prescriptions to French-speakers. When the Academy considers the French language they dictate that French is to have a certain set of properties covering the French lexicon, pronunciation and more besides. For example, in 1997 it was decided that the masculine form of the word “minister” (“le minister”) must be used for ministers of either gender, even though members of the French government had been using the feminine form to refer to female ministers.

Yet the French Academy doesn’t dictate as much about the languages of French-speakers as one might think. The kinds of rules that the Academy dictate are quite superficial from the generative grammarian’s perspective. The Academy specifies how French speakers *ought* to use the French language that they know. But a French speaker’s being able to put these instructions into action relies upon their knowledge of a bewildering range of linguistic constructions. These constructions are determined by underlying principles which “provide the substrate upon which those prescribed rules are parasitic”.²⁰² Even if the Academy provided a full list of official forms and provisions for style “they provide only examples and hints concerning the regular and productive syntactic processes.”²⁰³ Compare this with a *generative* grammar, which aims to be “perfectly explicit...it does not rely on the intelligence of the understanding reader but rather provides an explicit analysis of his contribution.”²⁰⁴

²⁰² Ludlow (2003) p.143

²⁰³ Chomsky (1965) p.5

²⁰⁴ Chomsky (1965) p.4

What the Academy offers is a set of proclamations about what the sounds and structures French speakers utter should be; something that would be palpably insufficient as an attempt to describe the full structure of the French language. These prescriptions don't cover the most basic properties of any language. Amongst those properties, perhaps the most fundamental is that human languages encompass an infinite variety of expressions, where each such expression is a sound-meaning pairing:

Human language is based on an elementary property that also seems to be biologically isolated: the property of discrete infinity, which is exhibited in its purest form by the natural numbers 1, 2, 3... There are three and four word sentences but no three and a half word sentences, and...they go on forever, it is always possible to construct a more complex one, with a definite form and meaning.²⁰⁵

This is a feature of all natural languages. The science of generative grammar aims to give some account of the membership of the class of this infinite variety of natural language expressions. A *generative* grammar is a fully explicit means of structurally describing the sentences of a language.²⁰⁶ A generative grammar for a language adequate to this task would “assign to each of an infinite range of sentences a structural description indicating how [the] sentence is understood” by speakers of the language.²⁰⁷

So generative grammars are supposed to correctly *describe* and *explain* the structural properties of sentences rather than *prescribe* what they ought to be. Generative grammars are not concerned with the linguistic proclivities speakers *ought* to have but with the full range of facts about the structures of natural languages as they *actually* are.

²⁰⁵ Chomsky (2000) p.4

²⁰⁶ Chomsky (1965) p.4

²⁰⁷ Chomsky (1965) p.5

2.2.2 Observational Adequacy

A generative grammar is *observationally*, or in Quine's terms "*extensionally*", adequate if it generates the right strings of words. A grammar that satisfies this condition for a particular language is also sometimes said to have *weak generative capacity*. But with perhaps one early exception²⁰⁸, the aim of generative grammars has always been something more than weak generative capacity. Rather the aim has been *strong generative capacity*. A grammar with strong generative capacity should include a generative procedure that assigns correct structural descriptions to the expressions of the language. If a grammatical theory correctly characterises the strong generative capacity of a language then it is said to be *descriptively adequate*. Chomsky argues that:

A fully adequate grammar must assign to each of an infinite range of sentences a structural description indicating how this sentence is understood.²⁰⁹

In this way, generative grammar differs from merely extensional approaches, like Quine's, which aim to characterise a language in terms of a set of well-formed strings. Such extensional approaches are indifferent to the generative procedure involved and the structural descriptions assigned, so long as the right set of strings is generated. Quine's extensional adequacy condition for grammars - that they generate the right strings - stops short of the aims of generative grammars. As Chomsky observes, generative grammars do not "keep to the only entities admissible from Quine's 'realistic point of view', namely 'the right totality' of well-formed English sentences".²¹⁰ Appealing to best theory arguments, generative grammars propose that some particular account of sentence generation is correct. This made little sense for Quine, given his view that the only reality to grammar is the totality of well-formed strings. From Quine's point of view, procedures of assigning grammatical structure are a kind of *choice*, like the choice of one or another set of axioms for a formal language. According to Quine's criterion of

²⁰⁸ Chomsky (1957)

²⁰⁹ Chomsky (1965) p.5.

²¹⁰ Chomsky (2003c) p.305

adequacy one can choose whatever set of rules for generating well-formed strings that one likes, chopping up the strings in whatever way these rules do, so long as one preserves the same totality of well-formed strings.

As Quine thought that the proper concern of linguistics was only extensional adequacy, he thought that two grammars which generate the same set of well-formed strings are empirically equivalent even if they differ in their generative procedure and the structural descriptions they assign. Two grammars that share weak generative capacity could only differ, for Quine, in terms of their simplicity. Where two different but extensionally equivalent grammars were equally simple, Quine deemed them grammatically equivalent.

Chomsky claims that Quine's approach simply ignores a basic fact about languages, namely, that they provide speaker's with a finite means to generate an infinite class of structurally articulated expressions. What the grammarian is concerned about is these structures rather than strings. The strings do not cut finely enough to tell which structures are part of a speaker's language. For example, the same string can be more than one structure of a speaker's language if it is ambiguous. Grammarians want to know how each string is structured as a sentence of the language. So they want grammars to strongly generate the language by assigning sentences their proper structural descriptions. Insofar as Quine was prepared to specify such a finite means for generating the structural descriptions, Chomsky thought Quine to be speculating about strong generative capacity: hypothesising about the generative procedure and not merely the set of well-formed strings that a language encompasses.

Quine claims that the choice of grammars to generate the well-formed strings of a language is *indeterminate* because the "business of syntax is the demarcation of strings of phonemes proper to the language" but "more than one battery of grammatical constructions and vocabulary will probably be capable of generating the same total output of strings."²¹¹ However, Quine wants to distinguish the indeterminacy facing grammatical theory from that facing translation: "there is no indeterminacy analogous to that of translation. Indeterminacy of translation

²¹¹ Quine (1990) p.49

consists rather in conflict between the outputs themselves.”²¹² For Quine the inscrutable differences between grammars that engender indeterminacy do not consist in their outputting different strings but in their doing so according to different procedures. By contrast, the difference between empirically equivalent translation theories consists in the translations that they output.

Chomsky agrees with Quine that *formal* languages don’t have a designated syntax. He claims “they just have a set of well-formed expressions; the syntax can be anything you like”:

There’s no right answer to the question: what are the *true* rules of formation for well-formed formulas of arithmetic? What are the axioms of arithmetic? The answer is: any set of axioms you like to generate all the theorems. It’s the theorems that are real, not the axioms; the axioms are just a way of describing them, one of many ways. Similarly, if you invent a computer language, it doesn’t really matter which rules you pick to characterize its expressions; it’s the expressions that are the language, not the specific computational system that characterises them.²¹³

But Chomsky denies that this is true of *natural* languages which are naturally occurring objects. With natural languages we *do* care about their structural complexity and the correct structural descriptions of their expressions. Natural languages have this structural complexity independently of our choosing a rule system to characterise them. The only argument that Quine has for limiting grammatical theory to weak generative capacity is an indeterminacy claim about the grammars of languages. Chomsky thinks that grammatical theory faces the same underdetermination of theory by evidence as other empirical theories. But Quine’s indeterminacy thesis only follows with the addition of behaviourist restrictions on the relevant evidence; restrictions which I argued in §1.3.1 that Quine fails to justify. So there is no cogent argument for the indeterminacy of grammatical theory that restricts it to considerations of extensional adequacy. It is Chomsky’s positive view that in the case of natural languages, there is something ‘in the head’, our

²¹² Quine (1990) p.49

²¹³ Chomsky (2002) pp.110-1

knowledge of language, or grammatical competence, that *is* the computational system and determines the structural properties of our language.

2.2.3 Descriptive Adequacy and Productivity

A DAG assigns structural descriptions to the sentences of a language indicating how they are understood. A sentence, but not a mere string of words, is a grammatically structured object. Generative grammar then is really the study of grammatical structures rather than strings. It must account for the grammatical structures of sentences, and explain the grammatical relations that we recognise between them as well as explaining why certain structures are ruled out of speakers' languages.

To take a classic example, an observationally adequate grammar will count the strings *John is easy to please* and *John is eager to please* as amongst the well-formed strings of English. But it will not explain why these strings, which are observationally very similar arrangements of words, are structured quite differently. The different structures of these two sentences can be brought out by considering their relations to other sentences. We can paraphrase *John is easy to please* as *It is easy to please John*. But we cannot paraphrase *John is eager to please* as **It is eager to please John*. *John is easy to please* is structured so that it is John that is being pleased. This is not true of *John is eager to please* in which it is John that is doing the pleasing.²¹⁴ This example shows that there are grammatical facts that observationally adequate grammars will not reveal.

DAGs are required to assign structural descriptions to an infinite range of sentences. Given any finite collection of sound-meaning pairings of a natural language, we can always construct another such pair using the finite stock of

²¹⁴ The analyses of these sentences fill the subject and object positions of the infinitival clauses with empty categories. "John" is interpreted in object position in "John is easy to please", which is understood as "John is easy (someone) to please (John)". Whereas, "John" is interpreted in subject position in "John is eager to please", which is understood as "John is eager (John) to please (someone)".

vocabulary and procedures for grammatical combination. This fact is often referred to as the *productivity* of language. To explain the productivity of the languages speakers know, generative grammarians postulate grammars with a finite number of recursive rules and a finite stock of lexical items.

If a grammarian were simply concerned with a language as a set of objects independent of speaker's knowledge, then there would be less motivation to assign languages recursive, combinatorial structure acting upon finite elements. We could think of the language as a set of structured sound-meaning pairings - (S_1, M_1) , (S_2, M_2) ... (S_n, M_n) - with an infinite number of members. There is no motivation to think that such independent sets must be built out of a finite number of elements and recursive procedures. One could just postulate the infinite pairs without postulating a recursive procedure. The idea that languages involve a finite basis is connected to their being known by speakers with finite limitations. The best explanation of the productivity of languages that speakers know, given speakers' finite limitations, is the recursive operation of grammatical rules on a relatively fixed stock of lexical items. So, it is only if one is interested in the known languages - in what speakers know - that the issue of productivity becomes an explanatory goal and receives an explanation.

The fact that speakers of productive languages have the potential to produce and understand an infinite number of new expressions is sometimes referred to as the *creativity* of our use of language. Language users can get a sense of the channelled but infinite expressive power that their language provides:

The core property of discrete infinity is intuitively familiar to every language user. Sentences are built up of discrete units: There are 6-word sentences and 7-word sentences but no 6.5 word sentences. There is no longest sentence (any candidate sentence can be trumped by, for example, embedding it in "Mary thinks that..."), and there is no non-arbitrary upper bound to sentence length.²¹⁵

Chomsky thinks that the normal *use* of language is creative in the sense "that much of what we say in the course of normal language use is entirely new, not a repetition of anything that we have heard before and not even similar in pattern... The number

²¹⁵ HCF (2002) p.1571

of sentences in one's native language that one will immediately understand with no feeling of difficulty or strangeness is astronomical."²¹⁶ If a grammar permits the generation of all and only the infinite grammatical expressions of the language, then a speaker's knowledge of such a grammar can be used in explaining the creativity exhibited in their linguistic behaviour. It is a primary concern of the generative enterprise to account for the productivity of the languages we acquire, and thereby explain this creativity in our use of language, on the basis that the productive languages are mentally realised.

Chomsky, and others, made this explanatory connection explicit from the beginning of generative grammar. Lees says that the explanatory goals of generative grammar ought not to be restricted to characterising the properties of some finite corpus "else it could not account for the fact that *speakers are able* to extend the corpus indefinitely."²¹⁷ Chomsky's stated aim in investigating generative grammars was to account for "a large store of knowledge...and a mass of feelings and understandings".²¹⁸ The "mass of feelings and understandings" Chomsky wanted to explain can be investigated via speakers' conscious awareness of their language and judgements. It is, in a sense, misleading to describe grammars that attempt to meet this condition as meeting only "descriptive" adequacy, and contrast this with "explanatory" adequacy, for grammars that aim to meet these conditions are already pursuing explanatory aims.

This is a sharp departure from the views of Bloomfield and Harris. Harris was sceptical of attempts to explain the creativity of language users in terms of their standing knowledge of a productive grammar. He writes:

[E]ven when our structure can predict new utterances, we do not know that it always reflects a previously existing neural association in the speakers (different from the associations which do not, at a given time, produce the new utterances)... [It] means only that the pattern or habit existed in the speakers at the time of the new formation, not necessarily before: the 'habit' – the readiness to combine these elements productively –

²¹⁶ Chomsky (1972) p.10

²¹⁷ Lees (1957) p.382 my italics.

²¹⁸ Chomsky (1955/75) pp.62-3

may have developed only when the need arose, by association of words that were partially similar as to composition and environment.²¹⁹

In contrast, Chomsky takes the creative aspect of language use *as* evidence that speakers of a language have enduring mental structures with the potential to generate the relevant novel sentences. Harris assumes an associationist psychology according to which human “habits” are shaped by indefinite processes of association. If one replaces the associationist psychology that Harris assumes with a cognitive psychology of enduring mental structures and principles then Harris’s point looks far less plausible. Where generative grammars predict the new utterances that speakers produce and comprehend, the natural cognitive explanation is that speakers have some corresponding cognitive structure in place, not that they develop capacities as required on each occasion.

It was a distinctively new concern of Chomsky’s conception of grammatical theory to produce DAGs for speakers’ languages, and to offer some explanation of the creativity of language use on the basis of their knowledge of a productive grammar. Attention to the grammars that speakers know raises further questions. A human child is exposed to only a finite amount of linguistic material; yet, barring environmental or developmental problems they all acquire such a grammar. A natural question is how our finite exposure to linguistic data can be squared with our infinite variety of expressions. We need to discover how the grammars speakers know are structured so as to make an infinite variety of expressions available. But such a theory of grammatical structure must be supported by an acquisition model on which speakers can acquire such grammars. The actual DAG that assigns structural descriptions to the sentences of our language over an unbounded range must be acquirable in first language acquisition if it is to constitute an explanation of our infinite expressive capabilities. Not all the possible DAGs for a speaker’s language are plausible candidates for the grammar the speaker knows. This is where *explanatory adequacy* comes in.

²¹⁹ Harris (1985) p.31

2.3 Explanatory Adequacy

2.3.1 Further Conditions of Adequacy?

A generative grammar is *explanatorily adequate* if it serves to determine which grammar the speaker actually knows from amongst the DAGs for his language. Insofar as an EAG is a hypothesis about which grammar a speaker actually knows it involves hypothesising about grammar acquisition. We know that speakers of all languages acquired in first language acquisition (i.e. the natural languages) develop the competence to extend their recognition and production of grammatical structures indefinitely. As a matter of course we all recognise sentences that may never have been uttered before by anyone, such as:

(6) Anyone who has been to Barry Smith's philosophy of linguistics seminar in London, at Stewart House, on Wednesdays, from two until four, more than once, and paid attention, has had the recursive aspect of grammar drummed into them.

Beyond this general recursive aspect of human language, there are other significant, language-specific generalisations that hold across all natural languages (see §5.4). It is a goal of generative grammar to explain the properties of linguistic structures in terms of these deeper principles.

One might wonder why Chomsky thinks that grammatical explanations will involve us in theorising about our acquisition of grammatical principles. Explanatory adequacy has always been a long range goal of generative grammar. But as Stephen Stich puts it, once we have achieved a DAG:

It might seem our job is finished. We set ourselves to giving an account of the grammarian's doings in building a grammar, and this we have done. But...such accounts go on to talk of *linguistic theory*, *acquisition models*, *evaluation measures* and other notions related to the question of how a speaker acquires his grammar. Moreover, the discussion of these notions is not a simple addition to the account of the grammarian's work in constructing a grammar. Rather it is an intrinsic part of that account. Yet why this is so is far from obvious. Constructing a theory of grammar acquisition is surely a fascinating project and one which would naturally catch a grammarian's eye. But at first blush at least,

it would seem to be a new project, largely distinct from the job of constructing grammars for individual languages. Why, then, do Chomsky and others view the study of acquisition as intrinsic to the construction of grammars for individual languages?²²⁰

In what follows I want to explain and defend Chomsky view that hypotheses about language acquisition are an integral part of the best grammatical theories of speakers' languages.

2.3.2 Grammar Acquisition and Universal Grammar

A cursory inspection of the properties of human languages reveals that there is much apparent variation as one compares different languages, such as English and Japanese. Yet speakers of each language are born with a uniform endowment for language. As Chomsky conceives of explanatory adequacy it is the attempt to explain the observed variation in acquired languages on the basis of a uniform endowment. This uniform endowment for language, UG, is hypothesised to admit only limited options:

[F]or H to know L is for H to have a certain I-language... UG is now construed as the theory of human I-languages, a system of conditions deriving from human biological

²²⁰ Stich (1985) pp.132-3 Stich agrees that the notion of explanatory adequacy has an important role to play in grammatical theory, he says that it is the EAG "which the linguist seeks to uncover." (p.136) But Stich gives an account of the role of explanatory adequacy which invokes an unjustified indeterminacy thesis about grammar and a general scepticism about mental representation. He sees explanatory adequacy as a practical measure for cutting down on indeterminacy and not as a means to the truth about the grammar the speaker mentally represents. Stich claims that though a grammar can be "true of" the speaker, there is no fact of the matter about which of the many grammars true of the speaker is the actual grammar of his language, and the speaker does not mentally represent any of them. Moreover, Stich construes the properties of UG as properties common to all possible DAGs rather than as those properties common to all the grammars that speakers have actually acquired. Stich effectively replaces the requirement that an EAG selects a DAG, on the basis of the primary linguistic data, with the requirement that it provides all the DAGs. See Chomsky and Katz (1974) for a critique of Stich's proposals.

endowment that identifies the I-languages that are humanly accessible under normal conditions.²²¹

Constructing a theory that is explanatorily adequate requires developing an account of UG, the innate grammatical principles common to all the languages that speakers acquire, and of how the parameter values of those principles are determined by experience. Building such a theory of grammar acquisition into generative grammar serves to explain in any particular case which grammar a speaker actually knows amongst those DAGs that might cover his language:

To the extent that a linguistic theory succeeds in selecting a descriptively adequate grammar on the basis of primary linguistic data, we can say that it meets the condition of *explanatory adequacy*. That is, to this extent, it offers an explanation for the intuition of the native speaker on the basis of an empirical hypothesis concerning the innate predisposition of the child.²²²

EAGs offer an explanation of the native speaker's judgements about the sentences of their language on the basis of their having acquired a grammar of a particular sort. Explanatory adequacy provides a reason for thinking that one particular account of sentence generation is the correct one because it is the one that the speaker has acquired and so it is the one involved in the provenance of their judgements.

As theories that aim for explanatory adequacy involve an empirical hypothesis about grammar acquisition, they can be falsified by showing that they fail to provide a DAG for a child who receives primary linguistic data from a particular human language. For children are not innately constrained to acquire one natural language rather than another. A theory is more explanatorily adequate when it provides more far reaching explanations of the natural languages.

Though this approach is radical in the context of pre-Chomskian linguistics, Chomsky suggests it is partly a reinvention of an old approach to language, as proposed by Beattie in 1788:

²²¹ Chomsky (1986) p.23

²²² Chomsky (1965) pp.25-6

Languages, therefore, resemble men in this respect, that, though each has peculiarities, whereby it is distinguished from every other, yet all have certain qualities in common... Those things, that all languages have in common, or that are necessary to every language, are treated of in a science, which some have called *Universal* or *Philosophical* grammar.²²³

The generative grammarian attempts to describe the facts about human languages rigorously and provide support from a theory of UG for theories of speakers' languages. As such, the grammar of a particular language is underwritten by UG, which:

accommodates the creative aspect of language use and expresses the deep-seated regularities which, being universal, are omitted from the [particular] grammar...It is only when supplemented by a universal grammar that the [particular] grammar of a language provides a full account of the speaker's competence.²²⁴

The aim of correctly describing the properties of *any and all* human languages means that generative grammarians make *universal* claims: in their fundamentals the theories must be adequate for any language that a human can acquire in first language acquisition. The perspective is one from which there is just one language, or just *language*, with lots of dialects or variations. Linguists immediately prior to Chomsky had not recognised the need to add a universal grammar to particular grammars in order to explain creativity and other deep principles of language. Chomsky thought such grammars were thus descriptively, as well as explanatorily, inadequate.

2.3.3 The Role of Explanatory Adequacy

Chomsky claims that aiming for explanatory as well as descriptive adequacy is critical if linguistics is to advance as a science. Just as descriptive coverage can be achieved by a large number of conflicting theories in other sciences, so descriptive

²²³ Beattie (1788)

²²⁴ Chomsky (1965) p.6

coverage of the data on languages might be achieved by large numbers of conflicting grammatical theories. DAGs state a procedure that generates the structured sentences of a speaker's language. As Stich points out, we might expect that in the absence of hypotheses about the grammar that a speaker has actually acquired, there will be more than one procedure that can achieve these structural descriptions. Stich says of a DAG that it is merely an axiomatised theory, and:

[I]t is a truism that a theory that can be axiomatized at all can be axiomatized in radically different ways... When the job has been done there are indefinitely many variants each of which captures the known intuitions equally well and predicts unprobed intuitions equally well (or poorly). Somehow the grammarian does come up with a single theory.²²⁵

Stich wonders "What principle can he use to guide his choice?"²²⁶ Chomsky agrees that if we consider languages as axiomatised theories then "it doesn't really matter which rules you pick to characterize its expressions."²²⁷ But he denies that this is true of natural languages which are naturally occurring objects. The grammarian wants to know about the actual structural complexity of natural languages and the correct structural descriptions of their expressions. To this end, Chomsky integrates the study of linguistic structure into a complex of empirical studies focusing on our acquisition and cognition of grammar. An explanatory benefit of doing so is that amongst the conflicting theories which can cover the data, a theory that can explain the data on the basis of an empirical hypothesis about the form of known languages is supported by more evidence than one that can only explain the data by ad hoc means. Chomsky's claim is that there is something 'in our heads', (namely, the grammar we have acquired) that determines the structure of our language and thereby determines which theory of our language is correct from amongst the DAGS.

Purely descriptive coverage of a language could be attained by postulating rules on a datum-by-datum basis, yielding DAGs that offer no theoretical insight.

²²⁵ Stich (1985) pp.133-4

²²⁶ Stich (1985) p.134

²²⁷ Chomsky (2002) pp.110-1

Chomsky notes that coverage of data alone is not in itself “an achievement of any particular theoretical interest”.²²⁸ He goes on to claim that:

A theory of generative grammar may be descriptively adequate and yet leave unexpressed major features that are defining properties of natural language and that distinguish natural languages from arbitrary symbol systems. It is for just this reason that the attempt to achieve explanatory adequacy...is so crucial at every stage of understanding of linguistic structure, despite the fact that even descriptive adequacy on a broad scale may be an unrealized goal.²²⁹

So Chomsky conceives of grammatical theory as appealing to a complex of data in order to decide between competing theories of the grammars speakers know and the innate information that provides a basis for all natural languages.

From this perspective, considerations of explanatory adequacy are vital to constructing grammatical theories. Suppose we have two theories T and T'. Both T and T' can adequately describe the intuitions data. T adequately describes the data but is *not* supported by an empirically confirmed theory about UG and the mechanisms for selecting a UG language. T' adequately describes the data and *is* supported by an empirically confirmed theory about UG and the mechanisms for selecting a UG language. T' affords us the more insight into the structure of the naturally occurring languages speakers know. Therefore, T' is more empirically justified. Fodor puts the point starkly:

[N]obody is interested in grammars that *demonstrably could not be learned*, though there is no reason why some such grammars shouldn't be [descriptively adequate grammars].²³⁰

²²⁸ Chomsky (1965) p.26. Chomsky (1965 p.36) says: “Although even descriptive adequacy on a large scale is by no means easy to approach, it is crucial for the productive development of linguistic theory that much higher goals than this be pursued.” Similar considerations might apply in the case of Devitt's example of bee dances (§1.3). The theorist does not just want any possible descriptively adequate rule system for the bee's dance of which there may be many, but rather that system of rules and properties which the bee is sensitive to.

²²⁹ Chomsky (1965) p.24

²³⁰ Fodor (1985) p.153

As Fodor suggests, generative grammarians would not consider a theory that is inconsistent with known facts about grammar acquisition. More moderately put, generative grammarians aim to discover the facts about the languages that speakers acquire and as such seek integration with empirical studies that investigate language acquisition (but also language processing, language deficits, and other branches of empirical theory). This is a consequence of targeting the *natural* languages: those languages that speakers actually acquire.

The interaction of descriptive and explanatory adequacy has been a driving force in developments in generative grammar. In order to accommodate the complex facts about particular languages, generative grammarians were driven to posit more rules and descriptive technology. In early generative grammars, prior to P&P, the transformational rules thought essential to meeting descriptive adequacy proliferated. These transformational rules in question helped explain the relation in meaning, but difference in structure, between sentences like (7) and (8):

(7) The dog chased the cat.

(8) The cat was chased by the dog.

Sentences (7) and (8) have the same interpretation but look and sound different. The proposed explanation was that sentences have more than one level of structure. It was hypothesised that sentences have a *deep structure* that fixes the interpretation of the sentence, as well as a *surface structure* that is much closer to the structure of the sentence that we hear.²³¹ Generative grammars were hypothesised to have two corresponding components: one to generate the deep structures and a transformational component that maps deep structures into surface structures.

²³¹ These theories were soon elaborated so that both deep and surface structure contributed to interpretation when it was noticed that transformations into surface structure can have truth-conditional consequences. This is evidenced in the difference between sentences like “Everyone in the room knows at least two languages” and “At least two languages are known by everyone in the room” where the sentences were hypothesised to share deep structure but the latter, unlike the former, may be false if the two languages known by each person in the room are different. In the “Extended Standard Theory”, deep structure fixed relations of subject and object whilst surface structure fixed relations such as scope, anaphora, focus and presupposition.

Sentences like (7) and (8) were thought to share a deep structure but map onto different surface structures; (8) involves a permutation of *dog* and *cat*, so that they are pronounced in different positions from those in which they are interpreted, and the insertion of further lexical items.

There were many other linguistic phenomena that could be accommodated with new transformations, including question-formation and structural ellipsis. But such phenomena were so widespread in languages and so complex that generative grammars began to posit more and more transformational rules and descriptive technology of a sort that was highly intricate and peculiar to each language. This gave rise to a puzzle:

The paradox was that in order to give an accurate descriptive account it seemed necessary to have huge proliferation of rule systems of a great variety, different rules for different grammatical constructions. For instance, relative clauses look different from interrogative clauses and the VP in Hungarian is different from the NP and they are all different from English; so the systems exploded in complexity...still, somehow children are reaching these states of knowledge which have apparently great complexity, and differentiation and diversity – and that can't be. Each child is capable of acquiring any such state; children are not especially designed for one or the other... But in that case it appears to be inconsistent with observed diversity and proliferation.²³²

It became implausible that children were determining amongst these highly intricate rule systems in acquiring their grammar. Simplification of the descriptive technology was a substantial step towards answering the question of how the child could select the grammar from an otherwise bewildering range of choices. Generative grammars expanded the phrase structure component of the grammar that generated the deep structures and greatly reduced the number and complexity of the transformational rules by subsuming them under more general principles. The new theories were more explanatorily adequate but also grammatically deeper because their principles were fewer in number and more unified, covering a greater range of phenomena than the proliferating transformations. The paradox was resolved by showing that:

²³² Chomsky (2002) p.93

the diversity of rules is superficial, that you can find very general principles that all rules adhere to, and if you abstract those principles from the rules...then the systems that remain look much simpler. That's the research strategy...that went on for a long time with efforts to reduce the variety and complexity of phrase structure grammars, of transformational grammars, and so on.²³³

In place of the large number of highly specific transformations (complex rules for complex constructions in different languages), generative grammars now appeal to a very small number of grammatical operations, extremely general principles with parametric options.²³⁴ The old rules have become artefacts of these deeper principles.²³⁵

The fact that there can be more than one DAG for a language is not a special problem for grammatical theory. It is the same situation that exists in other sciences where there are conflicting theories compatible with the evidence and the scientist looks for more evidence to decide between the theories from whatever sources are available. Beyond the intuitive judgements of native speakers, choices between DAGs can be justified on the basis of evidence from psychological theories of acquisition and, in principle, the brain sciences. Explanatory adequacy makes it a condition on an adequate grammatical theory that it meets this further evidence.²³⁶

²³³ Chomsky 2002 p.93

²³⁴ See my Appendix for a discussion of minimalist theories which attempt to explain grammatical structures in terms of the interaction of the properties of lexical features with some very simple structure building operations.

²³⁵ To take another example of explanatory adequacy at work, learnability arguments have been developed in favour of the widespread adoption of uniform binary branching, see Haegeman (1992).

²³⁶ Chomsky (1965 pp.35-36) points out that as well as considerations of explanatory adequacy being necessary to obtain a revealing DAG, some grip on descriptive adequacy is a prerequisite on an EAG because a theory of grammar acquisition requires we have some plausible theory about the properties acquired.

2.4 Arguments against Explanatory Adequacy

The goal of explanatory adequacy has been met with some philosophical resistance. The resistance builds on the thought that theories about language acquisition and theories about languages themselves are distinct theories. There is more than one way of developing this thought.

One way is to claim that theories about grammar acquisition, which are theories of *how* we acquire grammars, do not inform grammatical theories, which are theories of *what* the structural properties of languages are. This is the position that Katz holds. He says:

If, by some chance, one linguistic theory were to coincide to a greater extent than others with a psychological theory, this would have no more significance for linguistics than the coincidence of Riemannian geometry with physical theory has for pure mathematics.²³⁷

As discussed in §1.3.3, Katz does not think that any empirical evidence, such as evidence from grammar acquisition, is relevant to grammatical theory at all. He holds that grammatical theory is an “intuitional science” and that grammarians rely on rational intuition in the same way that he thinks mathematicians and logicians do.

There is a second, more moderate, view that offers resistance to adopting explanatory adequacy, according to which there is an *evidential* connection between generative grammars and theories of grammar acquisition, but it is not a *criterion of adequacy* on grammatical theory that it *involves* a theory of grammar acquisition. The idea behind this second position is that, whilst facts about grammar acquisition can inform grammatical theories, grammatical theories do not incorporate theories of acquisition. This is the position that Devitt endorses. Devitt wants to accommodate the important role of evidence from grammar acquisition within his linguistic conception without admitting that *grammatical* theories aim to incorporate psychological hypotheses about grammar acquisition. He says:

Concerning acquisition, evidence about nature and nurture showing that a language with a certain structure could or could not have been learnt by a person from the “primary

²³⁷ Katz (1981) p.238

linguistic data” is direct evidence for or against any theory that ascribes such a structure to a language that has been learnt by the person... But this does not make the theory of language psychological. It is one thing for a theory to have psychological *evidence*, it is another for it to *be* psychological. Any theory about anything can have evidence from anywhere.²³⁸

So Devitt wants to draw a distinction between a theory having psychological evidence and that theory being a psychological theory. The problem with Devitt’s view is that he has not provided us with a distinction of substance in the case at hand. The key question is whether the theories aim to determine the grammars that are actually acquired. If the answer is “Yes” then they have a theory of grammar acquisition as a goal. According to the view I defended in §2.3, it is not just the case that generative grammars *can* draw on evidence from grammar acquisition, rather they need to make hypotheses about grammar acquisition if they are to progress beyond mere descriptive coverage.

By drawing the distinction, Devitt thinks he can deny that it is an explanatory goal of grammatical theory to determine a grammar from amongst the possible DAGs on the basis of a theory of acquisition. He thinks he can deny this whilst also allowing that theories of acquisition are of “direct” relevance. Devitt maintains that “a grammar tells us absolutely nothing” about the facts in virtue of which speakers acquire the grammars that they do. But nevertheless, he suggests that generative grammars are constrained so that the hypothesised grammatical properties of a speaker’s language are ones he acquires.

Even supposing Devitt is right in the very general claim that “Any theory about anything can have evidence from anywhere”, that does not help him to make the substantive distinction he wants to make *in the case at hand*. If the very general claim is correct then the weather on Mars could be evidence for a grammatical theory and of course we do not want to hold that grammatical theory has as a goal a theory of the weather on Mars. But nothing follows from this very general claim about the role of theories of acquisition within generative grammar. The connection between a theory of grammar acquisition and a grammatical theory of a speaker’s language is clearly a lot closer than the connection between the weather on Mars

²³⁸ Devitt (unpublished ms. b) pp.9-10

and a grammatical theory. Devitt agrees with Chomsky that it is an important constraint on grammatical theories that they meet evidence from hypotheses about acquisition. Devitt requires some reason to resist the move to justify grammatical theories in terms of their explanatory adequacy. The only reason that Devitt has provided is that the latter goals would be psychological ones. But this is only a reason if, like Devitt, one already endorses a non-psychological conception of the goals of generative grammar. Without a further reason, it is unclear that there is any substance to Devitt's distinction and the consequent position on explanatory adequacy. After all, Devitt agrees with Chomsky that the theories must meet the psychological evidence or be deemed inadequate.

In the explanatory framework Chomsky proposes grammarians investigate the structure speakers intuit, imposing formal conditions on grammars that can generate these structures and proposing hypotheses about which of the possible grammars that generate the intuited structures has actually been acquired. The combinations of formal conditions and hypotheses about the grammars that are acquired can be empirically confirmed by their success in dealing with a range of natural languages. It is not merely the case that evidence from acquisition could be, or can be, evidence for grammatical theories because (as I argued in §2.3), the appeal to hypotheses about grammar acquisition is the central means of determining the actual grammar of the speaker's language.

Chomsky anticipated the kind of criticism of explanatory adequacy that Devitt offers. He thinks it apparent that the discussion over whether a theory of acquisition is a "necessary" part of linguistic theory" is "quite without substance".²³⁹ The substantive issue is whether grammarians should be content to formulate their theories with "little concern for justification" from psychological theories of UG and the way in which a grammar is selected from amongst the options UG makes available. The Chomskian alternative is that the grammarian intends to proceed from the study of facts about particular languages to explain the properties of the natural languages; those languages acquired in first language acquisition that humans speak.

²³⁹ Chomsky (1965) p.41

If the grammarian takes the former path, Chomsky argues that “since interest in justification has been abandoned” then “neither evidence nor argument (beyond minimal requirements of consistency) has any bearing on what the linguist presents as a linguistic description.”²⁴⁰ But if on the other hand, grammarians want to discover the actual grammar of the language a speaker is competent in, as Chomsky and Devitt do, then “he must concern himself with the problem of developing an explanatory theory of the form of grammar, since this provides one of the main tools for arriving at a descriptively adequate grammar in any particular case.”²⁴¹

For this reason Devitt’s position on explanatory adequacy does not constitute a genuine alternative to Chomsky’s. So I’m now going to consider Katz’s argument for his more radical position according to which we have two distinct and evidentially isolated theories: one of grammar and one of grammar acquisition.

Katz argues that Chomsky’s criteria of adequacy are question-begging against his Platonist conception because they include psychological objectives. Katz agrees with Chomsky that grammarians seek to explain the structural properties of natural languages and that an adequate grammar for a speaker’s language ought to explain linguistic intuitions. But Katz claims that the further condition of explanatory adequacy is question-begging as it can only be satisfied by an empirical theory of the speaker-hearer’s psychology: to meet explanatory adequacy we have to incorporate an empirical hypothesis about language acquisition.

Strictly speaking, Chomsky’s conditions of adequacy do not “beg” any questions as such because they reflect explanatory goals rather than substantive assumptions about any matter of fact. It does not beg a question to aim for theories that are not merely descriptively but also explanatorily adequate. But one can see what Katz has in mind. The inclusion of explanatory adequacy favours the psychological conception, yet Platonists claim to offer the right interpretation of generative grammars and Platonist theories need not be constrained to enumerate only those structural properties of languages acquirable by humans. So Katz rejects

²⁴⁰ *ibid.*

²⁴¹ *ibid.*

what he calls the “learnability constraint” which “imports” psychological hypotheses into generative grammar.²⁴²

Katz proposes an alternative set of conditions of adequacy on generative grammars which is intended to be “theoretically neutral” between competing conceptions and which he claims “conforms closely to our ordinary pre-theoretical conception of what we want a theory of language to be.”²⁴³ According to Katz’s criteria, a grammar G is an optimal grammar for a language L if and only if:

1. G implies every true evidence statement about L.
2. G provides the grammatical basis for explaining all grammatical phenomena in L that merit explanation.
3. There is no simpler grammar G’ which satisfies 1 and 2.²⁴⁴

Crucially, Katz classes speakers’ linguistic intuitions amongst the true evidence statements about L and thinks the intuited properties are amongst the “grammatical phenomena” that merit explanation. But for Katz, explaining speakers’ linguistic intuitions is not a psychological objective, because he thinks that the intuited properties that the grammarian ought to explain are properties of mind-independent, Platonic objects. The primary explanatory goal on Katz’s conception is to predict properties and relations, such as well-formedness and synonymy. A grammar predicts the fact that a sentence has a grammatical property just in case it defines the property and the structural description of the sentence by the grammar, together with the definition of the property, implies that the sentence has the property. Katz construes this as an explanation of logical and mathematical properties rather than psychological or causal properties.

Katz does see it as a goal of generative grammar to characterise UG. He thinks that a linguistic theory will be correct and complete only if it recursively enumerates the generative grammars for each natural language so that all the universal grammatical principles appear as clauses in the theory. But he construes

²⁴² Katz (1981) pp.53-4

²⁴³ Katz (1981) p.64

²⁴⁴ Katz (1981) pp.66-67

UG such that these universal facts are not facts about what competent speakers of languages are endowed with and determine the languages they acquire, but rather as grammatical facts that “in general merit explanation”.²⁴⁵ For Chomsky UG includes “those properties that must be assumed to be available to the child learning a language as...innate endowment.”²⁴⁶ For Katz the question of UG is “the non-psychological question of what are the essential properties of the languages that grammatical knowledge on the part of any intelligent creature is knowledge of.”²⁴⁷

Katz’s case against explanatory adequacy is weak. To do the argumentative work that Katz requires, his criteria ought to exclude considerations of explanatory adequacy. Katz himself does exclude the acquisition of grammar from the grammatical phenomena that “merit explanation”, but he does so on the grounds that acquisition is a psychological phenomenon. So it is unclear how Katz’s criteria are any more “theoretically neutral” than Chomsky’s if they must be interpreted so as to exclude explanatory adequacy on just such grounds. This is made apparent by the obvious availability of an interpretation of the criteria so as to admit grammar acquisition amongst the “grammatical phenomena”. Hence, the proposed criteria do not rule out explanatory adequacy unless interpreted from the outset so as to exclude psychological phenomena from amongst the “grammatical phenomena”. But this reasoning is entirely circular because it relies on excluding psychological phenomena rather than providing an independent basis for doing so.

What Katz wants is to establish a distinction between, on the one hand, grammatical phenomena and evidence that is directly relevant to generative grammar, and, on the other, psychological phenomena that are irrelevant. This is supposed to make palpable that the primary goal of generative grammar is to explain the “grammatical phenomena” and that any psychological import that the investigation has is merely derivative. But the only way to arrive at Katz’s conclusion on the basis of his criteria is to assume a proprietary sense of the *grammatical* or *linguistic* phenomena such that they are restricted to non-psychological phenomena. Thus, there are no “impartial rules of the game” that can

²⁴⁵ Katz (1981) p.66.

²⁴⁶ Chomsky and Halle (1968) p.4

²⁴⁷ Katz (1981) p.223

be extracted from Katz's criteria and which help us to decide between competing conceptions of the goals of generative grammar. Chomsky and Katz agree that generative grammar has descriptive adequacy as a target and that the linguistic intuitions of native speakers provide an important source of evidence for constructing DAGs. But there are no materials here to build an argument against the further justification that can be achieved by meeting the condition of explanatory adequacy that Chomsky motivates.

Suppose that there are a number of equally simple, but differently structured, DAGs for a speaker's language $g_1, g_2 \dots g_n$. They each offer structural descriptions of the sentences of a language consistent with a native speaker's intuitions. The grammarian could decide to be satisfied with the DAGs he has constructed. But he might want to know, further, which of the grammars characterises the language that the speaker actually knows. This interest will be particularly well-motivated if, as Chomsky claims, mere coverage of the data is not "in itself, an achievement of any particular theoretical interest".²⁴⁸

We might imagine that amongst the equally simple explanations of a speaker's linguistic intuitions, $g_1 \dots g_n$, there is one g_3 that contains three rules.²⁴⁹ Tomorrow the grammarian might discover a good reason to believe that speakers acquire the language in question in three distinct stages. In stage 1, the speaker recognises just those structures licensed by rule 1 of g_3 . Whilst in stage 2, he recognises just those structures licensed by rule 1 and rule 2 of g_3 . In stage 3, he recognises the full structure of the language. On the face of it, this provides us with evidence that g_3 rather than g_1 , or some arbitrarily selected grammar, g_{12} , is the grammar of the known language. We have evidence that it is the grammar acquired rule-by-rule: the language is acquired in three distinct stages, each seeming to correspond to the acquisition of a rule of the grammar.

The only move open to Katz is to deny that there is any evidence that we should opt for g_3 over the other DAGs. Katz would be making this move not because he has a saving hypothesis about how the language is acquired, or because he has a conflicting hypothesis about the speaker's grammar, but because he has

²⁴⁸ Chomsky (1965) p.26

²⁴⁹ This argument is due to Fodor (1985) pp.153-4.

placed an a priori limitation on the relevant data; placing the grammar that the speaker has actually acquired beyond the grammarian's sphere of interest. The theoretical interest is obviously a coherent one but one beyond Katz's proprietary sense of "grammatical phenomena".

Fodor suggests that we should be happy to grant Katz the term *linguistics* or *generative grammar* for the discipline that adheres to such explanatory limitations. But, as Fodor points out, "just down the road, there must be another science *just like linguistics* except that it *does* care about empirical truth because it cares about how the mind works."²⁵⁰ We might call this latter scientific enterprise '*generative grammar**', noting that it has the explanatory goals that Chomsky outlines for generative grammar. Generative grammar* aims for theories that do more than assign structural descriptions to all the sentences of a language. It aims to determine the actual grammar that a speaker knows, by characterising the structural properties of the possible human languages acquired by speakers in first language acquisition. It will be a condition of adequacy on a theory within generative grammar* that it offer an explanation of the intuition of the native speaker on the basis of an empirical hypothesis about language acquisition. Generative grammar* is inherently concerned with Katz's contingency: whether it is the case that "by some chance, one linguistic theory [coincides] to a greater extent than others with a psychological theory."

Katz claims that generative grammar* is an incorrect characterisation of generative grammar. He maintains that generative grammar has non-psychological goals and should be assessed according to non-psychological conditions of adequacy. So he needs to offer an account of the phenomena grammatical theories are explanatory of. Katz has to motivate the grammarian's appeal to properties of c-command, binding and the like as part of an explanation of some non-psychological phenomena. He cannot appeal to human cognition of these properties to motivate their interest, because citing our psychological capacities and their development would be to betray psychological goals. Consider sentence (9).

(9) Bill's brother loves himself.

²⁵⁰ Fodor (1985) p.160

As Collins puts the challenge to such conceptions, unless we are interested in what speaker's immediately and unreflectively cognise:

that Bill's brother c-commands himself is no more interesting a property of our example sentence than that the pairs <h, h> and <e, e> are cross-serial. The only conceivable reason to pick out c-command as opposed to any other conceivable property is that it, as opposed to an indefinite number of others, enters into an explanation of human cognition.²⁵¹

Those who oppose explanatory adequacy and avoid questions about the grammars that humans actually acquire have to offer reasons for the grammarian to focus on c-command, binding and the other properties that generative grammars appeal to as opposed to a range of other structural properties possessed by (9). Their reasons can't be that those properties are the ones that humans acquire and enter into an explanation of the cognition of language. But without appealing to these cognitive reasons, we need to know what makes such properties salient. What are salient to generative grammarians, looking at sentences like (9), are certain relations of command between constituents of specially categorised types. The generative grammarian is not so interested in many other properties and relations that can easily be found in linguistic material, and we could easily invent relations of structural command that are not the focus of generative grammar. There is no obvious reason, in choosing a domain for the theory, why grammarians need to attend to those properties of linguistic material that competent speakers of natural languages recognise, except for the grammarian's interest in the languages that speakers actually know.

Speaking of the putative linguist who eschews generative grammar*, Fodor says:

In principle, he might just as well attend to the construction of grammars that predict only intuitions about sentences with more than seven vowels, or sentences whose twelfth word is 'grandmother', or sentences that happen to be uttered on Tuesdays.²⁵²

²⁵¹ Collins (2007) p.4

²⁵² Fodor (1985) p.158

Fodor's thought is that linguistic material has a range of properties that generative linguists are clearly not interested in such as the number of letters in a sentence, or the relative number and position of vowels and consonants. Some of these are structural properties, such as the relative position of the sentences vowels and consonants. So there must be some Platonist suggestion about the target of explanation that motivates the selection of the properties that actually animate linguistic theory and according to which generative grammar is not targeting the language a speaker knows. As Collins remarks: "The very point of explanatory adequacy, as a condition on general linguistic theory, is to constrain the grammars speakers can know, to distinguish natural language from arbitrary symbol systems."²⁵³ If the Platonist cannot motivate the focus on these special properties then our No Violence Principle (§1.3) is flouted.

Perhaps the Platonist answer is that generative grammar just explains grammatical properties *period* - Katz's "grammatical phenomena that merit explanation" - and nothing more should be said. On this suggestion, properties like c-command predict grammatical properties like binding which linguistic structures have. This is correct, as far as it goes. Sentences have certain striking structural properties and the properties uncovered by the generative grammarian are explanatory thereof. There are two issues with this response, however, if it is being presented as an alternative to the psychological conception of the goals of generative grammar.

The first point is that it does not answer the question concerning the salience of the properties that grammarians focus upon so much as postpone it. The question for the Platonist was why certain grammatical properties such as c-command and binding were of special theoretical interest. These properties are of a different sort of interest to other properties that structures have, such as the number of vowels and consonants they contain, or properties that we could imagine or stipulate linguistic structures to have as they differ from those we immediately cognise. The proponent of the psychological conception says that we target these properties because they enter into an account of a range of cognitive phenomena centring on our immediate

²⁵³ Collins (2004) p.523

and unreflective cognition of language and how we acquired this competence. If the alternative is that grammatical theories aim to explain grammatical properties of linguistic structures, this is something that all parties (whether proponents of a psychological or Platonist conception) can agree with. But we wanted to know what makes these properties of linguistic material salient. Proponents of the psychological conception answer this question by conceiving of grammatical properties as those properties that the human faculty of language makes available to us.

The second point is that if generative grammar is understood to contain the theories offered within the generative enterprise then proponents of the psychological conception might claim with much plausibility that, de facto, generative grammar does not just explain the grammatical properties of linguistic structures *period*. Generative grammar offers some explanation of a range of cognitive phenomena and hence is pertinent to other questions “in particular those pertaining to the deeper systems that underlie behaviour.”²⁵⁴ Generative grammars aim to assign structural descriptions to an infinite range of sentences indicating how they are understood. Our linguistic judgements show evidence of productivity and special hierarchical structures. As such, generative grammars explain our linguistic intuitions, the productivity and the special organisation of the language we know. Like our generative grammar*, generative grammars will bear on empirical hypotheses about language acquisition, and the deep similarities between speakers of superficially different languages. Generative grammars are also part of the explanation of language use, insofar as an underlying knowledge of the hypothesised grammar is integrated with performance systems of perception and production.

The upshot of these two points is that the view that generative grammars explain the grammatical properties of linguistic structures *period* is uninformative and seems false if read prohibitively so as to exclude the psychological significance of grammar. The prohibition might serve to express a lack of interest in certain questions to do with natural language. But if Chomsky is right that descriptive adequacy by itself achieves “no particular theoretical purpose” then Katz’s proposal

²⁵⁴ Chomsky (1965) p.193 fn.1

amounts to a restriction of “generative grammar” to mere organisation of the intuitions data – the only data that Katz believes is relevant to the investigation of grammar. Generative grammarians, following Chomsky, typically ignore Katz’s restriction. The problem with the restricted conception of the goals of generative grammar that Katz is interested in may be, as Fodor once put it, that “deep down no one is remotely interested in it” because they are interested in the languages speakers actually acquire.²⁵⁵ The response I have considered on Katz’s behalf fails to motivate the linguist’s focus on the special properties of natural languages that speakers know and fails to accommodate the wider explanatory importance of generative grammar.

²⁵⁵ Fodor (1985) p.159

3. The Competence-Performance Distinction

In the last chapter, I examined the explanatory goals of generative grammar. I defended the view that the central explanatory target of grammatical theories is a generative grammar that meets conditions of descriptive and explanatory adequacy. In this chapter, I look at how grammarians investigate the generative grammars for speakers' languages. What grammarians require is a means to determine the grammatical properties of speakers' languages. The means suggested by Chomsky, and now widely adopted amongst generative grammarians, is to draw a fundamental distinction between a speaker's *knowledge* of a grammar, realised in a system of grammatical competence, and a speaker's *use* of the grammar he knows, which involves further systems of linguistic performance. This is a distinction between two aspects of a speaker's psychology. By seeking to isolate the speaker's grammatical competence from the extraneous performance factors, grammarians hope to get a much clearer reflection of the grammar of a speaker's language.

Theories of competence and theories of performance offer different kinds of explanation. As Fodor points out, competence theories look to account for facts about the organisation of a speaker's linguistic behaviour and capacities by reference to properties of his internalised grammar. Performance theories look to account for facts about a speaker's behaviour and capacities by reference to interactions between the internally realised grammar and other aspects of the speaker's psychology.²⁵⁶ For example, the linguist might look to explain a speaker's ability to understand or produce a novel linguistic form, embedding one sentence within another - as with *The man, the dog bit, died* - by reference to the productive operations of the speaker's grammatical competence. But linguists might want to explain the speaker's inability to understand centre-embeddings of an order of two (*The man, the dog, the cat scratched, bit, died*) or three (*The man, the dog, the cat, the wasp stung scratched, bit, died*) by reference to interactions between what the productive grammar allows and what non-grammatical resources, such as short term memory, the speaker employs in parsing the sentence.

²⁵⁶ Fodor (1985) pp.154-5

An inspection of speakers' abilities with respect to sentences like (10) is enough to suggest a difference between the grammars we know and our abilities to use those grammars in real-time comprehension and articulation.

(10) The man the cat the dog bit scratched died.

Sentence (10) strikes me, over and again, as lacking a full grammatical structure even once some prompting, extended attention or the explanations of others have led me to believe otherwise. After some concentration on (10), or some thoughtful prompting, I can recognise that it is a sentence of my language though I would never use it and struggle to understand it at normal speed. Once I recognise that (10) is a double centre-embedding, I can peel away the embeddings and its structure becomes apparent to me. It is structured so that *the dog bit the cat that scratched the man that died*. We can see this if we start with the sentence *the man died* and then embed the clause *the cat scratched*, which describes *the man*, to give us *the man the cat scratched died*. Then we can embed *the dog bit*, which describes *the cat*, into the embedded clause yielding (10). Running through that procedure as I read (10) I can keep its grammatical structure firmly in view but as soon as my attention lapses I lose that structure and (10) again strikes me as lacking it.

What's stopping me recognising the sentence as part of my language is the lack of attention and other resources required to process the embedding. In the case of (10), and a vast range of other cases, it is not the language I know that rules out the structure but the extraneous factors involved in using such grammatical information, in this instance to repeatedly centre-embed.²⁵⁷ The explanation on offer is that my use of language sometimes masks my standing knowledge of centre-embedded structure. The distinction is suggested by a wide range of phenomena concerning what one can immediately parse and what one can come to recognise with added performance resources such as time, extended attention or bracketing. But it is also suggested by such crude facts as that it's hard to see how I could speak, understand or judge of sentences of my language without such a body of standing information in place. Though I can decide what I want to say, I cannot decide to

²⁵⁷ I expand on this point in §3.3, §3.4.

speak and understand sentences of Arabic without drawing on some standing information that shapes my speech and understanding.

The key idea is that linguistic performances are the upshot of a number of factors amongst which grammar is only one. Grammarians name that aspect of the speaker responsible for the grammatical properties of performances, the system of grammatical competence. The competence system is considered as part of a *parser* for processing language that also includes distinct mechanisms for perceiving language. The perceptual mechanism works according to different principles to the grammar and the parser is subject to a variety of non-grammatical limitations. Insofar as a speaker's linguistic performances in uttering and parsing language obscure or deviate from the grammar they know, they constitute a kind of linguistic debris from the grammarian's perspective. As well as these systems for speaking and perceiving speech, grammatical competence is also integrated with conceptual-intentional systems for using language in thought. The *narrow* operation of the recursive grammatical competence, FLN (see §1.2), can be distinguished from the whole, *broad* package involving the performance mechanisms and systems for using language in thought, FLB.

In this chapter I argue that grammarians need a distinction between that aspect of the speaker's mind responsible for grammar, the competence system, and those other aspects of the mind that enter into the speaker's behaviour and capacities in order to determine the generative grammar of his language. On the basis that grammarians require a competence-performance distinction to investigate a speaker's grammar, I argue that it is theoretical apparatus that all conceptions of generative grammar must commit to. A conception of generative grammar that did not appeal to the competence-performance distinction to determine what amongst a speaker's behaviour and capacities reflects grammar and what reflects independent features of language use, and offered no other means to effect the disentanglement, would flout the No Violence Principle. For such a conception would have no resources to determine what amongst a record of speech, or amongst the intuitions data, is a reflection of the grammatical properties of a speaker's language and what is not. I thereby provide an argument that generative grammar is committed to a *psychological* distinction between different aspects of the mind/brain responsible for our linguistic behaviour and capacities, and that it attempts to characterise a

psychological system of grammatical competence. The fundamental importance of the competence-performance distinction suggests that grammatical theory is guided by psychological commitments and as such “is concerned with discovering a mental reality underlying actual behaviour.”²⁵⁸

3.1 Competence and Use

The notion of *competence* employed in generative grammar is not one that applies to our proficiency in using a language. Ordinarily, we talk about someone being competent at doing something or having the competence to perform a certain task. If we say that someone is a competent golfer, this is usually understood to be roughly synonymous with their playing golf well. Although, we might maintain that someone is a competent golfer when an injury or a bad set of clubs temporarily masks their abilities.

Possessing a grammatical competence, in the sense pertinent to generative grammar, is not meant to mark someone out as having an ability or doing something well. It is not meant to mark that someone is good at constructing and understanding grammatical sentences, for instance. Rather, in generative grammar, “competence” is used to designate a special psychological state: the structural organisation of the human mind that explains our potential for speaking and understanding language. Though possession of this psychological state, the “competence” system, is one of the enabling factors in our being competent users of a language, the ordinary notion of competent use is not what is being targeted.

In the linguist’s sense, two individuals might share exactly the same grammatical competence but differ greatly in their capacities to use it. And a person might increase their abilities to use their language whilst adding nothing to their grammatical competence. Not only in principle, but in real cases, subjects can possess this underlying system of knowledge whilst their ability to speak and understand language is impaired so that they are not proficient language users:

²⁵⁸ Chomsky (1965) p.4

Ability to use language may improve or decline without any change in knowledge. The ability may also be impaired selectively or in general, with no loss of knowledge, a fact that would become clear if injury leading to impairment recedes and the lost ability is recovered.²⁵⁹

This is because further psychological systems enter into the use of the information supplied by the competence system. With “competence” so conceived, a person who knows English but suffers cerebral damage - leaving the underlying knowledge intact but impairing their use of language in speech, comprehension or thought - retains their competence. We can imagine the effects of the brain damage receding and the person recovering the capacity to use their language. Though the person has no capacity to speak or understand in the interim, a certain mental structure that underlies speaking and understanding must have gone undamaged; unless the person re-learnt the language from scratch after the injury.

Such *aphasias* are real phenomena.²⁶⁰ Aphasia is the loss of the ability to speak or comprehend language due to damage to the brain areas responsible for these functions. The recovery of aphasics who lose the ability to speak and comprehend provides evidence for the integrity of the competence system, though none of the aphasic’s linguistic behaviour up to the point at which recovery begins evidences competent language use. It is this retained structure to the mind/brain that the grammarian wants to distinguish from the capacity to use language.

Further evidence for the distinct system designated as the competence system comes from the different modalities and mechanisms engaged in putting our knowledge of language to use in speaking and understanding. Pettito reports that grammar acquisition occurs in profoundly deaf children exposed only to sign, in hearing bilingual babies that acquire a sign language and a spoken language, and in

²⁵⁹ Chomsky (1986) p.9

²⁶⁰ There are well attested aphasias resulting from damaging to Broca’s area, which is responsible for the production of language, and to Wernicke’s area, which is responsible for language interpretation. Aphasia’s are not the result of general sensory, intellectual or psychological malfunction. Depending on the severity of the brain damage involved aphasics may be able to produce language but not comprehend it or even sing but not speak. Certain chronic neurological disorders such as epilepsy and severe migraines can also have transient aphasia as a symptom.

hearing children without any spoken input whatsoever, only signed input.²⁶¹ Hearing children exposed exclusively to sign exhibit normal development of a grammatical system, albeit that they put it to use in signing rather than speaking. They do so without using the brain's auditory and speech perception mechanisms and without the use of motor mechanisms involved in producing speech. They even babble using gestures. What this suggests is that this grammatical competence system is not primarily a set of mechanisms for producing and comprehending speech or sign. There is the same underlying grammatical competence across variations in the production mechanisms and behavioural capacities "all the while preserving linguistic structure across the [different] modalities."²⁶²

The charge might be levelled at the grammarian that he should not use the term "competence" if what he has in mind is a mental state and not an ability to exhibit a certain range of behaviours. The term "competence" was originally introduced to "avoid entanglement with the slew of problems relating to 'knowledge'".²⁶³ Chomsky has noted that "competence" is itself misleading insofar as it suggests an ability, and this is an association that he hopes to sever by pointing out that it is a technical term, encompassing:

all those aspects of form and meaning and their relation, including underlying structures that enter into that relation, which are properly assigned to the specific subsystem of the human mind that relates representations of form and meaning.²⁶⁴

Though the integration of the competence system with wider psychological systems for using language does support (roughly speaking) commonsense exercises of linguistic "competence", Chomsky could just as well talk about a component of I-language or invent some other technical term. The competence system clearly does bear on what the speaker can do, though it is not a characterisation of what he does. The key point is that the competence system shapes behaviour only insofar as it is integrated with systems for linguistic *performance*.

²⁶¹ Pettito (2005) pp.90-1

²⁶² Pettito (2005) p.92

²⁶³ Chomsky (1981) p.59

²⁶⁴ *ibid.*

3.2 Why Generative Grammar Needs a Competence-Performance Distinction

In Chapter Two, I alluded to the fact that both linguistic behaviour and linguistic intuitions are the upshot of an ensemble of distinctive underlying psychological systems. The problem that faces the linguist is how to focus on those aspects of linguistic intuitions and behaviours that reflect the grammar of a speaker's language. Stanley describes the linguist's predicament in the following way:

Ordinary discourse often involves the use of complex expressions which would be counted as ungrammatical even by the utterer's own lights. For example, some people regularly start a new sentence halfway through an utterance of another sentence... It is absurd to suppose we should count such discourse as grammatical, and thereby modify syntactic theory to account for it, and this despite its (statistically speaking) relative normalcy.²⁶⁵

It would be "absurd" because such discourse is clearly affected by factors other than the speaker's language; to incorporate all the factors that enter into a speaker's linguistic behaviour would expand the domain of linguistic theory indefinitely. The grammarian would be left with the thankless task of trying to explain all the properties of linguistic behaviour concurrently. For, presented with a record of speech, he would have no resources to distinguish the false starts, the deviations from rules, the changes of plan in mid-course or even the hiccups from the enduring linguistic forms that speakers recognise. So, it is a sensible theoretical move to separate out the factors responsible for linguistic behaviour so as to get a clearer reflection of the speaker's language.

The competence-performance distinction is just such a theoretical move to differentiate amongst the complex factors responsible for linguistic behaviour and judgement. As Chomsky puts it:

²⁶⁵ Stanley (2000) p.408

To study actual linguistic performance, we must consider a variety of factors, of which the underlying competence of the speaker-hearer is only one. In this respect, the study of language is no different from the study of other complex phenomena.²⁶⁶

Speaking very broadly, our use of language involves our pairing the sounds that we hear and utter with linguistic forms. So there is some core linguistic system, or ensemble of systems, responsible for integrating the two sets of structures. It is this system that is called the *grammatical competence* system. In order to isolate aspects of our use of our language in behaviour and judgement that reflect the grammatical forms of our language, linguists can investigate this pairing of sounds and linguistic forms that our grammatical competence licences. The linguist can then try to distinguish this project as sharply as is possible from the investigation of whatever other factors affect our use of these grammatical forms.

So understood, the competence-performance distinction is to be drawn *a posteriori*. The distinction is an inference to the best explanation of our linguistic judgements, as with our explanation of (10), and of the organisation of our behaviour. The linguist discovers which aspects of speakers' capacities and behaviours fall within the domain of a theory of grammatical competence, and which within the theory of performance, by fashioning the best explanations of the speaker's capacities and behaviours.²⁶⁷ Hiccups do not turn out to be part of the theory of grammatical competence, even though they occur within a record of speech, because they are produced by the interaction of the language mechanisms with other mechanisms that are known to operate independently of language. So an adequate theory would have to treat speech that contains hiccups as an interaction effect of the language systems and something else rather than attribute them to the speaker's grammatical competence.

It might be suggested that this explanatory commitment to the competence-performance distinction as the method to discern the grammatical properties of a speaker's language is overly hasty. The problem we started out with was that linguistic behaviour and intuition are complex interaction effects involving a range of factors other than grammar. It might be suggested that this only implies that

²⁶⁶ Chomsky (1965) p.4

²⁶⁷ Fodor (1985) pp.154-5

judgements, speech and comprehension are a *noisy* guide to grammatical properties. The fact that these phenomena are a noisy reflection of grammar, so the thought goes, is just the same situation as occurs in other sciences where the data are noisy. What is required of the linguist then is that he set aside the data that are too noisy, idealising away from the noise so as to get a clearer reflection of the underlying grammatical regularities.

This proposed alternative denies that the complexity of factors that enter into linguistic intuition and behaviour calls for the grammarian to develop a theory about the underlying grammatical competence system. The proposed alternative is that the complexity and noise only require the linguist to set aside seeming anomalies and to regiment the observed phenomena in such a way that grammatical generalisations can be extracted. On this alternative view, the grammarian need not theorise about an underlying grammatical system of the mind, as distinguished from performance factors, in order to reveal grammar. Rather the grammatical properties are discernible amongst the data once it is systematised and the noise is set to one side.

However, the grammarian who makes an explanatory commitment to the competence-performance distinction is not denying that the usual situation obtains, whereby there are aspects of the phenomena scientists observe that are relevant to their investigation but there is also noise. The purpose of making an explanatory commitment to the competence-performance distinction is to give some substance to the question of what amongst the observed linguistic phenomena reflects grammar: of what is relevant to grammatical investigation and what should be counted as “noise”. For any set of observations, from some explanatory perspective, there are properties of the observed phenomena that are noise. So, the grammarian needs some explanatory perspective from which to ascertain what amongst the observed phenomena reflects grammar, the underlying pairings of sounds and linguistic forms, and what is noise. The competence theory commits the grammarian to there being an aspect of the speaker responsible for the sound-linguistic form pairings, and by investigating that aspect of the speaker we determine what in his behaviour and capacities reflects his grammar. We can fix up messes, regimenting our observations, in any number of ways. But not all

systematisations of the observed phenomena will be revealing of the grammars of speakers' languages.

Devitt's view is that grammarians should study properties of outward linguistic behaviours, such as utterances, the "outputs" of the psychological systems for language, rather than study an underlying system of grammatical competence. He recognises that these linguistic outputs are not an unproblematic reflection of grammatical properties. Accordingly, he suggests that the grammarian is only concerned with a subset of the outputs. Just as "sometimes what a blacksmith produces is not a good horseshoe" so Devitt reasons that sometimes the psychological systems for language will produce utterances that are not well-structured grammatically. What he suggests is that "[t]he theory is only concerned with the nature of the outputs of a competence when it performs as it should" and "performance errors" should be "set aside".²⁶⁸

The weakness in Devitt's suggestion, considered as an alternative to the competence-performance distinction, is that he clearly uses the theoretical notions of competence and performance to discern amongst the relevant properties of the outputs. Devitt says:

[Chomsky] draws attention to the fact that many factors can enter into any performance other than underlying competence: there can be "noise", limitations of memory, etc. The theory of competence abstracts from that. And so does the theory of expressions that the Respect Constraint distinguishes from the theory of competence.²⁶⁹

But the only means Devitt offers to sort amongst properties of outputs, the outputs that result when the competence "performs as it should" and the "performance errors" that should be set aside, is the competence-performance distinction. His "theory of expressions" abstracts from "performance errors" that enter into the outputs. But to do that one needs a hypothesis about which factors enter into the performance events, his "expressions": about which aspects of utterances are the reflection of performance systems and which reveal grammar. Once one has made this distinction amongst the factors that enter into the etiology of outputs, i.e.

²⁶⁸ Devitt (2006) p.18

²⁶⁹ Devitt (unpublished ms.) p.16

between what is a reflection of performance systems (Devitt's "performance errors") and what is a reflection of the speaker's possession of a grammatical system, *then*, and only then, can one go about abstracting from performance factors.

Contrary to what Devitt supposes, our felt grasp of which behavioural phenomena are a reflection of "performance errors" and which a reflection of a speaker's grammar will only take us so far. In order to discern in an empirically motivated way what amongst the phenomena constitutes "performance error", we require an empirical theory of performance, as this is distinguished from a theory of the grammatical system of competence. So Devitt is anyway relying on a theory of grammatical competence and its interaction with performance systems. The notion of *grammar* is a highly theoretical one, and ultimately our theory of what a speaker's grammatical system licenses is going to inform our conception of what phenomena fall within the scope of grammatical theory and which are Devitt's "performance errors". A theoretical focus on the competence-performance distinction is an attempt to get some empirical purchase on what amongst the data reflects what; on what is a reflection of a speaker's grammar and what is just "noise" from the grammarian's perspective.

So the question of what is "messy data" or "noise", from the point of view of grammatical theory, is non-trivial. Devitt wants us to distinguish between the rules that characterise a speaker's grammatical competence and the structure rules "respected" by that competence. But, as should be apparent, adopting the competence-performance methodology, we find out about the rules that characterise a speaker's language by theorising about their grammatical competence and how it informs their linguistic performances. Devitt himself is explicitly committed to a theory of performance in appealing to "performance errors", and so, in my view, he tacitly supposes, the competence-performance distinction and a conception of grammatical competence as that which the performance factors are to be distinguished from. His "theory of expressions" relies on our finding out what constitutes "performance error" and what reflects the rules of a speaker's language. But the only way we know how to do that is to find out how the grammar which a speaker has internalised interacts with the performance mechanisms involved in language use.

Chomsky suggests that grammatical theory is primarily concerned with:

[A]n ideal speaker...who knows its language perfectly and is unaffected by such grammatically irrelevant conditions as memory limitations, distractions, shifts of attention and interest and errors (random or characteristic) in applying his knowledge of language in actual performance.²⁷⁰

So conceived, grammatical theory focuses on the language a speaker knows. As Fodor notes, it is the competence-performance distinction that Chomsky draws upon to introduce the idea of an idealised speaker.²⁷¹ Chomsky is imagining what a real speaker would be like if the grammatical system engaged in his behaviour was not constrained by other psychological states and processes. Of course, without the further states and processes there would not be any behavioural output at all. But we might imagine a speaker with infinite time, attention and short-term memory and so on. Then assuming, artificially, a theory of processing that appealed to only these three factors, we might suppose that the speaker would be able to interpret centre-embeddings of an arbitrary number, always tell whether a string, however long, was grammatical and so on. The theory of the ideal speaker is then really a theory of grammatical competence idealising away from the interaction of the competence system with other psychological factors:

We can even add this: grammars per se are theories about Ideal Speaker-Hearers. This sounds wildly deep and ontological and sexy, but actually it is trivial and harmless. All it means is that grammars are not, per se, theories of the interaction effects.²⁷²

The questions of fact are whether this broad division of theoretical labour is an explanatory one, and, if so, what the correct division of explanatory labour between grammatical competence and the other systems really is.

But a competence theory doesn't just abstract from the performance limitations. The theories of competence and performance appeal to different systems which exhibit different principles. Competence Theories do not ignore

²⁷⁰ Chomsky (1965) p.3

²⁷¹ Fodor (1985) p.155

²⁷² Fodor (1985) p.156

these performance factors or simply idealise performance. They are supposed to gel in some way with theories of the different performance systems insofar as competence is one of the factors that, along with linguistic performance systems and other cognitive factors, offer some explanation of our language use and judgements. Competence is not just idealised performance: when we leave performance *limitations* out of the picture this is really to idealise performance not to offer a theory of grammatical competence. The competence system generates linguistic forms that cannot be simply read off performance events, however idealised, because the performance systems work according to different principles. As we saw with (10), linguists need both sorts of principles to explain the data.²⁷³

Chomsky initially conceived of his “ideal speaker” as situated within a “homogenous speech community”.²⁷⁴ One might wonder whether the speaker’s place within such a community, rather than his individual competence, might help the grammarian determine what the grammar of his language is. The idea might be that the aspects of a speaker’s judgements and behaviours relevant to the grammarian are those that obey the grammatical principles of the speech community, whilst the rest, where the speaker is out-of-step with his community, is just noise.

Chomsky never foresaw the notion of a speech community playing this theoretical role. It is noteworthy that he has subsequently given up altogether on the idea that the notion is of any use to grammatical theory, adopting a thoroughgoing individualist perspective on grammatical theory (§1.2).²⁷⁵ Chomsky originally employed the notion of a homogeneous speech community as to abstract away from the differences between speakers of relevantly similar languages. The theory of a speaker’s competence was never informed by a theory of his membership of a speech community. Rather the object of inquiry has always been any individual speaker’s endowment for grammar and the way that their mature grammar is determined by a course of experience. Though grammarians talk about English-

²⁷³ I develop this point in §3.3, §3.4.

²⁷⁴ Chomsky (1965) p.3

²⁷⁵ Chomsky (1987) says: “The notion of ‘community’, in fact, requires clarification that has never been given.”

speakers, French-speakers and so forth, these communities are abstractions from the competences of collections of more or less similar individual speakers. To work out the grammatical principles that characterise the language spoken within a speech community, one needs to investigate the principles that characterise the grammars of its individual speakers.

Chomsky's view is that *languages* such as English, and speech communities such as the English-speaking community, are not only more abstract but also less well-defined than the grammatical competences of individual speakers. It is a well known quip of Max Weinrich's in linguistics that languages are dialects with armies and navies; interest-relative demarcations made according to historical, sociological or power interests.²⁷⁶ Much the same view is taken of linguistic communities. If one wanted to use the idea that there were grammatical rules that characterised the English-speaking community in generative grammar then one would first have to draw a boundary around English-speaking communities. The widely held view is that there is little empirical substance behind any such boundaries beyond the unity provided by similarities in speakers' grammatical competences. Moreover, it is widely assumed that there is nothing to be learnt about grammatical structure by drawing such boundaries.

For instance, a marginally different set of rules may underlie the English that one hears in the Highlands of Scotland, Glasgow, Newcastle and South London.²⁷⁷ We could classify all of the speakers in these regions as members of the same speech community, even if some of them are mutually unintelligible, or we could say that these marginal differences are sufficient to distinguish amongst the dialects giving us three speech communities. We would not have learnt anything about the grammars of speakers' languages by doing so.²⁷⁸ It is well known that

²⁷⁶ Compare Chomsky (1987 p.170): "[My speech community] is taken to be the community of speaker of 'my language' but here 'language' is understood in the unanalysed sense of ordinary usage, with its socio-political and normative-teleological dimensions, a notion that is obscure and in any event has no status in an inquiry into language and the use of language."

²⁷⁷ Kayne (2000 p.7) suggests that: "it is entirely likely that no two speakers of English have exactly the same syntactic judgements."

²⁷⁸ In this vein, Chomsky (1987 p.170) says: "Thus is George, or Kripke, or a speaker of Oxford English, or of Cockney, or of German, etc., a member [of my speech community]? There is no

grammatical differences in idiolects and dialects amongst the speakers of one “language”, such as Chinese, can outstrip the differences between dialects of different languages, such as Dutch and German. There are speakers of dialects of Dutch and German that are mutually intelligible whilst there are speakers of dialects of Chinese that are mutually unintelligible. Settling the boundaries between these commonsensical languages and the respective speech communities would not help us in investigating linguistic form because the answers to such questions seem to be a matter of decision or perspective, arbitrary from the grammarian’s point of view.

3.3 Grammatical Competence and Linguistic Processing

Theories of grammatical competence attempt to specify a function that our production and consumption of language meets. Our production and consumption of language meet this function insofar as the function yields the sound-linguistic form pairings that are partially explanatory of our linguistic judgements and behaviour. But such models of our grammatical competence are to be contrasted with *processing* models because they are not attempts to specify *how*, by what effective real-time procedure, we process linguistic material. Chomsky has always been clear that:

When we say that a sentence has a certain derivation with respect to a particular generative grammar, we say nothing about how the speaker or hearer might proceed in some practical or efficient way, to construct such a derivation. These questions belong to the theory of language use – the theory of performance... No doubt a reasonable model of language use will incorporate, as a basic component, the generative grammar that expresses the speaker-hearer’s knowledge of the language; but this generative grammar does not, in itself, prescribe the character or functioning of a perceptual model or a model of speech production.²⁷⁹

known way to answer these questions in any useful manner; note that the issue is not whether the terms are usable in ordinary informal discourse – of course they are – but whether they can be made clear in a way that will contribute to the effort to understand language and the use of language.”

²⁷⁹ Chomsky (1965) p.9. Radford (1988 p.132) offers the following analogy: “Municipal regulations specify certain structural conditions that houses must meet...What they do not do is tell you *how* to

Theories of grammatical competence investigate a structured state of the human mind/brain, its growth and maturation. One might think of this state as a more overarching structure to the mind/brain than the mechanisms effective in speaking and understanding. The grammarian imputes existence to this abstract structure of the brain in “the same sense as we impute existence to a program that we believe to be somehow represented in a computer”, hoping to find further evidence about the physical states that implement the program.²⁸⁰ The investigation of these states is informed by the more abstract study and, in principle, the converse is true.

There are phenomena that pervade language processing that do not directly concern the grammarian. Though people sometimes plan their utterances to some extent, typically they don't fully formulate sentences before they begin uttering them. This is evident from the existence of *disfluences* such as “uh...ums”, repetitions and false starts. People often say things like:

(11) Because...you see I, uh, didn't want to go because, uumm, I don't really like it there.

There are aspects of (11) such as the pause after *Because* and the “uh...uummm” sounds that are processing effects, which are not of direct relevance to linguistic structure. But there are processing effects where grammatical structure does look to be directly relevant. For example, agreement errors are more common after embedded that-clauses. Errors like (12) are significantly less commonly produced than errors like (13):

(12) *The claim about newborn babies were rejected.

(13) *The claim that the wolves had raised the newborn babies were rejected.

go about building a house...Phrase structure rules should be thought of as analogous to municipal building regulations: they lay down certain structural conditions which sentences must meet.”

²⁸⁰ Chomsky (1979) p.3

The relation between grammatical competence and parsing/production mechanisms is a huge area of empirical research. There exist a very wide range of suggestions about the role of competence theories in understanding language processing.²⁸¹ To take an illustrative example of the issues involved, some early work sought to build transformational grammar directly into the parsing mechanism. This was called the *derivational theory of complexity*: the theory that the processing complexity of sentences increases as a function of the number of grammatical transformations involved in their derivation. There were some initially promising results suggesting a correlation between the amount of memory required to process a sentence and the number of transformations involved.²⁸² The proposed deep structure for sentences like (14) had three occurrences of *the detective* whilst the proposed deep structure for (15) had only two:

(14) The governor asked the detective to cease drinking.

(15) The governor asked the detective to prevent drinking.

And in recall experiments, *the detective* was significantly more effective in prompting people to remember (14) than (15) which seems to correlate the higher number of transformations into surface structure with the memory resources drawn upon. But later work undermined the theory.²⁸³ The derivational theory of complexity predicts that subjects should find (17) and (18) more complex to process than (16) because they involve extra deletion transformations.

(16) Pat swam faster than Chris swam.

²⁸¹ See Neeleman and van de Koot (2008) for a recent discussion of the issues. They argue the grammatical system should not be understood as a knowledge base *consulted* by the performance systems in language use but rather as a more abstract description of the structure of the brain (at a cognitive level) distinguished from language computation. For a discussion of some of the wider explanatory issues implicated in whether we see the competence-performance distinction as a distinction between two (Marrian) levels of description of a single system or two distinct but interacting systems on one level, see the debate between Franks (1995) and Patterson (1998).

²⁸² Miller (1962)

²⁸³ Fodor, Bever and Garrett (1974)

(17) Pat swam faster than Chris did.

(18) Pat swam faster than Chris.

But in fact subjects process (16) more slowly than (17) or (18). This indicates a more complex relation between grammatical competence and the processing mechanisms than the derivational theory of complexity suggests.

Ultimately, we want to know what the relation is between a subject's competence, the structured information they possess, and their use of language in speaking and understanding. As Garrett puts it, we want to know "what relation holds between formal theories of grammatical structure and theories of real-time computational processes that underlie human language use."²⁸⁴ There are some broad features of the distinction that I want to press without pre-judging the empirical issues too much.

The fact that the competence system and the processing mechanisms work according to different principles can be illustrated by appeal to a few famous examples. We struggle to process doubly centre-embedded structures like (10) that we can come to recognise as part of our language. But self-embedding contributes far more radically to our inability to parse structures than applications of grammatical principles of similar complexity like nesting. Consider that sentence (19) is fairly comprehensible:

(19) Anyone who feels that, if so many more students whom we have not actually admitted are sitting on the course than ones we have that the room had to be changed, then probably auditors will have to be excluded, is likely to agree that the curriculum needs revision.

Sentence (19) contains six nestings but no self-embeddings, and compares favourably in how easy it is to parse with (10) where two centre-embeddings cause us a lot of trouble.²⁸⁵ We often fail to process such grammatical "garden path"

²⁸⁴ Garrett (1990) p.139

²⁸⁵ Chomsky (1965 p.196 fn.6) suggests that these phenomena provide evidence about the organisation of the parser beyond the triviality that it has finite available memory. The problem with

structures as *The horse raced past the barn fell*, where *raced past the barn* is a relative clause. And we systematically misconstrue such “depth charge” sentences as *No head injury is too trivial to ignore*, which consideration reveals to mean that however trivial a head injury is it should be ignored, rather than that that no head injury, however trivial, should be ignored. Speakers can recognise these special processing effects on what we say and hear, when we attend to them carefully, often appreciating fairly immediately upon intuitive consideration what we never notice in production and consumption.

When speakers produce or parse speech in ways that diverge from what their grammatical competence licenses, they are often then able to come to the right arrangements of words, rearranging them to form an acceptable sentence. Such judgements are not merely a matter of firing up the processing system again, running it as before and seeing what comes out. Something, a body of information, guides a speaker’s judgement of a more acceptable result.²⁸⁶

How might a grammatical principle like Principle A act as a constraint on linguistic processing? We know that by and large speakers do not produce sentences with reflexives that lack local antecedents and that if the parser tries to construe a reflexive as referentially bound by a non-local antecedent then this violates Principle A and the proposed structure will strike us as deviant in some way. We can observe quite radical divergences from what competence licenses when speakers are inattentive, extremely tired or suffer strokes that damage their ability to speak and understand. But we can also get good evidence for such principled constraints on what we produce and comprehend which are reflected in our sense of what constitutes a good sentence of our language.

These questions are obviously worth pursuing as we have real phenomena here that require an explanation. Smith argues that the explanation will require a two-tier architecture.²⁸⁷ His own tentative suggestion about that architecture is that we might think of the parser as a “fast and dirty” connectionist system trained up on

self-embedding might suggest that the perceptual device has a stock of procedures available to it, and that it finds it hard to execute a procedure P whilst it is already in the course of executing that same P.

²⁸⁶ See Smith (unpublished ms.) p.13 for discussion.

²⁸⁷ Smith (unpublished ms.) p.14

the parameterised principles of a grammar, which provide “templates” according to which parses can be checked and the parser can “correct or fine tune its outputs”. He offers the comparison of the process of shaking something through a sieve:

Were we to shake a jar of lexical items through a sieve – the templates...if a selection violates at least one of the principles, or if at least one lexical item fails to have its requirements met, then it fails to be a grammatical structure. On the other hand, if a sequence passes through the sieve untrapped it is a grammatical structure... What matters is that the expressions the quick and dirty processor puts together get checked against the conditions imposed by the grammar. In this sense, the principles can have an influence on output and input without being part of the processing.²⁸⁸

So understood, the competence does not produce outputs in any stage of a process but the generative procedure encodes structural constraints against which possible outputs of processing might be compared. One virtue of this sort of model is that it suggests how the competence system (at one level) *constrains* the performance systems (at another), whilst allowing that processing is governed by different rules to the competence.²⁸⁹

Devitt’s “minimal position on psychological reality”, (M), stated that all a grammatical theory tells us about grammatical competence is that there is something inside a speaker “we know not what” that is responsible for *producing* outputs. This something “we-know-not-what” is, on Devitt’s account, the competence system which is governed by a set of processing rules that respect the grammatical rules of a language. We can now see that this is to mistake the theory of competence for a processing theory, for the grammatical competence system does not produce outputs, at least not of the externalised sort Devitt suggests. The competence system shapes our performance but does not serve to process language and produce anything like utterances or parses, this being a performance matter. If we think of the structured state of grammatical competence as having “outputs”, in an extenuated sense, then these are the structured LF and PF representations. LF

²⁸⁸ Smith (unpublished ms.) p.17

²⁸⁹ Though see Neeleman and van de Koot (2008) for a critique of similar architectural proposals.

and PF are internal, mental structures not Devitt's outputs, which are performance events.

This distinction between the competence system that encodes grammatical structure and the processors is set out in Collins' Revised Distinctions I – III (§1.3) which distinguish a theory of the language faculty, revealing of grammatical structure, from a theory of the processors involved in performance events. The structures encoded by a speaker's grammatical competence, and targeted by generative grammarians, are distinct from processing mechanisms. Hence, a theory of competence is not a processing theory. But however the brain does process linguistic material, we expect it to be constrained by the conditions that our competence theory specifies, insofar as we take that competence theory to be part of the best explanation of our intuitive judgements and linguistic behaviour.

Hence, it is perfectly possible for Chomsky to adhere to Devitt's distinctions I – III that distinguish the investigation of grammatical structure from the investigation of language processing whilst maintaining that generative grammar involves investigating grammatical competence. Chomsky urges us to distinguish the study of the grammatical categories and relations to which the competence system is sensitive from the rules of language processing. With the competence-performance distinction in place, Devitt's distinction between a theory of the structure rules of a language and a theory of the processing rules involved in outputting utterances could be a distinction amongst theories of the mind/brain, marking a distinction that animates the psychological conception he is attacking.²⁹⁰

²⁹⁰ This is why Slezak (unpublished ms. p.12) claims that "Devitt is merely restating Chomsky's own competence/performance distinction as though he has advanced some novel insight." Slezak asks us to note the similarity of the following passages. Devitt (2006 p.38) says that it is not sufficient for knowing the processing rules that we know "there is something-we-know-not-what within a speaker that respects the rules of her language...We would like to go beyond these minimal claims." Whilst Chomsky (1986 p.197) says "we are keeping to abstract conditions that unknown mechanisms must meet. We might go on to suggest actual mechanisms, but we know that it would be pointless to do so in the present stage of our ignorance concerning the functioning of the brain." Both think that grammatical theory specifies abstract conditions that unknown mechanisms meet. This is Devitt's respect constraint and Chomsky's point about unknown ways in which grammatical principles are met by mechanisms. However, Slezak is wrong to suggest that Devitt's distinctions must be Chomsky's restated. Devitt thinks that grammarians theorise about the *environmental* outputs of

So if Devitt's view is that "a theory of a competence must posit processing rules that respect the structure rules of the outputs." then this is either an error or Devitt is using "competence" in a non-standard way to include grammatical competence and the processing mechanisms.²⁹¹ As Smith has pointed out:

Confusion arises...because Devitt glosses the claim that a grammar is about a speaker's linguistic competence as the claim that the grammar is about the processing system responsible for the speaker's performance. And this is clearly not what Chomsky means by 'competence'.²⁹²

The fact that Devitt is adopting a non-standard conception of grammatical competence is suggested by the following remark:

What more is there to linguistic competence than the *state* of embodying the remaining processing rules that are centre-stage in performance? If the grammar is not supposed to be about that state, what other mental state *could* it be supposed to be about?²⁹³

The answer to Devitt's first question is that the issue is about the role competence plays in processing not about whether the competence system "embodies" processing rules. An answer to Devitt's second question is that the competence theory is about a structured state of the mind/brain organised by parameterised principles. The theory of grammatical competence specifies what constraints the mind/brain of a competent language user realises. It internally determines conditions on the processes. We know that there are sounds and linguistic forms, and that the mind/brain serves to pair these structures recursively and over an unbounded range. The theory of grammatical competence investigates these systematic pairings, without specifying how the brain actually operates so as to solve this integration problem effectively. But the competence theory states precise

linguistic cognition distinct from psychology. Whilst Chomsky argues that we need a distinction between *psychological* systems of competence and performance to determine the grammars of languages, with the competence system specifying the structure of the mind responsible for grammar.

²⁹¹ Devitt (unpublished ms.) p.23

²⁹² Smith (unpublished ms.) p.8

²⁹³ Devitt (unpublished ms.) p.10

structural conditions, abstracted from the brain's behaviour, which any theory of what the brain does will have to take on board. There is no reason to take seriously the idea that the grammatical structure our competence system licences is a kind of epiphenomenon compared to processing, because grammatical structure is evident in our judgements and behaviour.²⁹⁴ The principles of this state are an inference to the best explanation of intuitions and the organisation of behaviour, as is the fact that this state is not the processor though it plays some role with respect to processing. We need both the competence system and the performance systems for the best explanation of judgements, production and comprehension. Moreover, grammarians have discovered a lot about the grammatical structures speakers acquire. So there is no plausibility to the claim that grammatical competence is more mysterious or less easily explicable than processing. Beyond these broad features of the competence-performance distinction which suffice to answer Devitt's rhetorical questions, the issues are to be answered by the best empirical theories of the distinction and of the respective contributions of competence and performance.

Though Chomsky holds that generative grammar is a part of psychology, he has never professed that grammatical theory is an investigation of psychological processing as such. Chomsky distinguishes theory of the grammatical structures recognised by competent speakers from theories of psychological processing whilst holding that *both* describe an aspect of our psychology. In a sense, Chomsky's aspirations have always been more modest than those Devitt attributes him. Even if we knew all the facts about the grammatical competence or information that speakers possess, we still wouldn't know exactly how they put it to use in speaking and understanding, how it is engaged in processing, and outside of these "online" processes in judgement.

Devitt could claim that investigating such grammatical structure to the mind/brain is not really a part of psychology. But this is highly problematic, for whatever theory of language processing is offered, we seem to require a theory of

²⁹⁴ Similarly, Neeleman and van de Koot (2008) conclude that the higher degree of abstraction involved in the theory of grammar as compared with the parser is no reason to question the reality of the grammatical code.

the grammatical structure to which the mind/brain is sensitive, more abstract conditions which the processing events meet and which explain our judgements. These conditions are the special recursive and hierarchical principles of our grammatical system. Such a theory is just what the theory of grammatical competence, required to determine the grammar of a speaker's language, provides us with.

Accounts of processing are informed by this account of grammatical structure. But one could maintain that only processing theories, and not theories of the structural constraints to which competent speakers are sensitive, constitute genuine psychology. However, one then has to defend an a priori conception of psychology as extending only to processing mechanisms, and not to states of knowledge and more over-arching structure. Moreover, if this is Devitt's view, then he is not attacking the more modest psychological conception that Chomsky defends. And *prima facie*, this would be to ignore an important class of cognitive states. Principle A characterises a state of grammatical competence insofar as *speakers are sensitive* to reflexive structure. But it is not a processing rule telling us how the constraint is effectively adhered to by the brain in parsing or producing sentences. It is sufficient for theories of grammatical competence to characterise an aspect of speaker's psychology that their acquisition of such a grammatical competence is part of the best explanation of such intuitive judgements and our adherence to such constraints in our behaviour.

Devitt's view is that generative grammarians impute "psychological reality" to generative grammar without sufficient explicitly psychological evidence. Though the phenomena I have discussed in this chapter would serve to falsify Devitt's claim, his claim suggests he has misunderstood the methodology of generative grammar at a more fundamental level. Generative grammarians investigate a speaker's system of grammatical competence *in order to determine* what the grammar of their language is. Without the insight into a speaker's grammar that our theoretical conception of the competence-performance distinction provides, we have no generative grammar to impute anything to. To suggest there is a prior question about the "psychological reality" of the grammarian's theories is misleading, for prior to developing a theory of the speaker's grammatical competence there is no generative grammar to impute any sort of reality to or to legitimate with evidence.

To summarise, generative grammarians, drawing on the competence-performance distinction, aim to tell us what categories and relations the mind/brain is sensitive to as it is responsible for linguistic production, consumption and judgement. These are abstract conditions that unknown mechanisms meet but revealing explanations can be framed at this abstract level. It is a further project to determine how exactly these conditions constrain processing and to determine the principles of biophysical organisation that allow us to speak and understand, of which we are currently ignorant.²⁹⁵

²⁹⁵ It is also noteworthy that there are other non-grammatical systems beyond the grammar, the parser and the production mechanism that are engaged in linguistic performances. Pragmatics aims for an explicit account of the further sub-personal systems involved in utterance interpretation (for an influential relevance-theoretic account see Sperber and Wilson (1986)). Sperber and Wilson (1996 p.461) suggest that amongst its fundamental tasks pragmatics must “explain how hearers resolve ambiguities, complete elliptical or otherwise semantically incomplete sentences, identify intended references, identify illocutionary force...and recover implicit import.” These are just some of the ways in which context-independent grammatical structures fall short of determining the interpretation of utterances in context. Like the parser, the pragmatic systems are generally considered from the perspective of real-time performance systems subject to non-linguistic constraints. But unlike the grammatical parser they are sensitive to all sorts of non-linguistic stimuli. See Carston (2000) for an illuminating discussion of the relation between generative grammars and pragmatic theories. Pragmatic theories are usually considered as cognitive theories that attempt to account for our performance of these extra-linguistic tasks in interpretation, building on the structures our grammatical system makes available. In some of Chomsky’s remarks on pragmatic systems (see Chomsky 1980), he talks about a pragmatic competence as a system of knowledge pertaining to conditions of the appropriate use of lexical and conceptual resources, and integrated with grammatical competence. What Chomsky suggests is a distinct kind of pragmatic system to the kind of processing systems that much of pragmatic theory focuses on (but see Horn 1988 and Prince 1988 for discussion of pragmatic competence systems). It is worth remarking that the role that results from grammatical theory have been afforded in these pragmatic investigations is in keeping with the idea that generative grammar targets a system of grammatical competence integrated with wider performance systems.

3.5 Katz's Case against "Competencism"

Amongst the conditions that a grammatical theory must meet are that it includes a description and explanation of all the grammatical properties of the sentences of natural language. Katz argues that a lack of abstraction from the psychological states of speakers means that a theory of grammatical competence will fail to meet these conditions on an adequate grammatical theory. Katz concludes that we should reject what he calls *Competencism*: the view that generative grammar should aim to characterise grammatical competence. Katz's thought is that although a theory of competence can accommodate the properties of the sentences that we might produce and recognise, it will make false predictions about the properties of sentences that lie beyond our psychological capacities for production and recognition.

Katz claims that, even if we set aside constraints on real-time performance, only the sentences that it is within our human capacities to produce or hear will be structured in the way that a psychological theory of competence predicts. He thinks that the competence theory will falsely predict facts about those sentences that it is beyond our psychological capacities ever to produce or hear because they are too incredibly long or complex. Katz says:

[T]he internalised rules might convert all strings of words above a certain very great length, n , into word-salad, so that the best theory of the speaker's competence falsely predicts that strings of English words exceeding length n are ungrammatical.²⁹⁶

What Katz is suggesting is that if the competence system might fail to accommodate a significant range of the sentences of natural language, for instance, turning all sentences over a certain length into word-salad, then it would be a mistake to make this a feature of grammatical theory. Given this possibility, Katz argues that we should understand grammatical rules not as the internalised rules governing a system of grammatical competence, which is Chomsky's view, but as rules defined over Platonic objects (see §1.3.3).

²⁹⁶ Katz (1985) pp.198-9

Though Katz professes to be setting performance issues aside in considering competence theories, in my view he is conflating an account of the structure of the mentally realised grammatical system with an account of what humans have the psychological capacities to utter and recognise. One consequence of the distinction between grammatical competence and the performance systems involved in putting that competence to use is that there can be grammatically structured expressions generated by the competence system that are completely unusable in performance. This is ensured by the recursive nature of the competence system which is structured so as to generate an unbounded number of sentences of indefinite length and the limited resources available to the performance systems involved in uttering and parsing sentences.²⁹⁷

The best theories of grammatical competence are structured so that the system licences an unbounded number of expressions without limits on their length or complexity. In doing so the grammarian allows that there are an infinite number of grammatical expressions that meet the grammatical constraints determined by a speaker's grammatical competence but which he will never speak or hear, their being unusable by him, and which the speaker would find unacceptable as sentences of his language. What the recursive principles, which ensure this result, achieve for the grammarian is an explanation of the special productivities distinctive of human language.

As Stich notes, one might think that we ought to restrict the account of what our grammatical competence licenses by "simply cutting off the class of sentences generated by a grammar at an appropriately high point."²⁹⁸ But there is no natural point to draw such lines, for there is no point at which the addition of another

²⁹⁷ Equally, there are usable forms that are not grammatical. Perfectly good meanings can sometimes be expressed by linguistic utterances that are not structures that the competence system licences, as with (a) and (b).

(a) He seems sleeping.

(b) No head injury is too trivial to ignore.

A speaker might assign an interpretation to (a) though it is not a full grammatical structure of their language and might assign a rather different structural interpretation to (b) than the one their grammatical competence determines.

²⁹⁸ Stich (1985) p.131

conjunct or another clause consistently turns a structure that a speaker recognises into one that they don't; this being a matter of their motivation, attention, memory and other processing resources that they apply on particular occasions. As linguists come to understand the factors involved in processing, it is apparent when these factors, which are independent of the structure of competence, are masking the grammar that the speaker knows.

One could pick an arbitrarily high cut off point but “this would leave the grammar as before with generated sentences that are unacceptable” to speakers.²⁹⁹ Any account that we provide of why the upper reaches of this cut off point are never produced and not parsable will appeal to factors other than the structure of the state of competence and would also serve to account for our inability to produce or parse sentences beyond the proposed cut off point. But then it seems clear that we are appealing to performance factors to explain the relevant facts about what speakers can produce and parse, rather than appealing to the underlying structure of the language that speakers know being insufficient. If we have such an explanation of why the grammatical structures licensed by competence diverge from the structures produced and parsed, then it will cover both the unacceptability of sentences of short length and also those of great length. The package of competence and performance can then explain all the broad data including productivity, special structures of branching, embedding and conjunction, and what speakers can come to recognise with extended attention and prompting (their powers of recognition extending with the addition of performance resources but within the structural constraints provided by competence).

So, the length and complexity of expressions may mean that speakers are unable to put them together in production or assign them structural descriptions in parsing but this does not imply, as Katz claims, that “the best theory of the speaker's competence falsely predicts that strings of English words exceeding length n are ungrammatical.” The competence is a structured mental state that encodes pairings of sounds and linguistic forms over an unbounded range, whilst the issues concerning our capacity to recognise very long or complex grammatical structures fall clearly on the performance side of the competence-performance

²⁹⁹ Stich (1985) p.131

distinction. What can be uttered or parsed is not a direct reflection of the structure of the system of knowledge that shapes these processes.

Hence, whilst it is true that processors fail to output a structure in production, or assign a structural description in parsing, for sentences of very long length this does not suggest that the underlying system of competence deems them ungrammatical or turns them into “word-salad” for it is not a mechanism for processing. To get a sense of why this is so consider (20).

(20) [[[[He caught the rat]₁ that bit the child]₂ that ate the food]₃ that cost five pounds]_{4...]}_n

For extremely high n , English-speakers cannot utter or hear such sentences. In the most extreme case, they might die before getting to the end. But it would be a mistake to attribute this inability to the structure of our knowledge, which makes provision for the branching at each step, when we know that they are depleting other resources that are in finite supply such as time, attention and memory. It is interesting that while we can easily generate unacceptable but grammatical structures using basic constructions of nesting and self-embedding³⁰⁰, multiple branching structures³⁰¹ create far less trouble, being nearly perfect in their acceptability at extremely long lengths.

³⁰⁰ Phrases A and B form a nested construction if A falls completely within B, with some non-null element within B to its right and some non-null element to its left, as where *who wrote the book that you told me about* is nested in *I called the man who wrote the book that you told me about up*. Phrase A is self-embedded in B if A is nested in B and A is a phrase of the same type as B. So *who the students recognised* is self-embedded in *who the boy who the students recognised pointed out*. Nesting of a long and complex element reduces acceptability. Repeated nesting contributes tends more definitely to unacceptability. Self-embedding contributes still more radically to unacceptability.

³⁰¹ A multiple branching construction is one with no internal structure as in the subject noun phrase “John, Bill, and Tom” in “John, Bill and Tom saw the film last night”. It has no internal structure because “John”, “Bill” and “Tom” are immediate constituents and stand in no further structural relations amongst themselves. Left-branching structures include indefinitely iterable structures such as “John’s brother’s father’s uncle’s friend went” and right branching structures go the other way as in “this is the cat that caught the rat that stole the cheese...”. There are no clear examples of

One could postulate that human beings have a huge number of ready-made sentences in their head rather than a recursively structured grammar with infinite generative capacity. But this would serve only to replace an explanation that attributes humans a finite stock of lexical items and recursive procedures with one that attributes to them an infinite memory (which we know they don't have). To see that the memory resources required would be infinite, consider that my desire to own books, however insatiable, can be matched by the sentences I can form to describe that desire.

(21) Gareth desires a book.

(22) Gareth desires two books.

(23) Gareth desires eighty billion, seventy-five million, three hundred and fifty-two thousand and eight books.

Our recognition of sentences containing the natural numbers illustrates the point that the number of sentences I have at my disposal extends indefinitely. But the point can be equally well illustrated without appealing to the natural numbers by embeddings in the context of a propositional attitude verb.

(24) Gareth laughed.

(25) Simon thought that Gareth laughed.

(26) Craig knew that Simon thought that Gareth laughed.

(27) Sophie wondered whether Craig knew that Simon thought that Gareth laughed.

(28) Jim didn't care that Sophie wondered whether Craig knew that Simon thought that Gareth laughed.

The embedding needn't end here because we can keep going like this, in principle, forever. In reality, we will be hampered by exhaustion, boredom or death. But the

unacceptability involving only left or right-branching constructions, although these constructions are unnatural in other ways, and can require special intonation.

point is that there is no upper limit on the number and length of such sentences that we could understand if exhaustion, boredom or death could be kept at bay. Each time we embed we form a longer sentence. Since the embedding could continue indefinitely, there are infinitely long sentences licensed by our competence system which encodes the grammar the speaker knows. Our ability to embed like this, in principle, comes from our grammatical competence which is structured so that we could keep on putting the embeddings together. So Katz is mistaken in his claims about what a competence theory would predict because competence theories do not predict that extremely long and complex sentences are ungrammatical or “word salad”.

But Katz has a second argument against competencism.³⁰² The argument starts not from the potential length of grammatical structures but from the fact that a grammar must assign an infinite number of structural descriptions:

(P1) Grammars must generate infinitely many structures.

(P2) There are only a finite number of linguistic tokens in the universe.

Katz concludes from (P1) and (P2) that:

(C1) Grammatical theory must be an investigation of types and not tokens.³⁰³

Katz then makes a general claim about mental states:

(P4) Mental states are physical objects and not abstract types.

³⁰² Katz (1996) p.278

³⁰³ This conclusion doesn't immediately follow. What the premises entitle Katz to is the claim that the infinitely many structures licensed by the rules are not identical to the finite tokens, not to the positive claim that they are types. See Devitt's response to such a challenge to his view that linguistics studies tokens in §1.3.2. Devitt appeals to the notion of lawlikeness to explain talk of infinitude. Katz argument turns on taking the appeal to either finite tokens or abstract types to be exhaustive. Katz (1996) has further arguments for this step. I'm happy to let this step in the argument stand, though it requires further defence, because the premise relevant to the issue at hand is (P4).

From (C1) and (P4), Katz draws the target conclusion that:

(C2) Grammatical theory is not an investigation of mental states.³⁰⁴

In support of the crucial premise (P4), Katz says that: “Given the finiteness and discontinuity of matter, of brain-matter in particular, there can’t be an infinity of mental/neural objects.”³⁰⁵ Katz takes talk of infinite collections to be a mark of a requirement for types, and as he thinks that there can’t be an infinite number of mental objects he surmises that mental states must be tokens rather than types.

Katz is right that proponents of the psychological conception think of the state of grammatical competence as physically realised in the brain. Chomsky holds that generative grammar investigates such I-language states, which are real objects, an aspect of the human mind/brain. But Chomsky is explicit (see §1.2) that he *is* conceiving of the relevant mental states as abstract types. He claims that the generative grammarian is working at a level of abstraction from physical objects, such as neural mechanisms. So Katz’s (P4) does not accurately represent Chomsky’s view that the relevant states of the mind/brain are abstract types, the architecture of the brain described at a level of abstraction. In this vein Chomsky says that it is quite true that “I-languages are not *parts of brains*; rather they are components of the mind...That is, they are elements of the theory of mind abstracted from [such] states.”³⁰⁶

Chomsky argues that issues regarding the abstraction involved in theories of grammatical competence raise “no conceptual or other problems beyond those familiar in discourse involving theoretical entities, which are not relevant here.” Problems concerning the notion of abstraction in science are not relevant here because, to the extent they are genuine, they pervade the whole of empirical science, even the hard sciences. Chomsky invites us to consider by analogy theories of neural nets that describe the brain in terms of net-like connections between neural

³⁰⁴ George (1996) construes the argument slightly differently.

³⁰⁵ Katz (1996) p.278

³⁰⁶ Chomsky (1987) p.182

points. Neural nets are abstractions in the sense that when we claim that a particular type of neural net is a part of the theory of mind, we are not concerned with particular molecules or the precise orientation of the connections between neurons, i.e. the physical objects Katz claims the theory of mind is about in (P4). Chomsky says:

A theory of neural nets is part of a theory of the mind, in my sense, and we can look forward to progress in the brain sciences that will discover the physical mechanisms that exhibit the properties of neural nets. I-languages are in the mind (abstracted from the states of the brain, as explained) in the same (appropriate) sense...The issue of so-called 'Platonistic linguistics' does not arise in this connection.³⁰⁷

Katz thinks that Chomsky overlooks that there “cannot be enough mental/neural sentences for all the generated structural descriptions”³⁰⁸ and ought to concede that grammatical theories are about abstract types and not about the mind which is realised in the “finiteness and discontinuity of physical matter.” Katz claims that given the infinite structural descriptions assigned by the grammars, the view that each of these structural descriptions is a “distinct mental/neural object” will have to be given up. All but finitely many of the structures assigned by the grammar must be abstract types, so Katz argues.

But with Chomsky’s actual conception of the mind/brain now accurately represented, it is clear that Katz has erected a false dichotomy. On Chomsky’s view, theories of the mind *are* theories of abstract types. Katz claims that Chomsky faces a problem because of the finitude of the concrete neural state tokens that Chomsky is supposed to be appealing to, whilst Chomsky in fact insists that competence theories proceed at a more abstract level: a level of explanatory generalisation in terms of types of cognitive structures rather than token neural objects.

But it is not even clear that the linguist’s concern with types rather than tokens should concern physicalists about language. Tokens of physical objects, such as the neural objects, are finite in number. But the physicalist might hold that theories about the properties of the physical world appeal to types of properties.

³⁰⁷ Chomsky (1987) p.182

³⁰⁸ Katz (1996) p.278

These types will be framed at a more abstract level than their realisation in finite matter. Katz's reasoning would mistakenly suggest that, as these theories are concerned with types, none of them are concerned with physical objects.

Katz's two objections to "competencism" in generative grammar both involve imagining that we have two grammatical theories, G and G*. G would be a correct grammar for natural languages whilst G* would make false or insufficient grammatical predictions though it is a correct theory of grammatical competence. He invites us to note that proponents of the psychological conception:

have to prefer G* insofar as the psychological evidence has shown it to represent the cognitive structures underlying the speaker-hearer's use of language. Yet in forcing them to prefer G*, the philosophical commitment [to competencism] forces them to accept a false theory needlessly.³⁰⁹

But as yet we have been provided with no other way of determining the grammar of a speaker's language, other than by theorising about that speaker's grammatical competence. Hence, without such means, the situation that Katz envisages is illusory and the broad thrust of his objections is misguided.

Devitt seems to make a similar assumption to the effect that we have a way of determining what the grammatical properties of natural languages are independent of a developing theory of grammatical competence. He says that: "explaining the syntax of sentences is a prerequisite" for a theory of grammatical competence.³¹⁰ But the opposite is true: the generative grammarian needs a theory of grammatical competence to determine what aspects of performance events and linguistic judgements are a reflection of the grammar of a speaker's language. Moreover, once we move to a level of explanatory adequacy (§2.3), we need a theory of the grammar that a speaker has actually acquired in order to justify a choice amongst the possible DAGs for a speaker's language.

Developing a closely related point, Smith notes that without a theory of a speaker's knowledge of language, there are infinitely many rule systems that could account for the regularities among any sample of a speaker's language that we

³⁰⁹ Katz (1981) p.88

³¹⁰ Devitt (2006) p.29

happen upon.³¹¹ Without a theory of a speaker's knowledge, the theorist must first discover the rules governing the fragment and then extrapolate to the sentences outside the fragment – the rest of the language. But these rule systems may extend beyond the sample in quite different ways:

Each language is infinite in expressive capability. But speakers only produce a finite number of sentences. So our way of resolving this question must yield verdicts that extend beyond the produced fragment of the language. How are we to determine what the actual language spoken by a speaker is?

Smith points out that this is the philosophical problem of defining the “actual language relation” between a speaker and the language that he knows, with which David Lewis was concerned. Smith continues:

In a slightly different context, Lewis suggested that if there are rules governing the meanings of sentences and how they are paired with sounds in the used fragment of the language then these could be extended to generate the sentences of the language as a whole.³¹²

The central problem for Lewis's method is how we could know that the grammar for the fragment of the language used would serve to generate all and only the sentences of a speaker's language. Having grappled with this problem, Schiffer has come to the view that there is no method to come to correct hypotheses about the syntax and semantics of observed fragments of speakers' languages without appealing to facts about a speaker's grammatical competence of the sort theoretical linguistics uncovers.³¹³ So one clear way of defining the actual language relation, the one that Smith defends, is to develop a theory of a speaker's linguistic competence. By discovering the rules governing linguistic competence we get an

³¹¹ Smith (unpublished ms.) p.33

³¹² Smith (unpublished ms.) is referring to Lewis (1983, 1992). The context is different because Lewis is concerned with sound-meanings pairings conceived as properties of convention-bound public languages and not with the pairings of sound and grammatical forms that concern the grammarian.

³¹³ See Schiffer (2006).

answer to the question of which are the actual rules to generate the expressions of a speaker's language.

Devitt, like Katz, seems to assume that no such difficulties arise in the case of grammatical theory. But the philosophical problem of the actual language relation mirrors those practical problems which would be faced by the grammarian unless he appeals to a theory of the language a speaker has actually acquired, a theory of the structures that the principles of grammatical competence prescribe.

4. Linguistic Intuitions as Evidence for Generative Grammars

In order to determine what the grammar of a speaker's language is, generative grammarians draw upon a distinction between what is licensed by a speaker's grammatical competence and what is an effect of extraneous performance factors engaged in putting that competence to use. In this chapter I want to examine how grammarians get a grip on this distinction empirically, what sort of evidence is brought to bear in generative grammar and what sorts of hypotheses it is brought to bear on. My primary aim is to defend what I take to be an orthodox model of linguistic intuitions as they form a central source of evidence in generative grammar. According to this orthodox view, linguistic intuitions are the reflection of a dedicated system of grammatical competence as it interacts with performance systems for perceiving and articulating language. So conceived, evidence from speakers' linguistic intuitions allows the grammarian to investigate the competence-performance distinction empirically and thereby determine the grammars that speakers are competent in. This orthodox model has been attacked by Devitt. In its place, he advances a model of linguistic intuitions whereby they are speakers' theoretical judgements about the properties of languages.³¹⁴ I aim to make clear the rationale behind the orthodox model and the inadequacy of Devitt's proposed alternative.

4.1 *Intuitions as Evidence*

Both psychologists and philosophers are sometimes interested in people's intuitions. But it seems that the locus of their interest may differ. Psychologists are interested in gathering data on subject's intuitions, sometimes in elaborately designed experiments, because they reflect the workings of the psychological systems of

³¹⁴ Devitt (2006, 2006a, 2006b)

subjects that have them. Philosophers may also be interested in people's intuitions because of what they reveal about the psychological states of people that have them. But philosophers sometimes seem to be interested in people's intuitions because they are revelatory of facts independent of those psychological states. This interest is well-motivated where there is good reason to think that the subject is well-positioned with respect to the facts in question, perhaps because he has some special knowledge.

Generative grammarians draw upon the intuitions of competent speakers. We all have such intuitions. If I say to you "John posted the letter to Bill" you immediately recognise that as a part of your language. If I were to ask you whether it was ok, a perfectly good sentence of your language, then, no doubt, you would say that it was. However, if I say to you "to posted Bill the John letter" you are likely to recognise the words as part of your language but recognise that there is something amiss in the way they have been put together to form a sentence. In fact we have very intricate intuitions about the linguistic forms of our language and their meanings. Grammarians use these intuitions to investigate grammatical structure.

Harris seems to have held that linguists cannot investigate the structure of languages by examining speakers' intuitions. He said:

We do not ask a speaker whether his language contains certain elements or whether they have certain dependencies or substitutabilities...they (speaker's habitual judgements) are not sufficiently close to the distributional details, nor is the speaker sufficiently aware of them. Hence we cannot directly investigate the rules of 'the language' via some system of habits or some neurological machine that generates all the utterances of the language.³¹⁵

Harris surmised that, rather than investigate speakers' linguistic judgements about their language, the grammarian has to investigate "some actual corpus of utterances" from which we derive "such regularities as would have generated those utterances".³¹⁶ But current scientific practice suggests that Harris was wrong about what could be learnt from speakers' judgements. Though the questions are not framed in the metalinguistic way that Harris considered, it seems that grammarians

³¹⁵ Harris (1985) p.45

³¹⁶ *ibid.*

do investigate grammatical structure by probing intuitions. The orthodox view is that these intuitions are yielded by special cognitive systems responsible for recognising and shaping grammatical categories in speakers' utterances.

Although there are a wide variety of intuitions brought to bear on generative grammars, I want to focus on one illustrative example of how intuitions have been exploited for theoretical hypotheses. It concerns *null expressions*, elements of linguistic structures that are hypothesised to constrain our interpretation of linguistic material but have no phonological properties and so are not pronounced. Here are some prima facie considerations in favour of positing PRO, an empty pronoun, taken to occur in structures like (29):

(29) I would like [*PRO* to be on the beach]

In (29), *PRO* is dependent on "I" for its interpretation, so that (29) is understood roughly as with (30) or (31), by analogy with (32):

(30) I would like [*I* to be on the beach].

(31) I would like [*me* to be on the beach].

(32) He would like you to be on the beach.

One piece of prima facie evidence for PRO is that English-speakers have the intuition that (29) gets roughly the same interpretation as (31) and certain other related strings of borderline acceptability such as "I would like, for me, to be on the beach". Another is that English-speakers interpret it on analogy with (32). But we can offer far more evidence from speakers' intuitions to buttress *PRO*. As we saw in §1.1, a general requirement on reflexives is that they must have a local antecedent. In contrast, a general requirement on pronouns is that they are not dependent on local antecedents. We can gather evidence for hypotheses about locality and referential dependence by probing the following sorts of intuitions. Consider (33) and (34).

(33) Arthur believed that Lee shaved himself.

(33a) *Arthur_i believed that Lee_j shaved himself_i.

(34) Arthur believed that Lee shaved him.

(34a) *Arthur_i believed that Lee_j shaved him_j.

Why don't we hear (33) so that *himself* is *Arthur* rather than *Lee* as in (33a)? Indeed (33) is only acceptable under the interpretation on which *himself* means *Lee*. Compare (34) and (34a) where we don't hear (34) such that *Lee* means *him*. We might recognise the alternative interpretations as possible in principle but not part of our language. But why aren't they? One hypothesis is that *Arthur* is not local in (33) where the reflexive needs a local antecedent but *Lee* is local in (34) where *him* can't take a local antecedent. So the interpretations suggested in (33a) and (34a) contravene these binding principles.

The requirement on reflexives explains why we have the intuition that (35) is acceptable whilst (36) is much less good.

(35) They want [Gareth to take *himself* to the beach].

(36)*They want [Gareth to take *themselves* to the beach].

The square brackets mark the local domains for the reflexives – roughly the smallest clause which includes the reflexive. The reflexives *himself* and *themselves* must find antecedents within the square brackets if we are to find them acceptable and this turns out ok in (35) because *himself* agrees in number with *Gareth* and so can take *Gareth* as its antecedent being understood as in (37):

(37) They want [Gareth to take himself (Gareth) to the beach].

But unless *Gareth* is understood as a plurality in (36) then it fails to agree in number with *themselves* and so can't serve as its antecedent. Hence, we don't understand (36) on the model of (38):

(38)*They want [Gareth to take themselves (Gareth) to the beach]

The only antecedent available for *themselves* is *they*. Why don't speakers find (36) perfectly acceptable with *themselves* bound by *they*? The explanation is that *they*

falls outside the local domain, so fails to provide an antecedent for *themselves*, and hence *themselves* does not meet the requirement on reflexives that they are locally bound.

Now consider (39):

(39) Gareth wants [to sun himself on the beach]

English-speakers have the intuition that (39) is perfectly ok. But *himself* requires a local antecedent to bind it, and the only possibility seems to be *Gareth* which falls outside the local domain. The proposed solution is that (39) really has the structure given in (40) with the null expression, where *PRO* is interpreted as dependent on “Gareth” and “himself” as dependent on *PRO*, as in the paraphrase in (41):

(40) Gareth wants [*PRO* to sun himself on the beach]

(41) Gareth wants [(that) *he* (Gareth) sun himself (Gareth) on the beach]

That’s how intuitions data can issue in such theoretical postulates as *PRO* and the hypothesis that a bit of structure that must be visible to the conceptual-intentional system, must also be invisible to the perceptual-articulatory system. *PRO* is supported by the intuitions data to the extent it’s the best available explanation of that data. The aim of the theory is to explain these intuitive judgements of acceptability and unacceptability, and the structural interpretations that speakers come to.³¹⁷

³¹⁷ The discussion of how intuitions data might be used to support *PRO* is intended to be illustrative of the methodology. I am not arguing for the claim that *PRO* is, in fact, the best theoretical explanation of these phenomena. See Jacobson (1999) for an alternative account of bound and unbound dependencies, drawing on the resources of categorial grammar and type logical semantics, which does not appeal to unpronounced variables. Jacobson argues that her account simplifies the analysis of a range of constructions and preserves a relation of “direct compositionality” between the syntax and semantics, whereby the semantic rules are very close to the syntactic rules. Direct compositionality serves to minimise the theoretical machinery involved in explaining the interaction of syntax and semantics.

In classifying the data, the term *acceptable* is used to refer to utterances that are relatively natural and easy to comprehend without any paper-and-pen analysis. Acceptability is a matter of degree. The *unacceptable* structures are relatively more difficult though they may be grammatical for all that. Chomsky gave the following characterisation of acceptability:

(L)et us use the term “acceptable” to refer to utterances that are perfectly natural and immediately comprehensible... Obviously acceptability will be a matter of degree, along various dimensions... The more acceptable sentences are those that are more likely to be produced, more easily understood, less clumsy, and in some sense more natural. The unacceptable sentences one would tend to avoid and replace by more acceptable variants, wherever possible.³¹⁸

We can use the term *interpretable* to refer to the fact that a string has a natural interpretation, though it may have more than one such interpretation. Where a string is associated with more than one structural interpretation, linguists say we have *structural* ambiguity. We can also have *lexical* ambiguity where a single set of phonetic properties may be paired with more than one set of semantic properties as in “bank” – the financial institution – and “bank” – the river bank.

Intuitions about acceptability and interpretability can dissociate, for example, where speakers find a string acceptable in principle but uninterpretable as in (42):

(42) Colourless green ideas sleep furiously.

Although speakers struggle to assign (42) an interpretation and find it odd, they recognise it is dislike (43):

(43) Ideas colourless furiously green sleep.

The string presented in (43) is “word salad”. Examples like (42) suggest that we can prise apart speakers’ grammatical sensitivities from their ability to find a literal meaning for a string. Speakers’ intuitions about (42) suggest that while they may

³¹⁸ Chomsky (1965) pp.10-11

not have direct awareness of the underlying structure of (42) they do have a sense of whether what they are confronted with has the structure of a sentence of their language.

Acceptability and interpretability as data sources are to be distinguished from the theoretical notion of *grammaticality*, and what is generated by *a grammar*. Speakers have no intuitions about what a grammar mandates, in the theoretical sense of a grammar that concerns linguists. This is reflected in the distinction between grammatical competence and linguistic performance. A speaker's grammatical competence is just one component amongst an ensemble of systems responsible for their intuitions about acceptability and interpretability. Acceptability elicits or classifies intuitions but it is not something that can get a full explanation from linguistic theory as it looks to involve a range of factors beyond grammar. These include processing factors, semantic and pragmatic factors as well as commonsense knowledge and contextual factors. If I say that *I called the man who wrote the book that you told me about up*, this might seem rather unwieldy. But it is grammatical, where that means that the best generative grammars assign it a structural interpretation in just the way they do my less unwieldy utterance that *I called up the man who wrote the book that you told me about*. As a set of rules the grammar will generate a set of structures but how that set bears on what we do and don't find acceptable is a theoretical matter.

At deeper levels of explanation, where grammarians are concerned with very abstract principles and aiming for greater generality, they are not merely checking principles off against the observed intuitions. Though this might trivially ensure descriptive adequacy, it would be merely to recapitulate the data. The grammarian is concerned with the explanatory adequacy of his theory and not just data coverage. At deeper levels of explanation theoretical virtues like the generality and simplicity of the grammatical principles will be more to the fore. But speakers' intuitions will still play a guiding role in the investigation insofar as linguists are concerned to explore the languages that speakers are actually competent in and not just to come up with simpler and more powerful grammars.

Though intuitions have a central evidential role in generative grammar, there is no suggestion that other forms of evidence are irrelevant in principle or in practice. A central component of what Fodor calls *The Right View* of generative

grammar is that there is no proprietary body of data such that we can tell a priori what evidence might bear on grammatical hypotheses.³¹⁹ According to *The Right View*, not only speakers' intuitions but also facts about language use, grammar acquisition, the neurology of speaker-hearers,³²⁰ "or, for that matter, the weather on Mars" could, in principle, bear on grammatical hypotheses. It is a consequence of Fodor's *Right View* that there is no a priori distinction between a proprietary body of "linguistic" data and the "psychological" data, or any other kind of data. He says:

Suppose that some very clever astro-linguist were to devise an argument that runs from observations of the Martian climate to some or other constraint on theories of human psychology and thence to the proper formulation of the English pseudocleft. *Surely* we would say, 'Bravo and well done' and not 'Ingenious but not pertinent'.³²¹

The alternative to *The Right View* of generative grammar, Fodor calls *The Wrong View*, according to which we can stipulate in advance what evidence counts as relevant to grammatical theory. Fodor finds *The Wrong View* implausible in light of the way that science is really conducted; to adhere to it, Fodor claims, would be to take exception to the methodological principles that characterise the more mature sciences.

Katz's view is a species of what Fodor calls *The Wrong View*: "The criterion we apply to determine the relevance of a factual source to theories in linguistics is whether the source concerns the subject-matter of linguistics, language."³²² For Katz, the evidence from speakers' linguistic intuitions has precedence over all the other sorts of evidence. Katz labels such evidence "direct",

³¹⁹ Fodor (1985) pp.147-151

³²⁰ Though only a little is known about how I-language is realised in the brain, some aspects of I-language seem relatively localised. There are two areas of the brain's left hemisphere which appear to be particularly implicated in our linguistic abilities: Broca's area and Wernicke's area. We know that these areas are important because sometimes they are affected by some pathology or trauma, and particular linguistic deficits arise. For example, people with trauma to Broca's area can have problems dealing with complex grammatical constructions though their pronunciation may be normal.

³²¹ Fodor (1985) p.150

³²² Katz (1981) p.71

or “linguistic”, evidence and contrasts it with other forms of evidence, such as that from grammar acquisition or psychological experiment, which only constitute “indirect” or “psychological” evidence. By Katz’s lights, a linguist may get clues about grammar, when the “linguistic evidence gives out, by discovering psychological or neurological facts about speakers” but “indirect evidence depends on direct evidence for its legitimization as a relevant source of facts and direct evidence has a prior claim over indirect evidence.”³²³ Katz thinks that empirical evidence can never compel us to revise or abandon a grammatical hypothesis that is supported “on the basis of unchallenged direct evidence.”³²⁴

By contrast, according to The Right View, there is no distinction between direct and indirect evidence for grammatical theories. There are just different sources of evidence that may be more or less useful in our current state of knowledge. In principle, a psychological experiment or a piece of neurological evidence might be relevant to working out the form of a speaker’s language, just as a speaker’s intuitive judgement may be. But in our current state of knowledge, evidence from intuitions is more readily available than neurological evidence for grammatical hypotheses. There is no reason that other forms of evidence couldn’t lead us to disregard particular hypotheses that have been supported by intuitions data.

If one is a Platonist, like Katz, then one may be unmoved by methodological morals about general scientific practice because one denies that linguistics is like the other sciences that draw on empirical evidence. If linguistics is a part of mathematics or logic, and in these mathematical sciences one can choose what is of interest, then one can stipulate that only a certain range of data are to be included amongst the “linguistic” data. Platonists can then focus on the mathematical problem of formally specifying a grammar that predicts a certain range of data, such as speakers’ intuitions, insofar as these intuitions are not hampered by memory limitations, lapses in concentration and other performance factors. However, there is no particular reason why Platonists should attend to just those grammatical properties speakers acquire.

³²³ Katz (1981) p.71

³²⁴ Katz (1981) p.83

Generative grammarians, according to *The Right View*, are interested in explaining and predicting, inter alia, speakers' intuitions. The difference is that on *The Right View* grammarians are interested in intuitions because they hypothesise, rather than stipulate, that intuitions are revealing of the target of inquiry, the grammars speakers acquire. On *The Right View* the intuitions evidence has no privileged status. So conceived, the grammarian wants to know what sort of grammars can be acquired and how, how speech is understood, how language enters into cognition, what aphasics and schizophrenics reveal about language, what we have that animals lack: "in short, all that stuff that got people interested in studying languages in the first place".³²⁵ Fodor warns proponents of *The Wrong View* that while they are free to adopt a proprietary, or a priori, conception of the "linguistic" evidence and pursue such an inquiry, "all the action is at the other end of town".

To take a schematic example of the sort of evidence that could be useful to grammarians, consider evidence concerning processing and how it could be used to help filter out the effects of the parser that exhibits a different organisation to grammatical competence. Let's suppose we had two differently structured grammars, G and G', that hitherto could both explain a speaker's intuitions, and a theory M of the organisation of short-term memory in human adults that has received some independent confirmation. As Fodor points out, if the conjunction of M and G predicted that triply self-embedded sentences are not construable by human adults, whereas the conjunction of M and G' predicted the contrary then we have evidence for preferring G to G'; though they might make the same predictions about the intuitions of speaker-hearers' independently of the evidence from short-term memory.

Linguists draw on evidence from grammar acquisition in meeting explanatory adequacy, evidence from pragmatics in discerning what falls within the core language faculty and what falls outside, evidence from pathological cases and evidence from work on the brain.³²⁶ But one might wonder why if linguists are

³²⁵ Fodor (1985) p.60

³²⁶ For a discussion of evidence from work on impairments to the brain that may support Chomsky's postulation of a level of phonological representation common across hearers and signers, see Pettito (2005) pp.97-8.

really interested in this broad array of data, they seem to ignore a lot of readily available psychological evidence from language processing.

The relation between evidence from psychological processing and generative grammars is delicate. To take an illustrative example, Quine once argued that the phrase boundaries that grammarians posit are just artefacts of their theories, as they would be with formal languages, rather than a reflection of anything real.³²⁷ Quine claimed that for formal languages, there is no “right” grammar; one can arbitrarily pick one that generates the right theorems, and by analogy, that generative grammarians can just pick a grammar because the only thing that is real is the set of strings that the rules generate.³²⁸ Quine argued that it was “folly” to assume that there is a real answer to the question of where the phrase boundary is in strings of the form ABC. He thought it could be between B and C or between A and B as one liked, so long as the same strings are preserved.

But later some psychological experiments were carried out called the “click” experiments which led Quine to change his mind. In the click experiments subjects were presented with sentences like (44) and (45). With the bracketed material included we get different readings of the non-bracketed material and seem to process the non-bracketed material differently.³²⁹

(44) [Your] eagerness to win the horse is quite immature.

(45) [In its] eagerness to win the horse is quite immature.

In (44) we leave a main break in between *horse* and *is*, whilst in (45) the main break comes after *win*. If click noises are placed in the same objective positions (between, say, *the* and *horse*) in the acoustical stream as each of these sentences are uttered, subjects re-position the clicks in different places to reflect the main phrase breaks.

³²⁷ Quine (1972)

³²⁸ Of course, those committed to the psychological conception of generative grammar will reject Quine’s comparison because they think that grammars are psychologically, and ultimately biologically, realised. They think there is something real in the mind/brain, a particular procedure for assigning information about sound, structure and meaning to expressions. The choice of a theory will then be no more arbitrary than that of any other empirical inquiry.

³²⁹ Examples from Collins (2006)

After the click experiments were devised, Quine changed his mind about phrase boundaries and said that they are real because the click experiments show how you could get evidence to decide between the competing rule systems that generate them.³³⁰

Chomsky thinks that this is a serious misinterpretation of what the experiments establish. As Chomsky sees it the work on clicks serves only to test an *experiment* and not to test for phrase structure.³³¹ The work on clicks can test whether clicks are displaced in a way that accords with phrase boundaries. But if the click experiments had been out of step with phrase structure in clear cases then it would have suggested not that phrase structure be revised to fit with click displacement but that the experiment was poorly designed as an indicator of phrase structure. One would not, for example, hypothesise that phrase boundaries come in the middle of a word on the basis of click displacements being heard in the middle of a word. On Quine's view of the evidential relation between click displacement and phrase boundaries, in such a case, one would have to say that the phrase boundaries were in the middle of the word. Chomsky thinks that it would suggest instead that the experiment is not fit for purpose because the displacements suggest the wrong structures in clear cases. We make robust judgements about where the phrase boundaries are and if the click experiments do not gel with these judgements then the grammarian has a reason to reject the connection between click displacement and phrase boundaries.

So it is not so straightforward to determine what can be learnt about the structure of a speaker's grammar from perceptual experiments. Thinking again about the click experiments, if the processing goes the way of linguistic judgement then nothing has been learnt. But if the results of processing and judgement conflict then it is unclear whether or not the grammarian will have evidence that the theory which explains our judgements is wrong. The click experiments might not accord with phrase structure in the best understood cases of phrase boundaries. If so, then they may not be revealing in the controversial cases. If upon intuitive consideration of linguistic material speakers make firm judgements of acceptability and

³³⁰ Quine (1986)

³³¹ See Chomsky (2002) pp.125-7

interpretability that reveal the structure they assign to sentences, the grammarian might then reject the connection between particular perceptual results and grammatical hypotheses.

If the aim of the inquiry is a theory of the grammatical competence system as it is situated within linguistic cognition then it is natural to seek data from the subjects whose cognitive capacities are the domain of inquiry, just as in other areas of psychology. The key issue that Devitt raises is whether generative grammarians, like psychologists, are interested in speakers' intuitions because they are data for theories about speakers, and more particularly their grammatical competence.³³² Devitt suggests, to the contrary, that grammarians are concerned with these intuitions because they are revelatory of a domain of non-psychological facts to which speakers may have access through empirical reflection. According to the orthodoxy, *PRO* is part of an explanation of a range of data concerning the proclivities of speakers. The data are that *they find* certain forms acceptable and that certain interpretations are *available to them*. On Devitt's view, the data bear primarily on the properties of the presented sounds and marks rather than on the cognitive states of the speakers that intuit them.

The fact that intuitions may have a different significance to psychologists and philosophers, and the fact that there are competing models of how linguistic intuitions work, may suggest that the nature of *intuition* as a general category is not well enough understood to provide a model of linguistic intuition. The term "intuition" has been applied across a range of domains and cognitive abilities, and there may not be the kind of singular phenomenon that the term suggests. As Fiengo suggests we might do well to focus for now on the question of "what linguistic intuitions must be like *if they are to be the data of Linguistics*"³³³, and then see how an answer to this question bears on the subject matter of generative grammars.

³³² See Devitt (2006) pp.95-125

³³³ Fiengo (2003) p.255

4.2 The Orthodox Model: Linguistic Intuitions as Data for Psychological Theories

4.2.1 How do Intuitions bear on Competence Theories?

Psychologists elicit intuitive responses to presented material in such diverse areas as “Theory of Mind”, moral cognition and vision science. Just as reports of visual impressions constitute data for theories of the visual system that processes visual information, so, on the orthodox model, linguistic intuitions constitute data for theories of the grammatical competence that constrains the linguistic forms that a speaker finds acceptable and how they can be interpreted. On this model, intuitions data are brought to bear on a theory about a core component of the language speakers have internalised: grammatical competence. The character of a native speaker’s intuitions leads us to ask:

What must her internalized grammar be like...for her to find these arrangements of words acceptable but not those; for her to be able to interpret a sentence in this way but not in that. To arrive at specific hypotheses about the internalized grammar we reason counterfactually: had the grammar been different, had it not respected a particular constraint then it would have been possible to hear certain utterances differently.³³⁴

One of Devitt’s charges against this orthodox model is that there is currently “no account of how the rules embodied in the language faculty could provide intuitions about syntactic facts”.³³⁵ In one sense, Devitt is correct. The orthodox model provides only a very partial explanation of how grammatical competence could issue in intuitive judgements of acceptability and interpretability. There is no complete explanation of how intuitions are produced available, only a partial explanation of the character of those intuitions in terms of an underlying system of grammatical information and systems for putting that information to use. As

³³⁴ Smith (2006) p.959

³³⁵ Devitt (2006a), see also Devitt (2006, 2006b).

Chomsky has pointed out “we do not, of course, have a clear account, or any account at all, of why certain elements of our knowledge are accessible to consciousness whereas others are not, or of how knowledge, conscious or unconscious, is manifested in actual behaviour.”³³⁶

The structure of the competence system provides some explanation of the form and character of the intuitions, and the performance mechanisms are intended to provide some explanation of how the linguistic forms licensed by the competence system are employed in judgement, speaking and understanding. But Devitt is right that there is no complete theory of how competence and the language mechanisms issue in linguistic intuitions. The explanation of this is that, for the reasons outlined, intuitive judgements of acceptability and interpretability (or indeed, conscious awareness) are not phenomena that can get a full explanation from a theory of grammatical competence. The broader empirical challenge is to try and understand all the different factors involved in linguistic judgements, in the same way that we might try to understand the factors that shape judgements in other areas of psychology. After all, there is currently no account of how the computations of the visual system issue in intuitive judgements about the properties of a presented scene. So, it’s not clear that there is a special problem.

Speakers’ linguistic intuitions are far more discriminating than one might expect. Take the following illustrative example. Sentence (46) is ambiguous between a reading on which *duck* and *swallow* are both nouns and one on which they are both verbs.

(46) I saw her duck and swallow.

Interestingly it is two ways and not four ways ambiguous in natural language. We can hear (46) as containing the two verbs *duck* and *swallow*. We can also hear it so that *duck and swallow* is an NP containing the nouns *duck* and *swallow*. But no speakers hear it as having mixed readings on which *duck* is a verb and *swallow* is a noun, or vice-versa, and it is never uttered with this meaning. It is logically possible that the sentence should have the mixed readings. And we could artificially stipulate

³³⁶ Chomsky (1986) p.270

that the sentence was to be understood in such ways. This would be to create a piece of artificial language, since no one naturally acquires such a language. The fact that English-speakers don't recognise these mixed readings can be taken as evidence concerning the organisation of their grammatical competence.

The explanation of the relevant intuitions is that the competence system is structured according to a *co-ordination constraint*. The constraint determines that we can only conjoin constituents of the same grammatical category. This hypothesis about grammatical competence, supported by the evidence from intuitive judgements about forms like (46), explains why speakers are unable to hear such logically possible, mixed readings. The interpretations that speakers can consciously hear, and then judge, such expressions to have, are crucial evidence for, or against, this hypothesis about their grammatical competence. I only hear (46) two ways, and I can only consciously hear, or attend to, one of those interpretations at a time. Once I recognise that (46) has these two readings I can consciously switch my attention back and forth between them. Though, in one sense, nothing in the sounds themselves or the written marks changes, something in my conscious experience changes as I shift my construal of (46) back and forth.

Though grammarians do not tend to distinguish explicitly between *intuitions*, *judgements* and *intuitive judgements*, it may be that *intuition* and *judgement* are picking out distinct aspects of speakers' engagement with language. The term "intuition" seems to refer to the unreflective take or awareness that the speaker has of linguistic form, whilst "judgement" seems to refer to the formation of a belief or report on the basis of that intuitive take or impression.³³⁷

It also seems that nothing very intellectualised is meant by *intuitive judgement* in this context. *Intuitive judgement* might suggest that the data take the form of a speaker-hearer judging *that* a linguistic form is grammatical, ambiguous

³³⁷ Fiengo (2003) makes a distinction between linguistic "intuition" as that access we have to the structures of sentences that involves no conscious reasoning and linguistic "judgement" as our evaluation of that to which we have the immediate access. I model linguistic judgements as involving a straightforward report on what we are aware of rather than evaluating the intuited properties. Of course, it might be that informants do both. But it seems to me that there are a core of cases where the speaker intuits some properties and reports on what they intuit without evaluating the intuited properties as such.

and so on. But in proffering their linguistic judgements speakers are not generally required to have linguistic concepts with which to express the status or structural interpretation they have assigned to linguistic material. As Collins notes:

We are interested in how speaker/hearers interpret strings, either their own or those of others. This covers a panoply of different attitudes. Most often, the data are simply that speaker/hearers find a string unacceptable. Period... Other times, we might be after a more explicit judgement, and so we ask, 'How many ways ambiguous is the sentence, *I had the book stolen?*'. Other times we might ask, 'Who is fixing the car in the sentences *Bill told Sam to fix the car* and *Bill promised Sam he would fix the car.*'³³⁸

To be capable of interpreting linguistic material, speakers need not have any metalinguistic concepts with which to categorise the material or any special expertise beyond competence in their language. No expertise is required, only an honest report of how things strike one. In this respect, the linguistic intuitions data is analogous to the data for other psychological theories, where "there is no relevant expertise about the data beyond the authority of the subject's own perceptions."³³⁹

On the orthodox model, a speaker's intuitions are simply data to be explained and eliciting a speaker's linguistic intuitions does not require attributing them any of the theoretical concepts that animate grammatical theory. If a linguist says that a speaker has the intuition that a reflexive must be locally bound, this is just a shorthand way of saying that a speaker has linguistic intuitions that can be explained on the basis of his possessing a grammatical competence, organised according to principles involving reflexives, locality and binding.

4.2.2 Linguistic Intuitions and Visual Impressions

Much of the evidence for computational theories of vision has come from subjects' responses to presented material, either in the form of reports on the way that things

³³⁸ Collins (2006) pp.7-8

³³⁹ Slezak (unpublished ms.) pp.33-34

appear or seem to them, or their use of such appearances to carry out visual tasks. Chomsky suggests a comparison between the way that speakers' intuitive responses to linguistic material are brought to bear on generative grammars and the way that subjects' reports in visual experiments are brought to bear on theories of vision:

A generative grammar attempts to specify what the speaker knows, not what he may report about his knowledge. Similarly, a theory of visual perception would attempt to account for what a person sees and the mechanisms that determine this rather than his statements about what he sees, and why, though these statements may provide, useful, in fact compelling evidence for such a theory.³⁴⁰

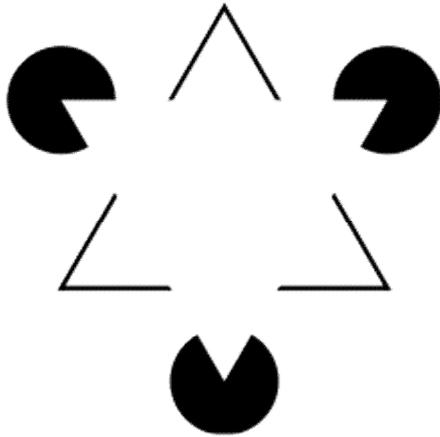
Chomsky takes the study of the computational operations of grammatical competence to be the study of “mental representations and computations, much like the inquiry into how the image of a rotating cube in space is determined from retinal stimulations.”³⁴¹

One similarity between the experimental investigation of vision and the investigation of speakers' linguistic intuition is that the intuitive takes speakers have on linguistic material are pre-doxastic in a way that compares with visual appearances. And the pre-doxastic nature of linguistic intuitions and visual appearances is of interest to the grammarian and vision scientist respectively. Upon presentation of a Kanizsa triangle as in Fig. A, subjects report an impression of an equilateral triangle with its corners in the circular (pacman-like) elements of the presentation.

Fig. A

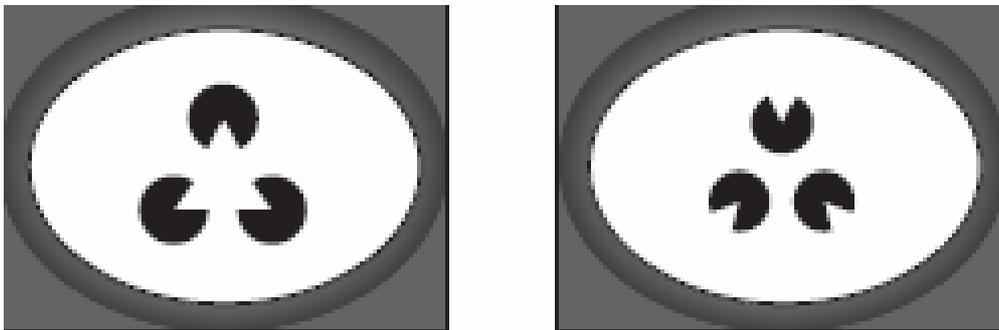
³⁴⁰ Chomsky (1965) pp.8-9

³⁴¹ Chomsky (2000) p.125



This impression of a triangle exhibits belief-independence because it can be had even by subjects who do not believe that there is a triangle there and who have seen how the illusion is created by comparing the two boxes in Fig. B.

Fig. B.



There are a large number and variety of visual illusions such as the Necker Cube and Muller-Lyer lines which can be enjoyed or suffered, even by those who do not believe in the veridicality of the appearances. They provide important evidence about how the visual system fills in and processes the information which is input to it. These visual seemings or impressions that are generated in the course of visual processing clearly encode more information than is given to the senses, and are of particular interest for precisely that reason. They are sometimes called *percepts* to

highlight that they are impressions or seemings, and distinguish them from genuine perceptions.

Such mental capacities as vision, which exhibit independence from belief and general intelligence are said to be *encapsulated*. As Fodor originally employed this notion, *informational encapsulation* meant that the computations which a system carries out are defined over a restricted base of information and not penetrated by central cognitive processes, such as those involved in belief-formation.³⁴²

Linguists' interest in speakers' intuitions is comparable along this dimension. With strings like (47), an impression of full structure can persist despite our coming to believe that (47) does not have a full structure:

(47) Many more people have been to France than I have.

If we try to fill in the structural ellipsis, we see that the sentence makes no sense: *Many more people have been to France than I have (been to France?)* Strings like (47), particularly if read at sufficient speed, can strike us over and over as having a full structural interpretation even once careful inspection has revealed to us that they have none, or we have come to believe as much on the basis of testimony.

Equally, (48) and (49) can strike me, again and again, as lacking a full structure even once prompting, extended attention or the testimony of others has issued in contrary beliefs:

(48) The man the cat the dog bit scratched died.

(49) The horse raced past the barn fell.

(49a) The horse raced past the barn.

The intuitions data suggest hypotheses about our grammatical competence and our ability to use it in real-time. As I suggested in Chapter Three, given some concentration, time or prompting, speakers can come to recognise (48) as a sentence of their language, though they tend to find it unacceptable at first blush and never

³⁴² Fodor (1983)

use that structure. Once I know that (48) is a double centre-embedding then I can pair off the embeddings and the structure becomes apparent: *the dog bit the cat that scratched the man that died*. What stops me recognising (48) as part of my language is the lapse of attention and ability to process the embedding. As I argued in Chapter Three, in this and a vast range of other cases, it is not the language I know that rules the structure out, but the extraneous factors involved in using the grammatical information that stop me from repeatedly centre-embedding. In performance something occludes with my standing knowledge of centre-embedded structures.

Sentence (49) strikes many English-speakers as leaving the verb “fell” dangling off the end of an otherwise good sentence (49a). The process by which I come to notice that the dangling verb is the main verb and *raced past the barn* is an embedded clause may forever erase this impression that *fell* is dangling. But it needn’t. The intuition can be robust and a parse of (49) on which *The horse raced past the barn* is a sentence rather than a determiner phrase can continue to suggest itself. Such structures “lead us up the garden path” and there is a residual impression of unacceptability. Though a parsing explanation is available on which speakers first find the tensed phrase (49a) and so do not structure (49) such that *fell* is the main verb, the explanation may in fact be partly grammatical. Compare (50) and (50a):

(50) The paint daubed on the wall stank.

(50a) *The paint daubed on the wall

Sentence (50) is not grammatically ambiguous in the way that (49) is. The paint can’t daub whereas the horse can race. The subject’s grammatical proclivities can be probed in this way by varying the presented material and seeing how the immediate intuitive take varies.

These examples suggest that linguistic intuitions provide evidence for investigating an encapsulated grammatical system and a distinction between a system of grammatical competence and integrated performance systems.

As with visual experimentation, there can be priming effects. If an ambiguous sentence such as (51) is presented in a certain context, the hearer may take it in a unique way and fail to see the ambiguity.

(51) Flying planes can be dangerous.

In such instances speakers may even reject the second proposed interpretation as unnatural or contrived. Nevertheless, the speaker's "intuitive knowledge is clearly such that both interpretations are assigned to the sentence by the grammar he has internalized in some form."³⁴³ This knowledge can be drawn out, sometimes in quite subtle ways to determine the actual form of the underlying competence. We can see this by taking a less transparent ambiguity like (52):

(52) I had a book stolen.

Few hearers will notice the fact that this structure is three ways ambiguous. But the fact that their internalized grammar provides three structural descriptions for the sentence (corresponding to my having the book stolen from me or for me, or my stealing the book myself) can be brought out by providing elaborations of (52).

(52a) I had a book stolen *from my car when I stupidly left the window open.*

(52b) I had a book stolen *from the library by a professional thief who I hired to do the job.*

(52c) I *almost* had a book stolen *but they caught me leaving the library with it.*

In bringing out the three-way ambiguity of (52), we do not have to present the speaker with any new information about his language, we only need to arrange linguistic material in such a way that the structures his grammatical competence affords him become available.

³⁴³ Chomsky (1965) p.21

Linguists have clever ways of controlling for pragmatic effects on linguistic judgements. Consider “minimal pair” experiments.³⁴⁴ Speakers in these experiments are presented with strings which are hypothesised to differ only in that one fails a certain grammatical constraint. The speakers are asked, simply, which is a worse sentence of their language. Naturally, such controls do not eliminate the intrusion of pragmatic factors, but rather aim to marginalise them. They reflect the fact that the grammarian is not so much concerned with what might be conveyed or implied by using a string in a particular communicative context. The experimental setting serves to strip away some of that context and leave the speaker to make a report revealing of the structural materials that are immediately available to them on the basis of the linguistic material alone.

There is evidence that the orthodox model I’ve outlined is precisely the model of linguistic intuitions as psychological data, analogous to visual reports, which Chomsky has in mind:

A grammar is a system of rules that generates an infinite class of “potential percepts”... In short, we can begin by asking “what is perceived” and move from there to the study of perception.³⁴⁵

The comparison has been noted by others. Slezak thinks that the familiar perceptual phenomena involving Kanizsa illusory contours and the like, where visual percepts are used to investigate perceptual constancies, are just like the intuitions reported on in linguistic judgement. He remarks that:

The two interpretations of the Necker cube known intuitively to a ‘visual virtuoso’ are closely analogous to the two meanings of an ambiguous sentence known as the percepts of the native speaker.³⁴⁶

Longworth has developed the same theme, comparing visual reports with reports of one’s intuitive take on linguistic material. He compares “quasi-perceptual”

³⁴⁴ See Crain and Thornton (1998).

³⁴⁵ Chomsky (1972) p.168

³⁴⁶ Slezak (unpublished ms.) pp.34

grammatical appearances with the role of perceptual appearances in vision science. Longworth considers visual experiments where subjects are presented with various patterns of printed marks and asked what they can make of those marks, whether some seem closer together than others, and so on. The key point is that the reports that subjects are requested to make are not reports on the properties that they *believe* the marks to have, for:

One may very well know that the marks are equally well spaced on the page. What one is asked for are reports about how the marks strike one, or how they seem to one, where how they seem to one is typically impervious to how one believes them to be.³⁴⁷

The reports are reports on one's experience. On this model, the intuitions data do not target grammatical properties conceived as properties of the extra-psychological environment. They serve as mental meter readings.

4.2.3 Are Linguistic Intuitions the “Voice of Competence”?

Devitt's main bone of contention with this orthodox model is highlighted by the name he gives it – the “voice of competence” view. Devitt construes the model such that it is committed to speakers having a direct access to the principles that organise their grammatical competence. He calls this “Cartesian access”, comparing it to the sort of direct access Descartes thought we had to the contents of our own minds. Devitt then wonders why, if linguistic intuitions are the voice of our grammatical competence, we cannot read off the properties of the grammar we are competent in from our intuitions. As he puts it, “if competence really spoke to us why would it not use its own language and why would it say so little?”³⁴⁸ Devitt thinks that if the source of linguistic intuition were our grammatical competence then we should have intuitions that give articulation to the very properties that characterise our competence. But our intuitions do not seem to give articulation to those

³⁴⁷ Longworth (unpublished ms.) p.11

³⁴⁸ Devitt (2006) pp.100-3, (2006a)

grammatical properties which are only revealed by the theoretical inquiry into grammar.

Devitt is correct that we do not have the kind of direct awareness of the underlying grammatical properties licensed by our competence system that he thinks the orthodox model appeals to. Fiengo's attitude to the idea that we have such direct awareness seems to me to be representative:

[I]t goes without saying that we have no such awareness. If one is in any doubt, all one need do is reflect on the fact that syntactic proposals for even the simplest sentences are often in debate... [If we had such awareness] much that is debated in Linguistics could be settled by appeal to the intuitions of speakers. We could ask them what the structures of sentences should be, and they could tell us.³⁴⁹

But there is a way to answer Devitt's question about why linguistic intuitions do not "use their own voice" and why they "say so little" that is suggested by the orthodox model. Grammatical competence does not "use its own voice" insofar as the properties of the sub-personal competence system are not available to mere personal-level reflection. We have to make a theoretical inference from a speaker's judgements of acceptability and interpretability to the structure of the underlying competence and its place within wider performance systems. The competence "says so little" because grammatical competence is not the only factor involved in linguistic judgement which engages systems of linguistic performance and more besides. Grammatical competence is not all that speakers bring to bear on presented strings. This has always been Chomsky's view:

The unacceptable grammatical sentences often cannot be used, for reasons having to do not with grammar, but rather with memory limitations, intonational and stylistic factors, "iconic" elements of discourse (for example, a tendency to place logical subject and object early rather than late) and so on...we cannot formulate particular rules of grammar in such a way as to exclude them.³⁵⁰

³⁴⁹ Fiengo (2003) p.258

³⁵⁰ Chomsky (1965) p.11

We don't know how conscious judgements are derived, or what causal role the linguistic systems play in issuing in these judgements. Linguists infer that a grammatical competence system shapes these intuitions but we know that linguistic intuitions are not an unproblematic reflection of the underlying competence.

In that sense, Devitt is correct that intuitions are not the "voice of competence". But then he is wrong that this is a commitment of the orthodox model. If grammarians do routinely think that linguistic intuitions are the "voice of competence", then it is apparent that they must think the voice is very muffled. The relation between the intuitive judgements and the structure of competence is not transparent; it is a highly theoretical matter to determine what it is.³⁵¹ As Fodor notes, in offering their intuitive takes on strings, subjects have access only to the upshot of their linguistic systems including the grammatical system and the performance systems. So their intuitive judgements will not give voice to the internal organisation of those systems. The internal organisation of the competence and performance systems, the yields of the systems taken individually and their manner of interaction will all be "completely opaque" to speakers as they respond to linguistic material.³⁵² So there is no reason, on the orthodox model, to expect a speaker's judgements to give "voice" to their competence in the way Devitt suggests. The intuitive judgements target the very broad properties of acceptability and construal, so do not "say" anything about the deeper, general and highly intricate properties of the competence system involved in their etiology. We may conclude, as Fiengo suggests that the fine-grained grammatical properties are not accessible to conscious intuition, keeping distinct "intuitions, which are conscious states, and those processes of which we are unconscious that perhaps underlie our intuitions."³⁵³

³⁵¹ This would involve, for starters, an account of the role of competence with respect to linguistic performance.

³⁵² Fodor (1983) p.60

³⁵³ Fiengo (2003) p.257

4.2.4 Are Linguistic Judgements and Visual Reports Disanalogous?

Devitt argues that there is an important disanalogy between linguistic intuitions and the perceptual reports drawn upon in vision science. He thinks that what the visual module delivers to the central processor is the impression on which a judgement about *what is seen* can be formed whilst what is delivered to the central processor by the linguistic systems is an impression of *what is said*.³⁵⁴ The important difference, according to Devitt, is that whereas judgements about what is seen are the ones of interest to the vision scientist, judgements of what is said are not the ones that concern the grammarian.³⁵⁵ The grammarian, Devitt claims, is interested in the grammatical properties of expressions and so he is not interested in speaker's intuitions about what is said. So, Devitt denies that the intuitions that concern the grammarian are derived in a way analogous to the way perceptual judgements are derived from the outputs of the visual module.

This argument against the orthodox model is unconvincing for two reasons. Firstly, it is unclear why the only intuitive materials made available to judgement by the linguistic systems are intuitions about what is said rather than intuitions about the acceptability of linguistic forms and their possible structural interpretations. The examples I've considered in this chapter suggest the contrary.

Secondly, it is unclear why intuitions about what is said are not of interest to the grammarian. These interpretability intuitions *are* revealing of linguistic form, because linguistic form acts as a constraint on speakers' interpretation of what is said, though these interpretations are informed by more besides, in particular by semantic and pragmatic information. The sorts of intuitions drawn upon by grammarians and pragmatists are not sharply discontinuous in the way Devitt suggests. The speaker has intuitions about what is said on a given occasion which are partially determined by his immediate recognition of the structure of the expressions of his language. It may be that Devitt thinks that such intuitions are of

³⁵⁴ It may be inaccurate to say that the visual module delivers an impression of *what is seen* to the central processor. Depending on how "what is seen" is to be understood, it may be that the visual module delivers something shallower than that.

³⁵⁵ Devitt (2006) p.24 fn.25

no use to the grammarian because they do not involve speakers' having intuitions about theoretical properties like c-command and binding, in the sense of making explicit mention of these theoretical properties. But if the intuitions are being used as evidence about the internal structure of grammatical competence and how competence is organised in terms of such properties, then the intuitions are of obvious use though they do not involve speakers' overtly considering such theoretical properties. The informant need have no way of describing sentences; he need only associate various first-order meanings with sentences. As structure constrains interpretation, it is then a theoretical matter to determine what reflects grammatical competence as opposed to other competences and performance factors.

4.2.5 The Orthodox Model and Introspection

On the orthodox model, the linguist is not asking speakers about their conception of their own language. But in asking them to report on their intuitive responses he is drawing upon evidence that goes beyond mere observation of their behaviour. Ludlow has raised the following concern for the orthodox model.³⁵⁶ As the orthodox view has it, the speaker is reporting on whether a string seems acceptable to him, which readings are available to him or how the material strikes him. One natural way of understanding this model of linguistic judgements is to think of the judgements as self-reporting or introspection. Moreover, it is very common for grammarians to consult their own intuitive takes on pieces of their language independently of the judgements of informants or publicly observable behaviours. If the intuitions themselves are private, then one might wonder if they are a suitable source of scientific evidence. The intuitions can be *reported on* by the linguist, or by his informant. But isn't the *publicity* of the data a requirement in science?

It seems to me that even if the grammarian is committed to linguistic intuition being a kind of introspection and to the publicity requirement, though he may not be, the worry about intuitions data can be allayed. The intuitions *are* publicly available and shared in at least three ways. Firstly, a *range of speakers* of

³⁵⁶ Ludlow (forthcoming) ch.3. Ludlow has his own way of dealing with the concern.

the same native language can report the same intuitive take. Second, the experiments can be *reproduced*. There is a regularity and robustness to the reports that is in keeping with underlying patterns in the speakers' interpretations. Thirdly, the intuitive takes are publicly *reported on*, and these reports are made in circumstances where we have no reason to suspect informants of misleading us.

It cannot be the *unobservable* nature of intuitions in and of itself that is a source of concern because, ultimately, linguistic behaviour is going to require explanation by unobservable states – inferred theoretical entities - just as scientists appeal to unobservables in the hard sciences. Of course someone might challenge the proposed connection between the unobservable intuitions and the reports but then they would require a saving hypothesis about the source of the reports.

We could imagine an extreme case where a linguist is faced with a pathologically dishonest subject. The dishonest subject reports that *him* must be *John* in *John shaved him* and that *himself* can't be *John* in *John shaved himself*. In practical terms, the dishonest subject throws a temporary spanner in the works because his intuitive reports are not a good guide to his grammar. But though the linguist could face such local issues about the reliability of his data, the possibility of such a subject does not threaten the broader methodology. There may be factors in the judgement, such as dishonesty, that it falls to other parts of psychology to explain. But at some level the subject's behaviour is just data to be explained, if not perhaps by the resources of generative grammar. The grammarian may be interested in aspects of the dishonest subject's behaviour, as his deliberately skewed pronouncements may still be revelatory of his competence, if we knew how to isolate the skewing mechanisms. The point is that while the situation is more awkward in this case, it is not different in kind: we have a number of cognitive factors, of which one is grammatical competence, interacting to yield a higher-level behaviour effect. And it's is the goal of the psychological sciences very broadly conceived to offer some explanation of such complex behaviour.

To summarise, Fodor argues that an adequate conception of generative grammar "should explain why it is that the intuitions of speaker/hearers constitute data

relevant to the confirmation of grammars.”³⁵⁷ We have seen that, on an orthodox model of linguistic intuitions, the psychological conception of generative grammar fulfils this condition:

It says ‘We can use intuitions to confirm grammars because grammars are internally represented and actually contribute to the etiology of the speaker-hearer’s intuitive judgements.’³⁵⁸

4.3 Linguistic Intuitions as Theoretical Judgements

4.3.1 Devitt’s Model and Belief-Independence

Devitt lists as the third major conclusion of his book, “Speakers’ linguistic intuitions do not reflect information supplied by the language faculty. They are immediate and fairly unreflective empirical central-processor responses to linguistic phenomena. They are not the main evidence for grammars.”³⁵⁹ Devitt has argued that speakers’ linguistic intuitions are not the upshot of a dedicated system of grammatical competence interacting with linguistic performance systems. Rather, on his account, linguistic intuitions are fairly unreflective or “low level”, theoretical beliefs about the grammatical properties of languages. They are “low level” in that speakers do not typically enter into much serious reflection upon the properties of their language or have knowledge of any scientific linguistics. They are “theoretical” in the sense that they involve central processing, or general intelligence, in working out the properties of external linguistic stimuli, albeit relatively immediately.

Devitt rejects the theoretical inference from the character of our intuitions to a competence system organised according to grammatical principles. He claims that his own explanation is more “modest” in appealing only to the generation of

³⁵⁷ Fodor (1985) p.152

³⁵⁸ Fodor (1985) p.152

³⁵⁹ See the glossary of Devitt (2006) and (2007a p.2).

intuitive judgements by central processing. And he points out that everyone should be committed to the existence of central processing.

But Devitt's view that linguistic intuitions are theoretical beliefs, derivative of central processor responses to external stimuli, is inconsistent with the pre-doxastic nature of linguistic intuition and its encapsulation, discussed in the previous section. The view that linguistic intuitions are amongst our theoretical beliefs seems unable to accommodate the persistence of impressions of grammaticality and ungrammaticality through contrary beliefs. So Devitt would have to try and explain these phenomena away somehow. Devitt would also have to explain why these intuitive impressions seem to be more than just "relatively unreflective". They seem to be *mandatory*. We can't help but hear the sounds of our language as structured and meaningful, forming an intuitive take on their form and interpretation independently of our choosing to reflect upon them.³⁶⁰ Further, our linguistic intuitions evidence special hierarchical and recursive principles that are highly language-specific. Devitt's view that these intuitions are central processor responses would have to accommodate these facts and compete with explanations that appeal to a dedicated competence system.³⁶¹ I'll now argue that the failure of Devitt's account reinforces the orthodox inference to the best explanation to the properties of the dedicated grammatical competence system.

4.3.2 Devitt's Model and Folk theory

Devitt's view of intuitions is "based on a view of intuitions in general"; that they are conditioned by empirical theory.³⁶² Intuitions, on Devitt's model, differ from other such theoretical beliefs "only in being fairly immediate and unreflective."³⁶³

³⁶⁰ The idea that the operations of a cognitive system are mandatory is often associated with the idea of a Fodorian *module* (Fodor 1983). Though the parser may behave like a module, Collins (2004) argues that it is wrong to think of FL itself as a module in this sense.

³⁶¹ Devitt (2006 chs.8-10) argues that these grammatical rules are psychologically real to the extent that they are rules of a more general language of thought.

³⁶² Devitt (2006) p.10

³⁶³ *ibid.*

Consequently, for Devitt, the grammarian is a good source of intuitions because he has spent a lot of time reflecting on language and has more theoretical knowledge:

If the person is a linguist then she will of course deploy her concepts from her linguistic theory... I think we should generally prefer the intuitions of linguists to those of the folk in seeking evidence.³⁶⁴

Ordinary informants are not such good sources of data, on Devitt's model, because they don't possess scientific theories involving concepts like c-command and binding; perhaps having only a little knowledge of verbs, nouns and the like. This is in stark contrast to the orthodox model, according to which speakers are not being asked for their opinions about such properties of linguistic material at all. They are being asked only to respond to linguistic material in terms of such broad categories as how acceptable and intelligible they find it, which interpretations of it they come to and how difficult it is for them to achieve certain readings. On the orthodox model, the intuitions gathered by linguists are just data to be explained. Once gathered they form the grammarian's observations rather than being statements of a theory. In contrast, for Devitt, a speaker's linguistic intuitions are amongst their theoretical beliefs and these beliefs constitute a less powerful theory than the linguist's.

To explain how ordinary speakers' theoretical beliefs could count as evidence for the *science* of language, Devitt once drew an analogy with our intuitions about physical reality:

Just as physical intuitions...can be produced by central processor responses to appropriate phenomena, so also can linguistic intuitions. These linguistic phenomena are not to be discovered by looking inward at our own competence but by looking outward at the social role that symbols play in our lives. When linguists do this now, they do not start from scratch. People have been thinking about these matters for millennia. The result of this central processor activity is folk, or otherwise primitive theory: the linguistic wisdom of the ages. The wisdom will be a good albeit not infallible guide to the nature of linguistic symbols.³⁶⁵

³⁶⁴ Devitt (2006a) p.22 fn.22

³⁶⁵ Devitt and Sterelny (1989) p.522

The analogy is unhelpful for Devitt. If it were good then our ordinary beliefs about physical reality could play an important evidential role in physics. But in physics one does not expect the folk's opinions to inform scientific theory, and there is no reason to assume that the concepts and constructs of ordinary thinkers carry over to scientific debate. Equally, there's no reason to expect folk opinion to constrain linguistics. As Neil Smith remarks:

In physics one does not expect folk views to inform the expert's theory construction, and while ethnoscience is itself an interesting field of inquiry, there is no reason to assume a priori that the concepts and constructs of pre-scientific debate should carry over unchanged into formal theories of I-language.³⁶⁶

In my view, it is therefore a consequence of Devitt's model of linguistic intuitions that speakers' intuitions should not be afforded the central evidential role that they are afforded.

Devitt might try to soften this result by maintaining that speakers' theoretical beliefs about grammar are largely correct. But if this were true then much that is debated by grammarians could be settled by appeal to speakers' intuitions, requiring little scientific theorising. Despite Devitt's previous commitment to "the linguistic wisdom of the ages", he has now recognised that his model would require some revision of existing methodology:

Where the judgements are those of the ordinary speaker, the theory will be folk linguistics. We do not generally take theory-laden folk judgements as primary data for a scientific theory. So we should not do so in linguistics.³⁶⁷

As Devitt agrees, it would be irresponsible to attribute so much significance to the folk's theoretical beliefs.

³⁶⁶ N. Smith (2000) p.xv

³⁶⁷ Devitt (2006) p.102

4.3.3 A modification to Devitt's Model

Devitt has modified his view that linguistic intuitions are like theoretical beliefs about other aspects of the world in two ways. Firstly, he now stresses that they are most comparable to intuitions we have about the outputs of other human competences such as “touch-typing and thinking”.³⁶⁸ Secondly, he now allows a role for grammatical competence in linguistic intuition. On Devitt's model, a speaker asked about a string of words first simulates the behaviour of attempting to produce or comprehend a string, and in doing so engages their grammatical competence. There is then some quick central processor reflection upon this experience in which speaker's employ their grammatical concepts to arrive at a judgement.

But even this modified version of Devitt's model is inadequate. Smith brings out the problem that remains with Devitt's model using the following example.³⁶⁹

(53) Bill believes that Bush is dangerous.

(54) Bill believes Bush is dangerous.

If we were to ask a speaker, presented with cases like (53) and (54), to do some quick reflection and say whether they believed the *that* in sentences containing *believes* is optional, they would probably say that it was entirely optional. But it is clear that this reflection is not what grammarians are targeting in probing a speaker's intuitions. When we elicit speaker's intuitive responses to strings like (55) we get an intuitive judgement which reveals their grammar but is indifferent to such central processor, theoretical judgements.

(55) *Bill believes that Hilary to be intelligent.

It is the language that speakers are immediately cognitively sensitive to, data of the latter sort, and not the irrelevant theoretical reflections that the generative grammarian is targeting. In plumbing the speaker's intuitions we want to find out

³⁶⁸ Devitt (2006a) pp.593-4

³⁶⁹ Smith (unpublished ms.) p.37

what the speaker can immediately recognise as part of his language, what interpretations he can get and what he will immediately produce. As Longworth puts it:

The subject may be especially well placed to report on how things seem to them, but should not be taken to be authoritative about whether apparent properties are determined by their language systems...In short, the linguist for the most part aims to treat subjects as objects of inquiry, rather than fellow inquirers.³⁷⁰

It is Devitt's false assumption, not shared by the orthodox model, that linguistic intuitions are a speaker's beliefs or opinions about grammatical matters that causes much of his consternation with the orthodox view. He asks us to compare the linguistic case to other cases of cognitive capacities, where a set of rules is somehow encoded in us such as thinking and typing. As Devitt rightly points out, there is no path from the embodiment of these rules in a subject to that subject's having correct *beliefs* about these rules.³⁷¹ Devitt infers that there is no such path from the grammatical rules encoded in the competence system to theoretical beliefs expressed in linguistic judgements. But as Longworth rightly notes, the orthodox model does not treat speakers as authoritative theoreticians about their grammatical competence because it does not treat them as theoreticians, the grammarian's "fellow inquirers", at all. The commitment to a theoretical model of intuitions is Devitt's own: *not* one the orthodox model shares. To be clear, proponents of the orthodox model should agree with Devitt that there is no path from the encoding of rules of grammatical competence in a speaker to their having correct beliefs about those rules. Correct beliefs about the principles of grammatical competence are what grammatical theory aims for, not what speakers are taken to provide for the grammarian. Therefore, the orthodox view creates no mysterious access to the principles that characterise speakers' competences.

³⁷⁰ Longworth (unpublished ms.) p.11

³⁷¹ Devitt (2006) p.118

4.3.4 Devitt's alternative view of the Evidence

Ultimately, Devitt's model of linguistic intuitions leads him to the following conclusion: "we do not generally take theory-laden folk judgements as primary data for a scientific theory. So we should not do so in linguistics".³⁷² Devitt argues that we should not give the linguistic judgements of ordinary native speakers a central evidential role in grammatical theory. Rather he claims we should seek evidence primarily amongst corpuses, what speakers would say and understand in linguistic contexts, and the intuitions of linguists. He says:

The main evidence for grammars is not found in the intuitions of ordinary speakers but rather in a combination of the corpus, the evidence of what we would say and understand, and the intuitions of linguists.³⁷³

Although linguists, like other scientists, have theoretical hunches (a sense of the sort of explanation certain phenomena might receive) that is not what they are interested in probing their native knowledge of language. Theoretical hunches, whatever role they do play in theory construction, are not treated as *evidence*. As Fiengo notes:

[W]e say, perhaps of a linguist, that the linguist has the *theoretical* intuition that that is the analysis which should be given of the sentence in question. The term 'intuition', in this case, has a sense rather like that of 'hunch'. Linguists say they have such intuitions or hunches, but they never constitute the data of Linguistics, rather they apparently occur among linguists during the practice of Linguistics, as they do among physicists during the practice of physics... And on the other hand, my intuition that the sentence 'Flying planes can be dangerous' is ambiguous is nothing like a hunch.³⁷⁴

Moreover, an investigation of the "evidence of what we would say and understand" upon presentation of linguistic material *is* part of the intuitions evidence on the orthodox model. Gathering evidence about what speakers would say or understand upon the presentation of linguistic material *just is* part of probing their linguistic intuitions, via the production and perception of speech. So I'm going to focus on

³⁷² Devitt (2006) p.102

³⁷³ Devitt (2006) p.100

³⁷⁴ Fiengo (2003) p.256

Devitt's suggestion that the corpus should be a central source of evidence for generative grammars.

The suggestion fits with Devitt's linguistic conception of generative grammars as theories of physical entities that form representational systems. For in suggesting that the linguist focus on the corpus rather than linguistic intuitions, Devitt is seeking to re-orientate the grammarian's attention away from speakers' proclivities towards properties of external outputs as the locus of grammatical reality. Devitt thinks the richest source of evidence for grammatical hypotheses is to be found in corpuses. But there are major problems with Devitt's suggestions that the corpus can play the central role in grammatical theory that has been played by the intuitions data.

I noted (§2.1 and §3.2) the problems with a focus on performance events for determining the structures of a speaker's language. For these reasons, speakers' intuitive judgements have been considered crucial data for generative grammars. Chomsky says:

Linguistics is characterised by attention to certain kinds of evidence...largely, the judgements of native speakers.³⁷⁵

Chomsky thinks we cannot determine the grammatical structure of languages on the basis of gathering corpuses of performance events. Why does Chomsky think that the intuitions of speakers are so central to investigating the grammars of languages?

Rather than drawing on corpus data in isolation from speakers' intuitions, corpus data are employed as part of a complementary package with speakers' intuitions. Depending on the hypothesis that one is investigating, one might, for example, want to examine a corpus to see if a certain construction ever occurs or for evidence about what is available in acquisition. As Collins suggests:

For example, take the hypothesis that children don't make 'errors' of a certain kind, say, 'Children don't move auxiliary verbs from relative clauses in the attempt to form interrogatives'...one can look at databases of child speech to test this. One can also look at

³⁷⁵ Chomsky (1986) p.36

adult speech to see how common certain constructions are, or whether children receive 'negative data',³⁷⁶

But it is crucial to hypotheses about grammatical structure that one uses pieces of a corpus in tandem with linguistic judgements, so as to work out how the expressions are actually structured rather than simply whether they occur or not. The mere occurrence of an expression by itself doesn't tell you about its grammatical properties. If one wants to know how it is structured that requires speakers making judgements about its interpretation. To this end, the grammarian's use of corpuses involves him, or his native informant, making intuitive judgements too. Insofar as Devitt's model ultimately suggests a preponderant role for corpus data over speakers' intuitions, it is mistaken.

Devitt rightly points out that linguistics textbooks are full of sample strings described as unacceptable, ambiguous and so forth. These notions have a cognitive ring to them but Devitt thinks they are best interpreted as marking properties of the written strings. When linguists say that English-speakers have the intuition that a string is ambiguous, Devitt thinks that this is employed as evidence that the string has the property of structural ambiguity. But as Devitt thinks that ordinary speakers are a fallible guide to grammatical properties, he suggests that we should go straight to the corpus: straight to the uttered and written strings, which he thinks are the primary focus of grammatical inquiry, and survey their properties rather than the properties of speakers' intuitions.

A corpus is a list of strings that have been uttered or written down.³⁷⁷ This description of corpuses might be challenged on the grounds that it is unnecessarily austere. Perhaps we should think of corpuses as imbued with all sorts of other interesting information about grammatical structure. But it is difficult to see how linguists could so imbue corpuses without drawing on evidence from the judgements of native speakers. Grammatical structures involve special hierarchical dependencies amongst constituents. We can't determine these special structures of

³⁷⁶ Collins (2006) p.7

³⁷⁷ A typical written corpus might be *The Times* or *The Wall Street Journal*. Spoken corpuses are collected in acquisition studies, sometimes capturing everything a child hears and utters over the course of months.

speakers' languages simply by enumerating the strings that they produce, where the latter are flat lists of words. This is one reason why intuitive judgements are so important in gathering evidence for generative grammars: because they can be used to determine the way that speakers structure linguistic material. Consider (56) and (57).

(56) Mary expected to leave by herself.

(57) Bill wondered who Mary expected to leave by herself.

The individual who is leaving can differ between the cases, as we can tell by the acceptability of substituting *herself* for *himself* in (57) but not (56). But the fact that the underlined material has different structural articulation in (56) and (57) is not obvious from looking at the strings themselves without bringing such judgements about their interpretation to bear.³⁷⁸ From looking at this mini-corpus, one might think that (56) is simply embedded as the *wh*-complement in (57) and retains its structure. Examining a corpus may not suffice for determining the difference in the structures. The difference in structure between the two occurrences of the underlined material is usually explained in terms of the difference in the empty categories, where a copy of *Mary* is the subject of the infinitival clause in (56) whilst a copy of *who* fills that position in (57). Linguists determined this by investigating the different interpretations speakers give these strings.

Moreover, an uttered string may have been an “error”, an ungrammatical utterance. If the linguist were to count these strings as part of the language then they would be counting in too much in constructing their grammar. So linguists and their informants make judgements to discern amongst the produced utterances. As Stanley observes: “Ordinary discourse often involves the use of complex expressions which would be counted as ungrammatical even by the utterer’s own lights.”³⁷⁹ Corpora contain no explicit information about which are the ungrammatical utterances, and such information is crucial to developing grammatical theories. The same is true in principle of written strings. We might

³⁷⁸ The example is from Collins (2006) p.8.

³⁷⁹ Stanley (2000) p.408

have reason to think that people are more careful about what they write than what they say. Yet we don't want them to be too careful. There are lots of things that speakers easily recognise as linguistic forms but wouldn't write down. Contractions of "want to" to "wanna" may be one such example. We would miss these forms if we relied solely on written corpus data.

And it would be a waste of time to wait until the strings of theoretical interest happened to just turn up in written or spoken corpuses when the linguist can construct them himself or enlist the judgements of a native informant when the language is not his own. The crucial cases to test the theory may not be in the corpus. Chomsky recognised these problems with relying on corpuses in the absence of judgements early on, saying:

A corpus may contain examples of deviant or ungrammatical sentences, and any rational linguist will recognise the problem and try to assign to observed examples their proper status...insofar as a corpus is used as a source of illustrative examples, we rely on the same intuitive judgements to select examples as we do in devising relevant examples with the aid of an informant or ourselves.³⁸⁰

Further, an uttered or written string, like "Flying planes can be dangerous, that turns up in a corpus, may instantiate *more than one* grammatical structure of a language. Sentences are structured objects but "do not wear their structures on their sleeves, so it can easily happen that distinct structures sound the same."³⁸¹ In such cases, the linguist limited to inspecting corpuses may miss out on structures that are part of the language.

The point is that a corpus contains a great deal of what is, for the purposes of investigating speakers' grammars, linguistic debris. This includes ungrammatical but interpretable utterances, false starts, mistakes, slips of the tongue, half-expressed thoughts, unfinished sentences, interruptions, and utterances affected by deficits in memory, attention and motivation. So corpuses taken in isolation don't provide a perspicuous guide to the linguistic forms that speakers recognise or the construal that they put upon them. Chomsky's suggestion (see Chapter Three) is

³⁸⁰ Chomsky (1965) pp.198-9

³⁸¹ Fiengo (2003) p.255

that to get some explanatory perspective on such a record of performance events, requires not only the further evidence of speakers making judgements but also a distinction between speakers' grammatical competence and the other factors that enter into these linguistic performances.

Devitt's products, the utterances and inscriptions that make up corpora, are an interaction effect amongst which only one factor is grammar. In arguing for the primary importance of corpuses, Devitt misses the point that generative grammarians look to separate out the different factors that contribute to performance events rather than study the properties of the corpus as they result from the motley. That is the point of the theoretical distinction between competence and performance: to try and determine the grammar of the language that a speaker has actually acquired. The broad target is the system responsible for our sensitivity to linguistic form. This system can be explored by probing speakers' intuitive responses to linguistic material (as I described in §4.2).

To summarise the major problems with Devitt's model of linguistic intuition: (1) It is unable to accommodate the pre-doxastic nature of linguistic intuition (§4.3.1). (2) Such folk theoretical judgements would not be afforded an evidential role in a science (§4.3.2). (3) In practice linguists do not target speakers' reflective judgements (§4.3.3). (4) The model has the consequence that a central evidential role should be afforded to corpus data which it cannot bear. Corpus data works as a complementary package with intuitive judgements (§4.3.4).

By contrast, if a major source of evidence for generative grammars is native speakers' pre-doxastic linguistic intuition, then a major source of evidence bears on the grammars that speakers have internalised in a system of grammatical competence. The orthodox model of linguistic intuitions and the evidential role they play looks in good shape.

4.4 Linguistic Intuitions as Observational Judgements³⁸²

4.4.1 The Observational Model

A third model of speakers' intuitive linguistic judgements that we might consider is one according to which they are judgements based on *observation* of non-psychological linguistic facts. On this model, linguistic intuitions are observations of non-psychological facts and the reports formed on these intuitions are akin to observational judgements. So conceived, linguistic intuitions are gathered in much the same way as that in which scientists in other fields gather their observations, namely, by observing facts in our environment. These observations are then employed as data for a theory about grammatical facts located in our environment.

On this observational model, the linguist is not interested in gathering evidence from speakers' intuitions because they can be brought to bear on a theory of grammatical competence. But neither is he plumbing a speaker's theoretical beliefs about language. On the observational model, the linguist collects the linguistic facts that he observes, and those observed by his native informants; facts especially accessible to observation by those that know a language.³⁸³ Grammatical theory then aims for a theory of why those linguistic facts are as they are.

On this model, when all goes well with a speaker, a speaker's intuition that a string is a good sentence of his language is his observation of the fact that this is so. Hence, a speaker's intuition that a string is a good sentence is, unless we have reason to think that the speaker has made an error, evidence that the string is a good sentence of his language.

The key difference between the orthodox model and the observational model is over the form that the intuitions data takes. On the orthodox model, the datum is that *a speaker finds* a certain utterance or written string to be acceptable or

³⁸² Timothy Williamson suggested a view along these lines to me.

³⁸³ The view that grammatical properties are especially accessible to be observed by those that know languages, has a point of contact with McDowell's view (§1.3.3) that meaning properties are especially perceivable by speakers who have a special sort of practical knowledge. The observational model under consideration here, however, makes no appeal to McDowell's notion of a *practice*.

interpretable in certain ways and so forth. On the observational model, the evidence takes the form *that a certain utterance or written string is a good sentence of the language or n-ways ambiguous* and so forth. On this observational view, the speaker's proclivities are not the salient data which is rather the non-psychological fact that the native speaker observes for the linguist. The grammarian's evidence consists in non-psychological facts that suitably equipped speakers can observe and is not to be confused with the data that speakers' linguistic intuitions simultaneously provide about those speakers' grammatical competences.

The proponent of the observational model might try to develop the analogy with observation in the non-psychological sciences in the following way. They might claim that the grammarian is no more seeking to gather data on the properties of a speaker's intuitions as they reveal his psychological organisation, than a physicist is seeking to gather data on the properties of an observer's judgements as they are revelatory of his mechanisms of observation. Making observations in the physical sciences requires an observer whose eyes function properly and who has an ability to read the instruments that he makes use of. Applying the analogy to linguistics, we might think of the native speaker as, so to speak, looking down the microscope at the facts about their language but making use of a special instrument, their grammatical competence.³⁸⁴

³⁸⁴ If, as Devitt holds, the non-psychological grammatical properties are partly determined by psychological facts, then the linguist's access to the grammatical facts about speakers' languages looks rather different to the physicist's observation of physical facts. The linguist's access to the grammatical facts is dependent on her access to speakers in a quite different way to that in which a physicist's access to physical facts is dependent on someone observing the physical facts. The linguist doesn't depend on the speaker merely because she, again so to speak, needs someone to look down the microscope but because the facts in question are *determined* by the psychology of the speakers under investigation. There would be no grammatical facts determined independently of speakers for the observer to observe in the way that there are independently determined physical facts for the physicist to observe. So conceived, if one takes speakers out of the experimental situation the linguistic facts vanish along with their determinants. The physicist needs someone to make observations of the pertinent facts but he doesn't need a human being in situ for there to be such a determinate set of facts. So the analogy works better with a mind-independent conception of non-psychological grammatical properties.

To take an example, suppose the presented linguistic material is (58) which is two ways ambiguous.

(58) I saw her duck and swallow.

On the observational model, the grammarian's data are not that speakers find (58) two ways ambiguous and can't hear it four ways ambiguous, leading us to ask how those speakers must be configured for this to be so. On the observational model, the data relevant to grammatical hypotheses are simply that the sentence is two ways ambiguous, an observation that can be made by English-speakers.³⁸⁵

4.4.2 Error, Belief-Independence and Types

It is a feature of the model that there can be errors in linguistic observation, where performance factors interfere with our capacity to observe the linguistic facts. Proponents of the observational model might claim that this is much the same situation as occurs in non-psychological sciences where an observer's inherent limitations cause them to make various sorts of errors and faulty observations. Though such errors in linguistic observation require explanation, proponents of this model will discard them as evidence for generative grammars, just as a physicist discards his observation when he mistakes a scratch on his particle detector for a real particle.

³⁸⁵ There is a difficult question about the necessity of observation to linguistic intuition. For, on the face of it, there need not be any external material at all presented to the subject for him to consult his linguistic intuition. Much of the data seem to be available to introspection and reflection on the contours of one's own language. And even where some initial observation is involved, once I have observed the string in (58), I can then fast-switch back and forth in my head between the two linguistic structures it instantiates without further observation of external material. In one sense we *observe* the same thing, the string that was presented to us in (58), though the grammatical properties of our experience change. We experience changes in the grammatical properties that we are attending to, though the object of our initial observation stays the same. I will not be able to give this issue the full treatment it deserves here. One thing that proponents of the observational model might suggest is that cases of linguistic intuition where external objects of observation play little or no role are akin to imagining observing something.

The observational model has a significant advantage over the theoretical belief model. The observational model can accommodate the pre-doxastic nature of linguistic intuition by allowing that the mechanisms of linguistic perception, including the grammatical competence system, are encapsulated.

To fill out the observational model a little, we need to say something about what is observed. What is observed in the hypothesised cases would have to be more than the physical entities, their physical properties and magnitudes. For I can hear the words produced by Arabic speakers, or see the marks produced by Chinese speakers, without observing any of the linguistic facts about their language.³⁸⁶ Moreover, in observing that something is a sentence of my own language, I would not be simply observing the familiar sounds or marks but observing it, so to speak, with a tick by it – *as* a sentence of the language that I know. To see or hear that something is a good sentence is to see or hear that it is a certain *type* of thing, a *linguistic type*.

These observed linguistic types would have to be fairly broad categories such that something is or is not a sentence of the speaker's language, that it is ambiguous or unambiguous, that it has a certain interpretation and so forth. It is clear that the fundamental structural properties of languages do not lie open to view to be observed in the strings that are outwardly presented to us. No one observes c-command relations and the other hierarchical dependencies, or the empty categories. Such grammatical properties make up a level of unobservable grammatical structure. The proponent of the observational model would hold that the observed linguistic facts receive their explanation in terms of a level of unobservable grammatical structure.

³⁸⁶ As Smith (2006 pp.949-50) observes: "We have difficulty even recognising the word boundaries in a foreign language since they do not correlate to breaks in the acoustic signal."

4.4.3 Grammatical Variation

There are arguments from different sorts of grammatical variation that might put pressure on the observation model.³⁸⁷ The external materials in which grammatical properties are discerned can vary widely, with much less constraint from their physical instantiation than one might expect. The grammatical relations are instantiated in the speech sounds of the different natural languages, in their different written forms – some of which mark word boundaries, some of which do not – but also across the sign languages, and are even recognised in patterns of vibration by subjects with large scale sensory deficits.³⁸⁸ If linguistic intuition is a matter of observation then it is a form of perception that is invariant across different sensory modalities.

But there is not only variation in the modalities involved in the perception of language whilst the grammatical intuitions remain invariant, there is also variation in grammatical intuitions amongst speakers. This raises a question mark over how analogous linguistic intuition really is to observation in the non-psychological sciences. For it is taken by some to be criteria of genuine observation that the facts observed must not vary from observer to observer. The thought that genuine observations must be such that they could be made by any observer, independent of their particularities, is clearly articulated by Tyler Burge:

Any observer could have been equally well placed to make any observation. Others could have made an observation with the same type of presentation of the scene, if they had been in the same position at the relevant time... Even though empirical commitments must be made by persons, nothing relevant to the justification of any empirical commitment regarding the physical world has anything essentially to do with any particular person making the commitment.³⁸⁹

However, the apprehension of grammatical properties in linguistic intuition cannot be enjoyed by all observers. Not all subjects are “equally well-placed”³⁹⁰ to intuit

³⁸⁷ See, for instance, Rey’s (2006) argument against linguistic realism.

³⁸⁸ Chomsky (2000) pp.121-2

³⁸⁹ Burge (1988) p.475

³⁹⁰ *ibid.*

particular grammatical properties. There are grammatical properties that English-speakers intuit upon presentation of the sentence “John shaved himself” that will not be available to speakers with their parameters set to Chinese upon “the same type of presentation of the scene, if they had been in the same position at the relevant time”. The grammatical properties that speakers can intuit are dependent upon the “particular person making the commitment”, for as Smith notes “Only those with a certain competence will have a richly articulated linguistic experience in response to the natural speech sounds of Chinese or Arabic speakers.”³⁹¹ The grammatical structures of a particular language, and their instantiation in speech sounds, may be available to only a very small proportion of speakers. We know this from our own experiences, and equally linguists require native speakers to intuit the properties of a language unless it happens to be their own.

So the intuited grammatical properties *vary* across speakers as their *particularised* grammatical competence systems vary. The dependence of which properties are intuited on highly specialised cognitive organisations might suggest, if we accept Burge’s criterion, that linguistic intuition could not be observation. The fact that so much more is “experienced in speech sounds by those that know the language than those that don’t” has suggested to some linguists that the words, phrases and sentences that we intuit are inner, mental objects.³⁹² Harris and Lindsay conclude from the dependence of intuited grammatical properties on our special psychological systems that what is experienced by one who knows the language is not something that is there in the sounds emitted by its speakers, rather it is “projected by means of articulations but is not embodied in them. The linguistic information read into, or onto, those sounds is simply part of ‘the specifically human way with sounds.’”³⁹³

There are, however, means by which the observational model of linguistic intuition might be sustained despite grammatical variation. As Longworth points out, the argument from variation in the intuited grammatical properties to the conclusion that linguistic intuition is not a matter of observation is “a reflex of

³⁹¹ Smith (unpublished ms.) p.21

³⁹² Smith (2006) p.951

³⁹³ Harris and Lindsay (2003) p.203

sensitivity to the modal dependence of the range of manifest grammatical properties, available to each of us, on the particularity of our language systems.”³⁹⁴ If we could show that the modal dependence of the intuited grammatical properties on the particularity of our competence system is at least consistent with the mind-independence of grammatical properties then the observational model could be sustained in the face of grammatical variation.

There are two broad ways of understanding Burge’s condition on observation. One could interpret it so that it is a condition on observation that an individual needn’t be the *particular* person who made the observation; it was available in principle to other individuals, though those other individuals would have to be of the same physical or psychological type. But a stronger interpretation is that it is a condition on observation that for any genuine observation, an individual needn’t be of the same physical or psychological type in order to make the observations, beyond the minimal requirement that they have observational powers at all. On the first interpretation Burge would be laying down a very plausible condition on observation. But if we combine it with variation it doesn’t suggest anything about whether there is observation taking place in linguistic intuition: it would be unsurprising if different subjects with the same physical and psychological constitution were capable of making the same observation whilst differently physically or psychologically constituted subjects had different observational powers.

But the second interpretation of Burge’s condition yields a far less plausible condition. Why should the observational be restricted to that which could be observed by anyone however constituted? This would mean that colour was not observable as one could have minimal observational capacities but lack colour perception.³⁹⁵

What is required to make variety and observation consistent is to drop this latter criterion for genuine observation and adopt a *pluralist* view of the observed

³⁹⁴ Longworth (2007) p.407

³⁹⁵ This condition has been called the “Martian Principle” by Travis (2002), it is targeted by Kalderon’s work on colour pluralism (see Kalderon 2007) and Longworth’s work on grammatical properties (2007, unpublished ms.)

grammatical properties. According to a pluralist view of grammatical properties, the environment sustains an abundance of grammatical properties; and different properties are selected by the different psychological systems of the speakers whose intuitions conflict. On the pluralist view, a given environmental item can sustain a plurality of grammatical properties and the instancing of a set of grammatical properties by an item does not serve to exclude its instancing other sets of grammatical properties. On such a view, the possibility of genuine conflict arises only where an environmental item is required to sustain incompatible properties drawn from the same family of grammatical properties; for it is amongst these families of properties that relations of exclusion hold. If amongst this plurality of grammatical properties there is a sufficient range then instances of variation could be explained in terms of what is available in the environment for different speakers to perceive on the basis of their “special psychological design.”³⁹⁶

If a pluralist view of grammatical properties could be sustained then the observational model would not be inconsistent with grammatical variation. So, although some have made a case against the observational model on the basis of the variety in speakers’ linguistic intuitions, that case is incomplete. Pluralism offers a defence of observation of non-psychological properties given variety. The next step in my argument is to determine which sorts of properties make up the grammarian’s explanans: the properties of grammatical competence revealed by speaker’s intuitions as per the orthodox model or properties observable in the external environment as per the observational model.

But two things are worth noting before we proceed. First, the possibility of an observational model in no way impugns the orthodox model, according to which speakers’ linguistic intuitions can be used to investigate the competence-performance distinction, a distinction which I argued (Chapter Three) the grammarian is required to investigate. Secondly, as on the observational model, which grammatical properties speakers observe is dependent on the particular configurations of their grammatical competence system, to determine what the

³⁹⁶ See Travis (2002) for this idea. See Longworth (2007, unpublished ms.) for discussion in the context of grammatical properties.

grammatical properties of speaker's languages are, the grammarian would still need to investigate the properties of their particularised competence system.

What we need to determine is what sort of properties – properties of a grammatical competence system or properties of a non-psychological domain - the generative grammarian is focused upon and are appealed to in grammatical explanations. It is this crucial issue to which I'll now turn.

5. Explanation in Generative Grammar

5.1 What sort of properties do Generative Grammars appeal to?

In the last chapter I defended an orthodox model of the linguistic intuitions evidence. Linguistic intuitions can be employed as evidence for investigating the competence-performance distinction as a means to constructing a theory of grammatical competence. I rejected an alternative model according to which the intuitions data should be treated as speaker's theoretical beliefs about language. On the orthodox model, grammarians are interested in linguistic intuitions because of what these intuitions reveal about speakers and their psychological system of grammatical competence. I also considered an observational model of linguistic intuitions, according to which linguistic intuitions serve as observations of linguistic facts located in the external environment and are used to support hypotheses about that non-psychological domain.

In order to determine which of these models is most appropriate to generative grammar, I will examine the question of what sorts of facts and properties generative grammars make explanatory appeal to in constructing their theories. Are they *psychological* properties of a speaker's grammatical competence system, as per the orthodox model? Or are they *non-psychological* properties of the external environment that might be observed by competent speakers? Or do grammarians require some *combination* of psychological and non-psychological grammatical properties in order to make their explanations work? In this chapter, I consider the theoretical constructs of generative grammar. I argue that the sorts of properties that must be part of the grammarian's explanans, if he is to meet the explanatory goals defended in Chapter Two, are psychological properties of the grammatical competence system, which I argued in Chapter Four can be investigated via speakers' intuitions. I claim that the grammarian's goals can be met

by a theory of grammatical competence to which grammarians make an explanatory commitment (as I argued in Chapter Three). Further, I argue that with respect to these explanatory goals, non-psychological grammatical properties external to our grammatical systems would be *explanatory danglers*, which would commit the theories to additional properties for no corresponding explanatory benefit.

5.1.1 Explanation and Supervenience

Though Devitt argues that generative grammar describes a non-psychological reality realised in our physical environment, he does not think that the reality the grammarian describes is mind-independent. Devitt is attracted to the idea that if there were no minds then there would be no c-command, binding principles, lexical features and so forth. He thinks that minds provide some of the unity to those entities which we categorise as c-commanders, pronoun and reflexives. So Devitt wants to accept that grammatical properties have psychological determinants. But he distinguishes between grammatical properties having psychological *determinants* and their *being* psychological properties. Devitt's view is that non-psychological grammatical facts supervene on a combination of psychological facts, social facts and environmental facts.³⁹⁷ So Devitt's "linguistic reality" is *mind-dependent*, though it constitutes a *non-psychological* domain of facts. In defending his linguistic conception, Devitt points out that this dependence on psychological determinants, "does not make linguistic facts psychological."³⁹⁸ Devitt argues that:

Even if symbols had their properties in virtue of certain mental facts that would not make the theory of those symbols about those facts and so would not make the theory part of psychology. Indeed, consider the consequences of supposing it would, and then generalizing: every theory – economic, psychological, biological, etc. – would be about physical facts and part of physics because physical facts ultimately determine everything. A special science does not lose its own domain because that domain supervenes on another.³⁹⁹

³⁹⁷ Devitt (2006) p.39

³⁹⁸ Devitt (unpublished ms. b) p.10

³⁹⁹ Devitt (2006) p.40

This is an important point because a number of Devitt's critics have highlighted the apparent dependence of grammatical properties on the psychological systems of speakers and pressed Devitt's view on this basis.

Smith argues that what transforms the physical entities (sounds and marks in our environment) into a representational system is not something located in the environment but something internal to the speakers who encounter the sounds and marks.⁴⁰⁰ Smith claims that a sound having the grammatical properties it does "consists in there being a relation between the sound and features of a speaker's mind/brain." Smith suggests that if the sounds do not have their grammatical properties in and of themselves, but only as they stand in a relation to speakers' minds, then facts that determine the "fine-grained linguistic detail and richness" the grammarian wants to capture, are psychological facts. Smith then questions:

Why not then concede that it is these domain-specific features of a speaker's cognitive organisation...that the linguist most needs to focus on. Sure, we can say that physical tokens of sound *have* linguistic properties, but crucially they do so because they stand in important relations to the psychological states of language users.⁴⁰¹

Smith is challenging Devitt's linguistic conception on the grounds that the major determinants of grammatical properties are on the psychological side of Devitt's distinctions. So, we might then wonder, as Smith does, why the grammarian should focus on what's going on over on the non-psychological side of the distinctions. Even if one grants Devitt's distinctions then one might think it reasonable for a linguist to focus on one side of them, the side where all the action is.

Devitt denies that the grammatical facts are determined solely by psychological facts because he thinks there is also a role for social and environmental facts. In §5.4 I offer a nativist argument against Devitt's claims concerning the role of social and environmental facts in determining the grammatical properties of speakers' languages. But Devitt's main line of defence against Smith's argument from the supervenience of grammatical facts on psychological facts, is to point out that even if grammatical properties *are* so

⁴⁰⁰ Smith (unpublished ms.) p.27

⁴⁰¹ Smith (2006a) p.11

determined “this would not make linguistics part of psychology any more than a similar determination makes psychology a part of physics.”⁴⁰²

Devitt thinks this challenge rests on a wrong-headed philosophy of science. Even if all the determination is on the psychological side of his distinctions, Devitt maintains that linguistics could still be about non-psychological facts. This is because the supervenience of one domain of facts on another does not suggest that we have only one domain of inquiry rather than two. If it did then, on the assumption that everything supervenes on the physical, we would have just one domain of inquiry: namely, physics. But all parties agree on the existence of the special sciences, so the argument from supervenience doesn't go through.

Devitt's general point here is sound.⁴⁰³ We cannot infer from the supervenience of grammatical facts on psychological facts alone to generative grammar being a part of psychology, any more than we could infer that the psychological sciences were a part of physics from the supervenience of the psychological facts on physical facts. Devitt argues that his critics mistakenly infer from the supervenience claim to a claim concerning linguistic explanations and what sorts of facts they traffic in. As Devitt notes, there is no such entailment. There are many scientific domains that supervene on lower levels. Ultimately, we might think they all supervene on physics. But that does not imply that these sciences have no proprietary domain of laws and generalisations or that they are all really about physical facts. Like Devitt, his critics are committed to the existence of the special sciences, such as psychology and linguistics.

⁴⁰² Devitt (2006a) p.11

⁴⁰³ Although I would question whether Devitt's critics, such as Smith, are making the straightforward supervenience argument, Devitt is right that such an argument would not go through. The supervenience of one domain of inquiry on another does not suggest we have only one domain of inquiry and not two. It is unlikely, however, that his opponents would commit to any such general principle. The suggestion is rather that in the case of grammar the revealing generalisations can be captured in psychological terms. Though Smith, in particular, presses on the supervenience claim, his argument (2006a, unpublished ms.) ultimately rests on the *explanatory* credentials of the proposed non-psychological properties. His focus is on whether the explanatory generalisations of generative grammar apply to Devitt's non-psychological phenomena or to a level of psychological organisation.

As Devitt suggests, questions about the proprietary domain of a science are not so easy to answer and cannot be inferred from supervenience claims. But we might try to answer such questions by making *explanatory* claims rather than *supervenience* claims. If the science of generative grammar states explanatory generalisations that can all be captured in psychological terms then we might wonder whether there is an explanatory domain of non-psychological grammatical properties. If such an argument could be supported *then* we would have good reason to think that generative grammar is a study of the psychological facts which determine the grammatical facts. This argument could go through even if we grant Devitt his point about supervenience and special sciences. We might hold that there is a special science of psychology, that captures generalisations that cannot be stated in the terms of physics or biology, and at the same time that generative grammar is a part of this special science of psychology.

What Devitt needs to argue is that there is a failure of *explanatory* determination between psychology and generative grammar.⁴⁰⁴ It might be true that everything supervenes on physics and yet also true that there are independent sciences of biology, psychology, geology and so forth, whose generalisations cannot be conjured from the resources of physics. But that doesn't in itself provide any reason to think that there are interesting grammatical generalisations that cannot be stated in psychological terms as generalisations about speakers that ought instead to be stated as generalisations about a non-psychological domain. I'll argue that Devitt has provided *no* reason to think that grammatical reality has an explanatory integrity that is lost at a psychological level of explanation.

It is important to emphasise then that my claim is an *explanatory* claim about the sorts of properties appealed to in grammatical explanations. It is a claim about what sorts of properties are required by theories that meet the generative grammarian's goals. My case for the psychological conception of generative grammar doesn't rest simply on the supervenience claim that grammatical properties supervene on psychological properties, but rather on the explanatory claim that it is psychological properties that do the explanatory work in grammatical theories.

⁴⁰⁴ Longworth (forthcoming) makes this point.

5.1.2 Psychological Reality, Linguistic Reality and Grammatical Explanation

A number of philosophers have raised concerns about Chomsky's proposal that the structures described by grammarians are realised in the mind; that they have "psychological reality".⁴⁰⁵ Chomsky's attitude to this question builds on his views about the psychological goals of grammatical theory and his employment of a competence-performance distinction amongst psychological machinery to determine the generative grammars of speakers' languages. Against this background, Chomsky appeals to a general scientific realism to defend his claims about the psychological realisation of the structures described in generative grammars.

Chomsky's thought is that the psychological targets of the theories (descriptive and explanatory adequacy) and the adoption of the competence-performance distinction ensure the "psychological reality" of the theory's posits, insofar as the theory is a true theory:

[W]e impute existence to certain mental representations and to the mental computations that apply in a specific way to these mental representations... We attribute "psychological reality" to the postulated representations and computations. In short, we propose (tentatively, hesitantly, etc.) that our theory is true.⁴⁰⁶

Chomsky notes that there is no hesitation in according "physical reality" to the theoretical constructs employed in physics to explain physical phenomena. Equally, he claims that there should be no question about according "psychological reality" to the theoretical constructs postulated in generative grammar, given that grammarians attempt to explain psychological phenomena: our best ideas about what is psychologically real being simply a reflection of our best theories of psychological phenomena like grammatical competence and its acquisition.

⁴⁰⁵ See, for example, the discussion of Searle's view in §1.2.

⁴⁰⁶ Chomsky (1976) p.3

As such, it is reasonable for Chomsky to be unmoved by scepticism about whether the structures described by generative grammars are “psychologically real”. If the theory has such psychological targets, then such scepticism is either general scepticism about scientific realism or it is unwarranted. Chomsky likens such scepticism to doubting an astrophysicist who develops a theory of the interior of the sun on the basis of observations of light at the sun’s periphery. Our sceptic might question the astrophysicist about the “physical reality” of the structures that his theory describes. He might accuse the astrophysicist of having explained the available evidence for his theory but having failed to provide any further ratification that the structures he describes have any “physical reality” in the sun. As Chomsky argues, this would amount to no more than a roundabout way of questioning whether the astrophysicist’s theory really was a good explanation of the evidence because:

It is senseless to ask for some other kind of justification for attributing “physical reality” to the constructions of the theory, apart from considerations of their adequacy in explaining the evidence and their conformity to the body of natural science, as currently understood. There can be no other grounds for attributing physical reality to the scientist’s constructions.⁴⁰⁷

As Chomsky argues, given the psychological objectives of generative grammar, his position with respect to “psychological reality” is not different in kind to the astrophysicist’s with respect to “physical reality”. The only substantive difference, as far as the reality of their respective posits goes, is that the astrophysicist is “actually postulating physical entities and processes, while [grammarians] are keeping to abstract conditions that unknown mechanisms must meet.”⁴⁰⁸ Though the grammarian might suggest actual physical mechanisms, Chomsky thinks “it would be pointless to do so in the present stage of our ignorance concerning the functioning of the brain.”⁴⁰⁹

⁴⁰⁷ Chomsky (1976) pp.4-5

⁴⁰⁸ Chomsky (1976) p.9

⁴⁰⁹ Chomsky (1980) p.197

Stich argues that attributing a psychological realisation of grammatical structure to speakers is little more plausible than attributing an internal realisation of the laws of physics to a projectile when those physical laws hold true of its behavior.⁴¹⁰ By analogy, Stich supposes that the grammatical laws do apply to speakers, and are *true of* them just as the laws of physics are true of the projectile. But Stich argues that the grammatical laws do not characterise any internal structure to the organism, they are only *true of* the organism.

The difference between the cases is that the grammatical laws, postulated to explain speakers' intuitive judgements, only apply to an organism and explain the data insofar as that organism is postulated to have an internal realisation of the principles of grammar.⁴¹¹ The grammatical laws of a particular speaker's language explain his linguistic judgements and so apply to him, or are "true of" him in Stich's phrase. But they do not apply to a rock, a chimpanzee or a speaker of a different language, so it is necessary to attribute to the speaker some internal linguistic structure not possessed by a rock or a chimpanzee. This is what is meant by an internal realisation of grammar. We do not have to attribute an internal realisation of the laws in the case of the physical laws which apply to all physical objects and can explain their behavior without being realized as internal structure.

So the analogy with a projectile is misconceived because, at least as far as Stich's example goes, there is no requirement that the projectile has a particular internal structure in order for the laws of physics to apply to it. However, we can imagine situations in which we would need to attribute special internal structures to the projectile. If we obtained evidence that the projectile's course was being adjusted, honing in on a particular target, or moving according to complex patterns, we might need to postulate that the projectile possessed some special internal structures in addition to its adhering to the laws of physics. We might then try to determine the character of this internal structure by modifying the projectile's environment, if this were possible, and seeing how the projectile's movements vary so as to make a theoretical inference to its internal structure.

⁴¹⁰ Stich (1985)

⁴¹¹ Chomsky and Katz (1974) p.362

The existence of a dedicated grammatical system as a part of human psychology, its nature and its role in cognition are all empirical issues. But no particular problem has been raised about the “psychological reality” of grammatical structure that does not depend on the kind of unwarranted methodological dualism discussed in §1.2. Moreover, there is a de facto argument that, as Collins puts it, “the best way, at present, of studying linguistic cognition is just to do linguistics as currently practiced”, i.e. within the methodological framework Chomsky has proposed.⁴¹²

So given the cogency of the theoretical inference from speakers linguistic intuitions and behaviours to an internal structure for grammar, we can return to the question at hand: is this internal structure, the system of grammatical competence, the subject matter of generative grammars as Chomsky claims it is?

Devitt argues that the strategy Chomsky proposes for learning about linguistic cognition is not at odds with his linguistic conception of generative grammar. But he proposes that non-psychological grammatical properties are the grammarian’s primary focus and have an “explanatory priority over a theory of the psychological reality underlying language.”⁴¹³ As described in §1.3, Devitt argues that generative grammars are better interpreted as being about something other than the psychological structure of speakers. He claims that they are best interpreted as being about a non-psychological domain of grammatical facts, which he calls a *linguistic reality*. On the linguistic conception, the grammatical reality that generative grammarians target is not a psychological reality underlying the sounds and marks we produce but rather consists in properties that are realised in the sounds and the marks themselves. Devitt draws his distinctions I – III in order to separate psychological and non-psychological linguistic phenomena. He claims that the properties appealed to in generative grammars, the explanans of the theories, fall on the non-psychological side of the divide.

For Devitt’s argument to be convincing, he needs support for the claim that the explanatory work in grammatical theory is done by the properties of a non-psychological domain. Then we would have a motive for thinking that grammatical

⁴¹² Collins (2007) p.2

⁴¹³ Devitt (unpublished ms. b) p.18

theory falls on the non-psychological side of Devitt's distinctions and, hence, that we should adopt a non-psychological conception of generative grammar over Chomsky's psychological conception.

Devitt claims to adduce just such reasons. He claims that non-psychological grammatical properties are required as part of grammatical theories that aim to meet descriptive and explanatory adequacy. And further, he claims that non-psychological grammatical properties are part of the explanation of linguistic communication (a view that I'll consider in §5.5).

But Devitt also makes a claim for the fundamental explanatory precedence of non-psychological grammatical properties over the psychological properties that Chomsky claims to be describing. Devitt asks: "How could we make any progress studying the nature of competence in a language unless we already knew a good deal about the language [conceived non-psychologically]?"⁴¹⁴ Devitt's claim is that we could not investigate a speaker's grammatical competence without first knowing what the grammatical properties of his language are. For, Devitt asks, how else would we know which grammar a speaker is competent *in*?

Given my argument in Chapter Three, Devitt has things back-to-front. The grammarian tries to ascertain the grammar of a speaker's language *by* investigating his grammatical competence. In doing so, he need make no prior theoretical claims about the language a speaker is competent in. We do not know by what prior means he might do so. He simply probes the speaker's grammatical competence and constructs a theory of the language that the speaker knows. There is no mystery to this. I have argued that the grammarian requires a theory of grammatical competence to construct a generative grammar for a speaker's language. To obtain a revealing theory of the speaker's grammar, grammatical theories must aim to go beyond descriptive coverage of the speaker's intuitions and be justified as an account of the grammar a speaker has acquired. Within the methodological framework I've defended, theorising about a speaker's language is not a task that is fruitfully undertaken antecedently to theorising about acquired grammatical competence. If this is correct then there is no basis to Devitt's argument that we

⁴¹⁴ Devitt (2006) p.29

must first have a theory of a speaker's language before we can develop a theory of their grammatical competence.

So proponents of the psychological conception should be unmoved by Devitt's claims to this effect. Their view is that if we follow Devitt in construing "language" and "the syntax of sentences" so that they are phenomena external to the grammatical competence system then such phenomena are neither targeted, nor presupposed, by generative explanations. In what follows I aim to adjudicate between psychological and non-psychological conceptions of the explanatory posits of grammatical theory and determine the correct conception of the grammatical reality that grammarians describe. My argument is that generative grammarians need a competence-performance distinction and that the associated theory of acquired grammatical competence serves the goals of meeting descriptive and explanatory adequacy (§5.3 and §5.4). I argue that, as a consequence, the proposed non-psychological grammatical properties become dispensable to grammatical theory, they are "explanatory danglers". I then consider and reject Devitt's argument that non-psychological grammatical properties are indispensable to an explanation of communication (§5.5).

5.2 The No Violence Principle

Before developing my explanatory argument, I'd like to assess Devitt's claim that his linguistic conception is consistent with the empirical status and results of generative grammar. For if Devitt's linguistic conception of generative grammar flouts the No Violence Principle described in §1.3, then we have a clear argument against that conception. An account of what generative grammar is about ought not to impugn its empirical successes. If Devitt's conception does, then we have an inference to the best explanation for the psychological conception (assuming for the moment that these two options are exhaustive). But a number of difficulties suggest themselves when we try to construe the explanatory properties described in generative grammars as properties of Devitt's "products". These difficulties have suggested to some proponents of the psychological conception that the properties

described in generative grammars cannot be realised in the physical environment as Devitt's conception presupposes.

On Devitt's linguistic conception, the grammarian describes properties that are realised in physical sounds and marks. They are high level, relational properties of the sounds and marks but are realised in the physical environment. In claiming that they are high level, Devitt is claiming that they are not brute physical properties. What makes them relational for Devitt is that they are not intrinsic to any physical symbol but rather are determined by structural relations between the symbols and the relations in which these symbols stand to psychological states.

To take an example of what Devitt has in mind, a grammatical property that is realised by structural relations amongst linguistic elements is the property of c-command (described in §1.1). Devitt's claim that c-command is high level is motivated by the fact speakers recognise the property in multiple and heterogeneous physical elements. C-command relations can be discerned amongst a wide range of elements in the physical environment. This is apparent when one considers all the conceivable acoustical signals, orthographies, hand gestures, and other entities in which we might discern such tree-like structures. As the Saussurean linguist Hjelmslev put it, speaking of the structure of Danish:

Danish when spoken, Danish when written, Danish when telegraphed by means of the international flag code of the navies, is, in all these cases essentially one and the same language, and not essentially four different languages. The units of which it is composed differ from one of these cases to another, but the framework of relations between these units remains the same and this is what makes us identify the language...the actual manifestations of the framework are immaterial to the language.⁴¹⁵

Devitt accommodates this fact by conceiving of grammatical properties as high level properties realised in the variety of physical symbols. What make an element a c-commander, for Devitt, is that it is part of a representational system partly determined by relations to psychological states. Being a c-commander is a property an element has in virtue of relations it stands in to other linguistic elements in the system, or "symbols" in Devitt terms. So, on Devitt's view, grammatical properties

⁴¹⁵ Hjelmslev (1985) p.164

are relational properties of sounds and marks; properties of the way in which the symbols that make up the representational system are arranged in the physical space of uttered noises and inscriptions.

Devitt's grammatical properties of symbols are *configurational* properties. A configurational property of a constituent is one that is determined by the constituent's position with respect to other constituents. One constituent c-commanding another is a configurational property because it is determined by the respective positions of the elements in a hierarchical structure. Devitt's view is that these configurational properties are realised in "the very marks on this page" or the corresponding sounds, when we utter the sentences.

Amongst Devitt's opponents there are those who deny that properties like c-command could be realised in the sounds and marks we produce. Rey is sceptical that grammatical structures are environmentally realised:

[W]hat thing in space and time possesses such structure? Not, evidently, any *noises* anyone makes: none of the wave forms produced by people when they speak have a tree structure in the way that, for example, a real tree, or river, or network of neurons might.⁴¹⁶

Upon consideration of the kinds of properties that generative grammars appeal to, Smith draws much the same conclusion as Rey: that generative grammarians are not talking about properties realised by physical entities.

The physical sounds we produce have none of this articulation. They do not even include word boundaries or the distinct articulation of phonemes that enable us to identify words.⁴¹⁷

If it is true that the physical sounds we produce could not realise such grammatical articulation as linguists describe then this would serve to undermine Devitt's linguistic conception.

⁴¹⁶ Rey (2006) p.5. Rey's own view is that nothing in the world – physical or mental – realises linguistic structures. Rey thinks standard linguistic entities are intentional inexistents like Pegasus or Hamlet.

⁴¹⁷ Smith (unpublished ms.) p.19

If generative grammars appeal to properties that cannot be sustained by external speech and written systems, then a natural alternative is to locate the required structure that generative grammars postulate in the mind of speakers. And here we have an argument for the psychological conception of generative grammar according to which the relevant grammatical structure is cognitive structure, involved in assigning structural descriptions to sounds and marks. This argument against Devitt can be formulated as follows.

- P1: Grammatical structures $S \dots S_n$ are essential to generative grammars correctly characterising languages. (For these structures are part of the identity of sentences that occur in natural languages.)
- P2: These structural properties $S \dots S_n$ described by generative grammars cannot be realised in the physical environment.
- P3: The structural properties described by generative grammars are either realised in the physical environment or they are only realised in the part of the speaker's mind responsible for grammatical structure.

From P2 and P3,

- C1: The structural properties described by generative grammars are realised in the speaker's mind.

From P1 and C1,

- C2: The psychological conception of generative grammar according to which grammars describe mental structures engaged in speaking and understanding is true.

Support for the crucial premise P2 is supposed to come from the fact that when we examine the noises that people produce, considering them as physical events, they lack the required articulation to be the realisers of grammatical structures. Though speakers find grammatical structure in sounds and marks, so the argument goes, the more fine grained grammatical properties that explain the structures that speakers intuit, are not realised in the sounds and marks considered as physical events.

Devitt's response is to deny P2. He admits that grammatical properties, like c-command, cannot be determined from inspecting the physical properties of

sounds and marks, but he denies that this implies grammatical structures are not realised by the sounds and marks. Devitt's move is afforded him by his further claim that grammatical properties are high level, relational properties. Devitt claims that the grammatical properties cannot be read off the physical occurrences because the structures relevant to grammarians fall into higher level categories, partly determined by psychological states, which include such hierarchical dependencies amongst constituents. The argument assumes that for grammatical properties to be physically realised there would have to be type correlations between grammatical properties and the lower physical levels; an assumption which would be hard to justify.⁴¹⁸ Devitt compares high level grammatical properties to the non-obvious, relational property of being Australian. The property of being Australian is instantiated by heterogeneous physical objects and cannot be determined from an object's physical characteristics, though there is a unity at a higher level amongst all those physical entities that fall under the category *Australian*.

Perhaps the most difficult cases for Devitt's claim that the grammatical properties are realised in a non-psychological domain by the physical sounds and marks, arise on consideration of the empty categories, such as PRO (discussed in Chapter Four).⁴¹⁹ These are constituents of grammatical structures generative grammars posit which have no phonological properties at all. Hence, it is natural to think they have no realisation in sounds and marks, even as those sounds and marks stand in relation to each other or to psychological states.

As described in Chapter Four, the subjects and objects of English infinitival clauses are thought to be such empty categories, as in (59) and (60).

(59) Simon is easy [_{e1} to please _{e2}]

(60) Simon is eager [_{e1} to please _{e2}]

We need the empty categories in these structures to explain why (61) and (64) but not (62) or (63) are acceptable to English-speakers.

⁴¹⁸ See Fodor (1974).

⁴¹⁹ See Collins (unpublished ms. a, 2007a, 2008a) and Smith (2006, 2006a, unpublished ms.) for arguments from Copies and PRO. See my Appendix for discussion of Copies.

- (61) It is easy to please Simon.
- (62) *It is eager to please Simon.
- (63) *Simon is easy to please Jim.
- (64) Simon is eager to please Jim.

In (59) *Simon* has moved out of the object position and hence *Simon* can replace the empty category as in (61). But *Jim* cannot as in the unacceptable (63). In (60) *Simon* has moved out of the subject position so *Simon* cannot take up object position as in the unacceptable (62) but *Jim* can as in (64). Devitt must accommodate the empty categories if he is to adhere to the No Violence Principle. So he has to explain how the empty categories, constituents that have no phonological properties, are realised in sounds and marks to sustain his account.

The argument against Devitt's linguistic conception that exploits the nature of PRO is supposed to run as follows:

- P1. No Violence Principle: An account of what linguistics is about ought not to jeopardise its explanatory successes.
- P2. The explanatory successes of generative grammar are dependent upon positing empty categories.

From P1 and P2,

- P3. An account of what generative grammar is about must explain how empty categories can figure in generative grammar.

At this point there are two possible continuations of the argument. The first goes via P4a and concludes with Ca.

- P4a. Devitt hasn't sufficiently explained how empty categories can figure in what generative grammar is about.

Therefore,

- Ca. Devitt hasn't done enough to show that his account of what generative grammar is about is acceptable.

The second stronger continuation goes via P4b and concludes with Cb.

P4b. We cannot explain how PRO can figure in what generative grammar is about on any account according to which generative grammars are about the physical environment.

Therefore,

Cb. No such account of what generative grammar is about is acceptable.

There is a case for Ca on the basis of the *prima facie* considerations about PRO raised above. But it seems to me that it is far easier to make a case for Ca than a convincing case for Cb. The case for Ca can be made as follows. There are grammatical properties possessed by the constituents of linguistic expressions; these properties include properties like carrying singular number, being a nominal and so on. According to Devitt, grammatical properties are realised configurationally by physical entities that form representational systems. But the empty categories are not themselves constituents that are realised by elements of utterances or written strings since they are phonetically null. So, it is difficult to see how they are realised configurationally in the physical environment. Hence, Devitt appears to have a problem on his hands with the empty categories and needs to offer some further explanation beyond the materials he has offered thus far.

But it would be unsatisfactory to leave things here. Unless there is some reason for thinking that proponents of views like Devitt's couldn't easily fix or supplement their view in order to accommodate the empty categories, then we haven't advanced much of an argument against the linguistic conception so much as presented some challenging cases.

The challenge for Devitt is that on his account grammatical elements are realised, somehow or other, by physical entities. This must be true of *all* grammatical elements; otherwise the linguistic conception cannot accommodate the full range of generative grammar's results. So Devitt must account for how the full complement of grammatical properties of the phonetically real elements are realised in the sound stream, and whatever problems he has in doing so will be more acute where the elements are null so are not even phonetically articulated by any of the elements of physical strings. Though empty categories seem especially problematic

for someone who holds a view like Devitt's, no arguments have been presented that categorically rule out placing the empty categories in our physical environment.

Here's one way that Devitt could elaborate his view. He says that the grammatical properties of symbols have complex determinants. The symbol's place within a representational system is partly determined by relations it bears to psychological states. But the grammatical properties the symbol has are also determined by its configurational properties determined by its relations to other physically realised symbols. Notice that these two aspects of Devitt's view do not commit him to *each* constituent of a grammatical structure mapping onto a discrete element of the physical occurrence that realises that structure. Otherwise Devitt would be committed to the view that there are type correlations between the higher grammatical levels and lower physical levels, but this is not generally true with the special sciences, so the relevant notion of realisation is not type correlation. Perhaps there could be constituents of grammatical structure that are realised by relations amongst phonetically overt elements though they themselves are not phonetically overt. Devitt gestures at such a possibility in a reply to his critics, when he says:

There is nothing to prevent one of those other properties, arising from a particular arrangement, being one that the larger part has *as if* it had a part with a certain property even though it does not in fact have a part with that property.⁴²⁰

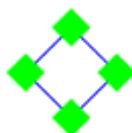
In this passage Devitt also says that it is possible for there to be grammatical elements that are not overtly realised because "there is nothing to stop there being a convention of this sort."⁴²¹ Devitt's view that grammatical properties are determined by convention raises further problems which I'll discuss in §5.4.⁴²² But without appealing to conventions, here is a rough suggestion on Devitt's behalf about how he might develop his account to accommodate the empty categories. The

⁴²⁰ Devitt (2006a) pp.599-600

⁴²¹ Devitt (2006a) p.599

⁴²² In my view, Devitt actually weakens his case by claiming that the realisation of empty categories in physical elements should be explained by appeal to 'grammatical conventions' whereby the multiple copies and empty categories in grammatical structures are conventionally associated with strings. He himself is clear that he has no explanation for how this could be so (Devitt 2008).

issue with the empty categories is that they appear not to be realised by any element of a physical string. A natural suggestion, then, might be that empty categories are not realised by any particular element of the sound stream or written text but instead by relations amongst a collection of produced symbols taken as a whole. By way of analogy, consider a square. Then consider the amodal completion in the diagram below as it contrasts with a full square.⁴²³



One can see that with the lines completed appropriately, the diagram would realise a square. We might think of these missing corners as occluded in the diagram by the smaller squares. Although no parts of the physical lines in our diagram correspond with the corners of a square, one might still think that the figure realises a square shape.

If our figure could realise a square shape, though it is not a complete square, then we might have provided an analogy to the case in which an empty category in a grammatical structure is physically realised though no physical element corresponds to that empty category. Just as there are missing physical parts to the square shape instantiated by the incomplete figure, so there might be absent physical elements for the empty categories though physical parts realise a grammatical structure with null expressions. Perhaps such completions could provide Devitt with a model for how empty categories could be realised by relations amongst physical sounds and marks.⁴²⁴

⁴²³ From Michael Bach's collection of optical illusions at:

<http://www.michaelbach.de/ot/index.html>

⁴²⁴ Though it is not clear to me that any decisive objections from empty categories have been raised against Devitt, given that such elements are rife in grammatical structure, anyone who wanted to defend a conception like Devitt's ought to do more than show that their conception is not logically

There are further grammatical properties that are difficult to accommodate within Devitt's linguistic conception on which grammatical properties are realised by *relations* amongst entities in the physical environment. For not all grammatical properties are configurational in this way. There are grammatical properties that are not *realised by* the relative structural positions of constituents, though they *do* play a role in *determining* constituent configuration. As these properties are not configurational properties of linguistic structures at all, they cannot be realised in relations amongst physical entities.

Lexical features are such a class of grammatical properties. Two important sorts of lexical features are those belonging to number, gender and case (called ψ -features) and those pertaining to the arguments or thematic roles assigned, particularly by verbs (called θ -features).⁴²⁵ I'll focus on number features for the purposes of illustration.

There are number features on lexical items, such as their being singular or plural, which are not configurational properties. Whether a lexical item bears the grammatical property of being singular or plural is involved in *determining* the admissible structures into which it can enter. But whether the lexical item has the grammatical property of being singular or plural is not a matter of its structural position relative to other constituents. Rather it is a feature specification on the lexical item. Such lexical features determine the combinatorial possibilities but are antecedent to grammatical combination and so are not realised by structural relations. They are *lexical* features, features that the items which combine to form structures have. If minimalist theories (see Appendix) are on the right track then grammatical derivation is driven by just such features.

On Devitt's account, grammatical properties are realised in relations amongst symbols and, hence, are extrinsic. As Collins has pointed out, Devitt's account struggles to make sense of the lexicon or the fact that configuration is

ruled out by such elements. They ought to make a positive case for how *well* their conception accommodates and explains the existence of these elements. See Collins (2008b) and Devitt's (2008a) response for discussion of Devitt's attempt to elaborate his picture by appealing to grammatical conventions.

⁴²⁵ See Adger (2003) pp.22-52 for ψ -features, pp.77 – 90 for θ -features.

determined by the properties of Merged lexical items (see Appendix).⁴²⁶ So Devitt needs to explain how lexical features are *realised* in the physical sounds and marks.

Devitt might respond by claiming that there are also difficulties with locating grammatical features psychologically. Devitt would be right to note that there are little understood issues about how grammatical properties are mentally represented, ultimately in virtue of properties of the brain. But these are problems shared by the whole of cognitive science. As Devitt is an advocate of cognitive science, he will share these problems because the linguistic conception is committed, as all accounts are, to there being an explanation of the psychological design that allows us to recognise and structure grammatical expressions. The best explanations of how we do this involve our having a mental realisation of grammatical features. Devitt's distinctions commit him to the existence of such mental machinery, for he is committed to the existence of linguistic processing that respects linguistic structure. Moreover, given the argument of Chapter Three, all parties are committed to a distinction between grammatical competence and performance in investigating generative grammars, where the grammatical competence encodes the speaker's sensitivity to grammatical features and categories.

Proponents of a non-psychological conception of generative grammar might point out that this difficulty with lexical features only arises for the configurational view that Devitt suggests. It only follows that lexical features aren't physically realised by being configurationally realised. For the difficulty is generated by the claim that grammatical properties are realised in *relations* amongst symbols, and not by the claim that they are realised non-psychologically as such. Of course, we still need an explanation of how the properties *are* realised in the external environment. But one cannot claim to have established the inadequacy of such non-psychological conceptions without a more principled and thorough-going argument.

The line of argument that I'll develop against non-psychological conceptions does not challenge the possibility of such non-psychological realisations of grammatical properties but rather challenges their explanatory credentials. My claim is not that it is impossible to account for the realisation of grammatical properties in the physical sounds and marks but rather that there is no

⁴²⁶ Collins (2007) p.3

explanatory pay-off for grammarians in so conceiving them. Generative grammars have to meet certain conditions of adequacy. To construct a generative grammar that can meet these conditions, grammarians require a distinction between grammatical competence and linguistic performance, where the theory of grammatical competence reveals the grammar of the speaker's language. I'll argue that this required psychological theory of grammatical competence is sufficient to meeting descriptive and explanatory adequacy and that no explanatory benefits accrue to grammarians from the postulation of non-psychological grammatical properties beyond the properties of the competence system. As Collins remarks of his conclusion in favour of the psychological conception, it is "not premised upon a rejection of the idea that linguistic properties are realised by external concreta."⁴²⁷ It rather turns out that the properties that successful grammatical explanations trade in need not be conceived as realised externally to the mind.

5.3 Meeting Descriptive Adequacy

The condition of descriptive adequacy requires that "A fully adequate grammar must assign to each of an infinite range of sentences a structural description indicating how this sentence is understood."⁴²⁸ In Chapter Three I argued that in order to determine the infinite range of structural descriptions assigned to the sentences of a speaker's language we need to develop a theory of their grammatical competence. Proponents of the psychological conception argue that such a theory of grammatical competence, revealing the infinite range of structural descriptions assigned to the sentences of the speaker's language by his internalised grammar, serves to meet this requirement of descriptive adequacy. But proponents of non-psychological conceptions of generative grammar deny that a DAG should be conceived as a theory of grammatical competence. They argue that a generative grammar describes an infinite range of non-psychological grammatical properties.

⁴²⁷ Collins (2007) p.2

⁴²⁸ Chomsky (1965) p.5.

Given my argument in Chapter Three, generative grammarians need to appeal to a theory of a speaker's grammatical competence in order to determine a DAG for a speaker's language. But do generative grammarians also need to appeal to a domain of non-psychological grammatical properties to meet descriptive adequacy, as proponents of non-psychological conceptions claim? In this section I want to argue that they do not. My broader aim in this section and the next is to develop a parsimony argument against non-psychological conceptions of generative grammar.

A parsimony argument in favour of the psychological conception can be developed as follows. We *require* a theory of a speaker's grammatical competence to work out what the grammar of his language is and thereby construct a theory that meets descriptive adequacy. Moreover, whatever non-psychological realisation of grammatical structure we posit, the competence systems of those that speak and understand the language must be organised so as to licence and be sensitive to just those structures. The speakers require an internal realisation of the grammar in virtue of which they are able to recognise such structures as part of their language. But, at least as far as meeting descriptive adequacy is concerned, once we have such a theory of grammatical competence, to then posit a non-psychological realisation of the properties would accrue no further explanatory benefit beyond that which accrues to the theory of grammatical competence. For it is that theory of competence we use to construct a DAG. So, a generative grammar, as far as descriptive adequacy goes, is best construed as the psychological theory of grammatical competence, and the putative non-psychological grammatical properties are theoretically dispensable.

It is a general maxim is to keep theories as simple as possible, without losing any explanation, so as to see what is really important to theories and what is not. Accordingly, a scientific theory ought to posit as few objects and properties as it requires. This is the familiar principle of Ockham's Razor. This principle is important within linguistic theory itself (see Appendix) but its application is just as apt at a meta-theoretical level.⁴²⁹ Linguistic theories should be interpreted so as to

⁴²⁹ However, the import of the principle as applied theory internally is different to its import at a meta-theoretical level. Parsimony within linguistic theory means appealing to the fewest

explain a complex of phenomena, but our conception of the sorts of properties they appeal to should be kept as lean as possible consistent with their explaining these phenomena.

I should emphasise that I'm not concerned to deny the very *existence* of non-psychological grammatical properties. It may be that there *are* grammatical properties 'out there'. But I will argue that they are of no theoretical significance, at least as far as generative grammar goes. They are theoretically dispensable, and so are not what generative grammars are about.

I assume the following general principle for discerning what theories are about, which I've called the *parsimony principle*.

Parsimony Principle: If one set of properties P is required by a theoretical explanation, and P can explain the targeted phenomena as well as the conjunction of P with another set of properties P', then an interpretation of the theory in terms of P is preferable to an interpretation of the theory in terms of P+P'.

The interpretation of the theory as appealing to P+P' would be *unparsimonious* according to my principle. The theoretical interpretation that appealed to only P rather than P+P' would have the following virtues. The notion of simplicity in theoretical machinery is widely employed within sciences as a way of dealing with theoretical underdetermination. But just as we could have a number of grammatical theories, all of which were consistent with the received data, so we could have a number of interpretations of grammatical theory that might be made consistent with the details of the theory. Within grammatical theory itself, a maxim is explicitly adopted not to multiply the theoretical apparatus beyond necessity, so as to choose between non-equivalent theories that cover the data. This maxim guides the theories away from mere descriptive coverage of the data towards less numerous but more

unmotivated assumptions rather than the fewest properties per se. This embodies a commitment to principled explanation. But as Ockham's Razor is generally understood it guides us not to multiply entities and their properties beyond theoretical necessity. When considering the meta-theory of linguistics, it is any theoretically dispensable entities or properties that would be unmotivated assumptions.

explanatory principles. The same principle applied to the interpretation of the theories guides us to the properties that are really doing the explanatory work, and ascertains which properties are mere explanatory danglers.

Descriptive adequacy can be achieved very naturally by a theory of the speaker's grammatical competence because it is this very theory of grammatical competence that the grammarian uses to determine what the structures of a speaker's language are. Suppose we start with some evidence from speakers' linguistic intuitions.⁴³⁰ Speakers have the intuition that the reflexive in (65) means *Simon's father* but not on *Simon* or *father* alone.

(65) *Simon's father* shaved himself.

According to the psychological conception, the explanation makes no appeal to non-psychological properties and runs as follows. The structural interpretation of (65) that the speaker's language assigns is revealed by the organisation of that speaker's grammatical competence. The speaker's grammatical competence is organised so that the reflexive is dependent for its interpretation on *Simon's father*. The principles of the speaker's grammatical competence require that the reflexive must be bound by a local antecedent that c-commands it. To determine which is the c-commanding constituent, the competence system arranges lexical items into a hierarchical structure. On the basis of this assignment of hierarchical structure, the competence system determines that the lexical items *Simon* and *father* do not c-command the reflexive, but *Simon's father* does. So the competence system assigns a structural description according to which the reflexive is referentially dependent on *Simon's father*. As the competence system has infinite generative capacity, such structures can be assigned for the infinite range of sentences in a speaker's language.

On the non-psychological conception, the attempt to meet descriptive adequacy would go as follows. We first discover the relevant language and its grammatical structures by investigating the speaker's grammatical competence. Once we know which grammatical structures he recognises and the structural interpretations his competence assigns, and what aspects of his linguistic intuitions

⁴³⁰ See Collins (2008) for this style of argument.

and behaviour are performance effects, we can extrapolate to the grammatical structures of his language. These grammatical properties are then claimed to be realised non-psychologically. We can then say that it is a non-psychological grammatical property of the language that reflexives must be bound by the whole DP and are not bound by either of the constituents of the DP because the reflexive must be c-commanded by its antecedent. As (65) is such a sentence, realised in a non-psychological domain, containing a reflexive, *himself* must be bound by the whole DP *Simon's father* and not by either of the constituents of the DP, *Simon* or *father*, which fail to c-command the reflexive. Therefore, the non-psychological grammatical theory will assign (65) such a structure. And speakers that have an internal realisation of this grammar will be sensitive to the non-psychological properties so as to make the intuitive judgement that (65) is an acceptable sentence of the language, with *himself* meaning *Simon's father*.

The theory of grammatical competence assigns the sentence the correct structural description indicating how it is understood in a speaker's language. So does our grammatical theory given a non-psychological interpretation. But notice that our non-psychological explanation appeals to a theory of grammatical competence to determine what the grammar of a speaker's language is. It uses a theoretical conception of the speaker's competence, as this grammatical competence is distinguished from other aspects of linguistic performance, to determine what the structures of a speaker's language are.

Moreover, the non-psychological conception then has to appeal to the speaker's internal realisation of these grammatical properties, in a system of grammatical competence, to explain how the speaker can intuit these properties. For we know that the speaker must be sensitive to the described structure so that they could come to the relevant interpretation. Even if the grammatical properties are 'out there', the speaker must be configured so as to immediately cognise them and assign them the structural interpretation on which the interpreted elements c-command one another. The claim that these grammatical properties are realised non-psychologically adds nothing as far as meeting descriptive adequacy goes. For the purposes of explaining the evidence from a speaker's intuitions and assigning correct structural descriptions to sentences of a speaker's language indicating how they are understood, the postulation of non-psychological grammatical properties is

superfluous. And hence, as Chomsky has argued, a grammar “is descriptively adequate to the extent that it correctly describes the intrinsic competence of the idealised speaker.”⁴³¹ Though we can and do talk of items in the environment having grammatical properties, this has no clear consequence for generative grammar as an explanatory theory; the really interesting question, as Collins suggests, “is whether linguistic properties so construed enter into theoretical explanation.”⁴³²

5.4 Meeting Explanatory Adequacy: Nativism

An EAG serves to determine which grammar a speaker has actually acquired, from amongst the possible DAGs for a speaker’s language. An EAG offers an explanation of speakers’ linguistic intuitions on the basis of an empirical hypothesis about grammar acquisition.⁴³³ A theory that meets this condition is justified to a greater extent as a theory of a speaker’s language than a merely DAG.

Hypotheses about grammar acquisition and the innate grammatical predisposition of the child are hypotheses about human psychology. So meeting explanatory adequacy requires a *psychological* theory about grammar acquisition. But proponents of non-psychological conceptions might claim that this theory of grammar acquisition, in turn, requires the grammarian to appeal to non-psychological grammatical properties. This is part of Devitt’s attack on the psychological conception. He says:

Some of a person’s language may well be innate but she learns a good deal of it. On my view, this learning is a matter of acquiring conventions... Once again [non-psychological] linguistic properties have a causal role... Without those entities, language learning becomes a mystery.⁴³⁴

⁴³¹ Chomsky and Katz (1974) p.348, see also Chomsky (1965) ch.1.

⁴³² Collins (2007) p.419

⁴³³ See Chomsky (1965) pp.25-6

⁴³⁴ Devitt (2006) p.189

For, it is a natural thought that grammar acquisition is a kind of learning. And if children are to learn their grammar from experience of their linguistic environment, then one might think that there had better be non-psychological grammatical properties in that environment from which children can learn. In this section I want to argue against the view that grammar acquisition involves the child learning their grammar from experience of grammatical properties in their environment, and in favour of a *nativist* view of grammar acquisition. Nativist theories appeal to an innately determined grammatical framework in explaining grammar acquisition and make no essential appeal to a domain of non-psychological grammatical properties.⁴³⁵

5.4.1 Theories of Grammar Acquisition: Nativist Theories and Learning Theories

In Chapter Two, I defended explanatory adequacy as a condition of adequacy on generative grammars. A consequence of this condition is that general grammatical theory should be regarded as an abstract theory of grammar acquisition. It describes, at an abstract level, the principles that enable every child to attain their mature grammar on the basis of primary linguistic data (PLD). As such, generative grammars are an explanatory hypothesis about the form of naturally occurring languages and the principles that make their acquisition possible. To meet such

⁴³⁵ Again, I should stress that whilst nativist arguments put pressure on non-psychological interpretations of the science of generative grammar, which aims to meet explanatory adequacy, they are not intended as a demonstration that non-psychological grammatical properties do not *exist*. There are, for example, nativist arguments regarding geometrical knowledge. But they do not show that there are no non-psychological geometrical properties. Perhaps there are good philosophical arguments that might lead us to believe in the existence of non-psychological grammatical properties. But my point is that, irrespective, non-psychological grammatical properties are not required to meet descriptive or explanatory adequacy, whereas psychological properties of the grammatical system are. So, if non-psychological grammatical properties do exist, they are still not an explanatory commitment of generative grammars.

explanatory objectives, the grammarian clearly needs a theory of the psychological properties involved in grammar acquisition. But, in explaining acquisition, does the grammarian also appeal to a domain of non-psychological grammatical properties, distinct from the psychological properties? First, we need to clarify what is required of a theory of grammar acquisition by setting out some broad facts about language acquisition.

At around six months old children start babbling and recognising the prosodic properties of words and phrases. At around ten months they begin pairing words and meanings. Between ten and twenty months they are at the one and two word stage, showing some understanding of these words and acquiring a large vocabulary before they are able to structure sentences. Between 20 and 24 months children enter a period of language acquisition, sometimes called the “syntax spurt”. By only three or four years, they have developed the full recursive aspects of grammar and a relatively stable grammatical competence, roughly equivalent to that of the mature speakers in their speech community. The fact that normal children go through this rapid period of grammatical development is well attested in their ability to understand and produce novel sentences, their recognition of acceptable and unacceptable expressions of their language, their discernment of ambiguities, and their sensitivity to relations of paraphrase and entailment.⁴³⁶ These facts suggest that grammar acquisition begins within a fairly rigid timeframe and that it occurs within a limited critical period.⁴³⁷ So we know that as a human child develops, their brain responds selectively to stimuli from their environment in ways that other animals’ brains do not. The aspect of the human brain responsible for grammatical structure starts out in some initial state and develops to some steady state of grammatical competence.

Though exposure to linguistic stimuli in the critical stages of grammar acquisition is vital, interestingly it seems that the external input need not come

⁴³⁶ See Crain and Thornton’s (1998) 10-year study of the competence of under-fives.

⁴³⁷ That grammar acquisition occurs within a critical period is supported by the case of Genie, and other similar cases, in which a victim of severe neglect and abuse develops only limited linguistic abilities and never acquires anything like mature grammatical competence. Genie was discovered aged thirteen but after eight years of study never developed mature knowledge of grammar. See Curtiss (1988) for the case of Genie and other important cases which suggest similar conclusions.

through any particular sensory modality. Blind children and deaf children acquire languages of the same structural complexity and with much the same ease as other children. Deaf children exposed to signed languages from birth acquire these sign languages in the same stages and on the same time schedules that hearing children acquire their spoken languages. The child sign language exhibits much the same semantic, discourse and pragmatic complexity as hearing children's language, and:

[T]heir expressions' conceptual content, categories, and referential scope demonstrate unequivocally that their language acquisition follows the identical path seen in age-matched hearing children acquiring spoken language.⁴³⁸

Pettito found that bilingual children who can both sign and hear, and are exposed to both a signed and spoken language from birth, demonstrate no preference whatsoever for speech. She also found that hearing children exposed only to sign reached all the same milestones at the same times as their peers, even "babbling on their hands".⁴³⁹ In fact, there is evidence that grammar can be acquired by subjects "with no sensory input beyond what can be gained by placing one's hand on another person's face and throat":

The analytic mechanisms of the language faculty seem to be triggered in much the same ways whether the input is auditory, visual, even tactual, and seem to be localised in the same brain areas, somewhat surprisingly.⁴⁴⁰

We can call the aspect of a child's brain that is responsible for their acquiring a grammar, a *Grammar Acquisition Device* (GAD). The idea that there is a GAD is itself sometimes regarded as controversial. But this is a mistake: "GAD" just labels whatever aspects of a child's brain supports their acquisition of a grammar, making no further assumptions about its specificity or content. The GAD takes the child from an initial state, which is just an expression of the human biological endowment, to a steady state of grammatical competence, given some input (PLD). It is then an

⁴³⁸ Pettito (2005) p. 89. See Pettito (2000) for more extensive discussion.

⁴³⁹ Pettito (2005) p.90

⁴⁴⁰ Chomsky (2000) p.122

empirical matter to determine the nature of the initial state, the GAD, the steady state of grammatical competence, the stages in between and the role of PLD.

Nativist Theories

A common theme in arguments for nativist theories of grammar acquisition is that children ignore a wide range of possible grammatical principles that are suggested by their PLD, whilst projecting beyond the PLD in ways that it does not suggest. To explain this, nativists make detailed proposals about a rich UG that excludes an indefinite number of conceivable grammatical principles and thereby excludes an indefinite number of possible assignments of structures to expressions. UG is an *innately* endowed set of linguistic principles. The broad framework within which recent nativist theories of UG have been offered is the *Principle and Parameters framework* (P&P), in which the initial state of the language faculty consists of a fixed set of universal grammatical principles with open parameters that require setting. Grammar acquisition involves the setting of these open parameter values to one of a limited range of options specified by the principles, but never deviating from the boundary conditions that the UG principles impose.

The interaction of only a few parameters can lead to a great deal of divergence in the sentence structures which particular grammars allow. If there are only a few parameters, then there are only a few possible analyses of the linguistic stimuli that children are presented with. For this reason, nativists claim that such a theory stands a good chance of explaining how children acquire grammars so quickly; on the basis of data which, they claim, does not contain the right information to learn a grammar of such a very special kind.

The most fundamental commitment of the *Poverty of Stimulus* (POS) arguments for nativism, which I'll defend in §5.4.2, is to the invariance of grammar acquisition with respect to the richness or poverty of PLD over wide bounds. The conclusion of the POS arguments is that speakers do not acquire their grammar by learning from properties that they find in their individually variable PLDs. Instead, a rich and innate framework of universal, grammatical principles determines the

humanly acquirable grammars and acquisition involves selecting amongst the options these innate principles specify.⁴⁴¹

POS arguments for nativism appeal to the impoverished nature of the data available to children in first language acquisition: ‘impoverished’, if we consider the data as a basis for *learning* the principles that children acquire. On the strength of these POS arguments, Chomsky claims that “in certain fundamental respects we do not really learn language; rather, grammar grows in the mind.”⁴⁴² He compares the development of a grammar to the development of a visual system and other physical organs, saying:

In both cases, it seems, the final structure attained and its integration into a complex system of organs is largely predetermined by our genetic program, which provides a highly restrictive schematism that is fleshed out and articulated through interaction with the environment.⁴⁴³

The comparison with the growth of organs might lead one to wonder in precisely what sense nativists think that the UG principles are ‘innate’. The most important point about grammar acquisition from a nativist perspective is the negative one that the parameterised UG principles are *not learnt*.⁴⁴⁴ The acquisition of a grammar is argued to be innately constrained insofar as the character of acquired grammars is

⁴⁴¹ Traditional UG was a theory of *substantive* universals, claiming that certain grammatical categories (like noun, verb, and so on) are found in the grammatical structures of the sentences of any natural language, and that these categories provide the underlying grammatical structure of each language. But Chomsky is keen to point out that it is also possible to search for universals of a far more abstract sort. He calls these *formal* universals and it is this sort of universal that has come to play a more prominent role in generative grammar than the substantive sort. Claims about formal universals are claims that the grammar of every language meets certain specified formal conditions, which tend to be far removed from the commonsensical grammatical categories. The truth of a hypothesis about formal grammatical universals “would not in itself imply that any particular rule must appear in all or even two grammars.” (Chomsky 1965 p.29) The formal universals involve the *character* of the rules that appear in grammars and the ways in which they can be interconnected.

⁴⁴² Chomsky (1980) p.134

⁴⁴³ *ibid.*

⁴⁴⁴ This is the point that Chomsky (1986) argues for: that the core grammatical principles are not learnt by induction, theoretical inference or statistical inference.

determined by UG principles and the limitations they impose on parameter settings, rather than by learning from the PLD. POS arguments suggest to the nativist that we are innately constrained to develop grammatical competence systems with a certain form, integrated into wider biological systems, in a similar way that we are innately constrained to grow limbs and bodily organs. This is why some nativists call the language faculty, the *language organ*. Chomsky says “Language acquisition seems much like the growth of organs generally; it is something that happens to a child, not that the child does.”⁴⁴⁵

There are, of course, hard questions about innateness. There are questions about biological endowment and development generally which obviously fall outside of the grammarian’s scope. But there are also special questions pertaining to the human brain and psychological function, and little understood questions about the neurobiological realisation of computational systems. If asked how the cognitive structures of the grammatical competence system could be innately endowed, the nativist grammarian should note that these questions also fall outside the scope of his inquiry. He might defer to the core biological sciences for some understanding of biological endowment, and to the neurobiological sciences where investigations of cognition and the brain are still in their early stages.

It may be that there is more that the nativist grammarian can say about innateness before he defers to other sciences. The notions of *canalisation* and *psychological primitiveness* have played a part in some nativists’ conceptions of their theories. A psychological system is canalised if its development is insensitive to the environment across a very broad range, and relatively insensitive to genetic variation, hence unfailingly acquired under normal conditions. A psychological system is primitive if it is not acquired by any psychological process of inference or perception.⁴⁴⁶

We might expect some convergence between the canalised psychological systems and the primitive psychological systems because primitives tend to grow and growth tends to canalise. And non-primitive systems may be acquired by

⁴⁴⁵ Chomsky (2000) pp.6-7

⁴⁴⁶ POS arguments are directed at the claim that grammatical principles are learnt by inference, it is assumed that they are not learnt by simple acts of perception of an instance.

psychological processes that are sensitive to environmental contingency, and hence, not be canalised. But there may be instances where canalisation and primitiveness come apart. We might have principles inferred directly from innate principles that are highly canalised but not primitive. Equally, where there are unstable biological factors in play there might be primitive systems that are not canalised.⁴⁴⁷

In explaining how grammatical parameters are set, some nativist theories appeal to a distinction between learning from environmental stimuli and *triggering* by external stimuli.⁴⁴⁸ These are different ways in which PLD might be necessary for grammar acquisition. The stimuli to which a child is exposed might serve as a basis from which he could generalise to the grammatical principles of his language. But alternatively the linguistic stimuli to which the child is exposed might play merely a triggering role in driving the GAD along predetermined paths. On the triggering model that these nativists appeal to, PLD plays a role in initiating and facilitating development of the innate principles and their parameter settings. But this triggering is a causal process by which one amongst a limited range of highly organised developmental trajectories is determined. There is no assumption that the stimulation provided by the PLD shapes the way in which the GAD functions or plays any role in shaping the grammatical principles that characterise the mature competence.

The nature of the PLD clearly does matter to which grammar a child ends up with. Children brought up in Italian speaking environments, acquire Italian grammar. Children brought up in English speaking environments, acquire English grammar. Italians can, for instance, use null subject sentences: they acquire a *pro-drop* language. But English is not a pro-drop language. But the availability of such limited options is not an argument against nativism. As Collins puts it:

Obviously, the human mind/brain is not designed to acquire English at the expense of Japanese, etc. But it does not follow that syntax is not innate; it simply follows that a given range of parametric values is not fixed, but the range of possible values *is* fixed. For

⁴⁴⁷ See Samuels (2002) for more detailed discussion of these matters, and in particular primitiveness. See Segal (forthcoming) for emphasis on the negative aspect of the nativist's claim - that grammar is *not learnt* – and relevant discussion.

⁴⁴⁸ See, for example, Gibson and Wexler (1994) or Nowak et al (2001).

example, if there is a head parameter, then children in London get it set to head-initial, while children in Tokyo get it set to head-final. The options are innate, the decision is made by experience.⁴⁴⁹

English children settle on a grammar that puts direct objects after verbs in transitive constructions whilst Japanese and Korean children settle on a grammar that puts them before. So this aspect of their mature competence seems to be sensitive to the nature of the PLD, although there appear to be a limited range of options.⁴⁵⁰ A common nativist theme is close attention to cases where the child, on exposure to a little English, say, tries out options that characterise Korean or another UG language, even absent evidence of such options in the PLD. What are often described as childish errors by parents are conceived by nativists as explorations of the limited grammatical options actually realised by possible human grammars. This is called the *Continuity Hypothesis*.⁴⁵¹ For example, English children sometimes try inserting an extra *wh*-word in long distance questions as in (66).⁴⁵²

(66) What do you think what pigs eat?

These structures are well-attested in adult German, Irish and Chamorro though it is not part of English. The nativist claims that, in such cases, English children adhere to a constraint from another language, compatible with UG, though they do not try out UG-incompatible constraints. Other things being equal, learning theories should predict that “children (insofar as they diverge from adults) will initially employ

⁴⁴⁹ Collins (2008a) p.30

⁴⁵⁰ Just as Japanese objects precede verbs, objects of prepositions precede prepositions. Generative grammarians have come up with a parameter to characterise the VO and OV languages. They claim that it is a principle of the X-bar theory of phrase structure that every phrase must have a head, but it is a matter of parametric variation whether the complement of that head follows or precedes the head. If the language picks the ‘precede’ option, it will be like Japanese. If the language picks the ‘follow’ option, it will be like English. The parameter applies to all phrases, so what is true of VPs will be true of NPs, PPs and APs. Baker (2001) argues that this sort of explanation can be multiplied to accommodate all sorts of phenomena across languages.

⁴⁵¹ Pinker (1984), Crain (1991), Crain and Pietroski (2001)

⁴⁵² See Crain and Pietroski (2002) pp.12-14.

constructions that are less articulated than those employed by adults. Complexity in the child's hypotheses about the local language should be driven by what the child hears; otherwise, complex hypotheses will look like reflections of a mental system that imposes certain structures more or less independently of experience."⁴⁵³

Nativists, who think that the dependence of grammar acquisition on PLD is a matter of triggering, think that PLD matters because it exerts a brute causal pressure on the different parameter settings. But there are competing nativist theories according to which statistical learning rather than triggering is involved in setting the parameters amongst the predetermined options. Yang has offered a nativist account which relies on the child's uptake of statistical information in parameter setting.⁴⁵⁴ According to Yang, the grammatical principles with their open parameters are innate and highly domain-specific, but the process by which the parameters are set is domain-general statistical learning. Yang holds that a grammatical knowledge system develops "in the head", insensitive to environmental factors, but an experiential process of examining the environment (using the innate categories and relations) is involved in shifting the child around the parameters of the system.⁴⁵⁵

It is not my aim to adjudicate between triggering models and statistical inference models. But one point about Yang's account requires emphasis. Yang argues that "a full explanation of children's grammar development must abandon the domain-specific acquisition model of triggering, in favour of probabilistic learning mechanisms that might be domain-general *but nevertheless operate in the domain-specific space of syntactic parameters.*"⁴⁵⁶ Yang is a proponent of P&P and

⁴⁵³ Crain and Pietrokxi (2002) p.11

⁴⁵⁴ See Yang (2004).

⁴⁵⁵ Yang's view is that the presence and effects of universal and innate grammatical principles has been successfully tested in young children. He says (2004 p.453): "The hypothesis space for grammatical structure within which that learning takes place is "the grammars and parameters defined by innate UG".

⁴⁵⁶ Yang (2004) p.451 my italics. Yang here draws an important distinction between knowledge and mechanism. As Fodor (2001 p.85) has pointed out nativists can maintain that we have innate, domain-specific knowledge of grammatical structure "while remaining entirely agnostic about the domain specificity of language acquisition *mechanisms*". Pinker (1999) defends an associative model for the acquisition of irregular verbs within a strongly nativist framework.

makes it clear that the statistical inference involved in parameter setting “does not refute UG”.⁴⁵⁷ Yang defends his view with evidence from children’s setting of the Pro-Drop parameter, all the while emphasising that grammar acquisition involves the development of innate grammatical principles. These principles allow learners to pick up on the specific, relevant aspects of the input, such as expletive subject sentences in the case of Pro-Drop.

Yang claims that children’s success in statistically inferring particular hierarchical structures “strengthens rather than weakens” the nativist case.⁴⁵⁸ He thinks that good arguments for nativism can be built on the fact that children must employ specific types of representation, such as grammatical constituency and hierarchical phrase structure, rather than generalise over “linear strings of words, or numerous other logical possibilities.” Yang argues that where children record statistical information pertaining to hierarchical structure, any conclusion they draw from the findings “must presuppose that children already know *what kind* of statistical information to keep track of.”⁴⁵⁹

Learning Theories

The main alternative to nativism is to develop a theory of how the child learns the grammatical principles from the properties of their PLD. Such theories are usually called *empiricist* theories, or *learning* theories. The central commitment of such theories is that the child learns its grammar by some form of inductive, statistical or abductive inference from PLD. On this approach, grammarians examine what is available in the PLD, using corpus data to ascertain its properties, and then attribute to the child learning mechanisms by which the grammatical principles could be extracted.

⁴⁵⁷ Yang (2004) p.451

⁴⁵⁸ Yang (2004) p.452

⁴⁵⁹ *ibid.* Yang claims that an “infinite range of statistical correlations” exists in the PLD, and hence, the fact that children can use statistical learning mechanisms to acquire a grammar in the relevant timeframe requires that “at minimum, they know the unit of information over which correlative statistics are gathered... the learner is sensitive to specific types of input evidence relevant for the setting of specific parameters.” (pp.452-4)

The principles of inference that such theories appeal to are normally called *general purpose learning mechanisms* (GPLM). The mechanisms are *general purpose* because too much grammar-specific information or bias built into the learning mechanisms would concede to the nativist that the child comes equipped with innate grammatical principles that determine his acquisition of grammar.

A conclusion of the POS argument I defend in §5.4.2 is that we lack a clear idea of how a child might employ GPLM to learn the grammatical principles of their mature competence, and that nativist theories constitute a better explanation of grammar acquisition. Without the promise of learning explanations, describing how children might learn the grammatical principles from PLD, the challenge from learning theories amounts to only the logical point that alternatives to nativism are possible.

Lappin attempts to refute claims that in light of POS arguments there is an absence of promising learning theories.⁴⁶⁰ Lappin highlights proposals concerning machine learning techniques and how theories of machine learning point the way to alternatives to nativist theories and UG. Lappin concludes that: “Recent research on unsupervised machine learning of grammar offers support for the view that knowledge of language can be achieved through general machine-learning methods with a minimal set of initial settings for possible linguistic categories and rule hypotheses.” Lappin claims that the learning priors of grammatical categories and rules required by such machines are more “minimal” than those proposed in nativist theories of UG (though they *do* assume binary-branching structure), and that machine learners have had encouraging successes learning to parse certain grammatical constructions.

Though the issues Lappin raises concerning what can be learnt from PLD are interesting, he is yet to tackle the issues about how his findings fit with children’s actual patterns of grammar acquisition. Given that these issues (which I discuss in §5.4.2) form the core of the nativist’s case, Lappin has not made a good case for the relevance of his findings to nativist arguments about the child’s acquisition of grammar. As he admits, the results he cites “do not, of course, show us anything about the processes that human learners actually apply in acquiring

⁴⁶⁰ See, for example, Lappin and Schieber (2007).

natural language.” A proper assessment of the results he cites would be beyond the scope of this thesis, but it should be noted that the best mechanisms Lappin and Schieber discuss currently have a much higher ratio of mistakes in parsing English sentences than it would be plausible to attribute to a child.⁴⁶¹

Nativists claim that, given compelling grammatical analyses, they have the *best* explanations of how suitable grammars were acquired. From their perspective, the major debates are not between nativist theories and learning theories but between the dozens of proposals within the P&P framework which offer detailed and empirically testable proposals about the nature of UG and its role in grammar acquisition.

One might wonder whether the options of nativism or learning theory are exhaustive, or whether there are further options for explaining how we acquire grammar. One possibility, sometimes raised in philosophical discussions influenced by Wittgenstein, is that grammar is acquired by training or inculcation into a set of linguistic practices (see §1.3.1). The problem with this suggestion is that there is a lack of evidence for treating grammar acquisition as the honing of an ability or practical skill. There is little evidence of training and children who receive no training acquire grammars. The evidence I’ll present supports Smith’s assessment that:

The forlorn idea that we do all this by analogy with the repetitious learning of a manual skill is a non-starter and does not even merit serious discussion. There is no evidence that such a practice takes place or that mistakes of the kind expected in such training actually occur.⁴⁶²

This is not to deny that training or practices have any role in understanding language more broadly construed, but the evidence I’ll document do not bear out their involvement in grammar acquisition.

⁴⁶¹ Lappin and Schieber (2007) pp.11-12

⁴⁶² Smith (2006) p.957

5.4.2 Poverty of Stimulus Arguments for Grammatical Nativism

The argument from POS to grammatical nativism runs as follows.

POS1. All normal children unerringly acquire G, where G is some grammatical principle or set of grammatical principles.

POS2. They acquire G either via GPLM or using innately specified, grammatical information.

POS3. The available stimulus – the PLD – is too poor for them to unerringly acquire G by GPLM.

POS4. So they do not acquire G by GPLM.

So, by POS2 and POS4,

POSC. They acquire G using innately specified, grammatical information.

If the premises of the argument can be defended, the grammarian aiming at explanatory adequacy should focus on detailing and empirically testing theories of innately-specified, grammatical information.

Such POS arguments are not used to defend the claim that it is *impossible* that the child could learn G given any possible GPLM and PLD. If the linguist can discover the UG principles, and the parameter settings of the individual languages, by collecting data and deploying scientific method then, at the very least, it is conceivable that children do so. We can, however, use grammarians as a measure of the difficulty of the learning task of recovering some candidate G from PLD.⁴⁶³ The grammarian is something like a GPLM but one with significant advantages over a child. Grammarians already possess a language and can be expected to possess a high level of general intelligence. They also work in large group enterprises using focused and systematic data. But the grammarian's task is incredibly hard and it is an ongoing project to work out those principles that the child acquires effortlessly.

⁴⁶³ Segal (forthcoming) p.6 makes this suggestion. The obvious reply, which Segal notes, would be for learning theorists to make a case that children employ a sophisticated and subtle learning mechanism in language acquisition that is not available to linguists. The evidence I present in defence of the POS suggests that any such case for a learning theory faces severe obstacles.

Yet PLD does not provide such targeted evidence as drives grammatical theory and the child has only a few years to carry out the learning task which has engaged grammarians for decades. As Chomsky, considering the empty categories, describes the situation:

It is no simple matter for the scientist inquiring into language to discover that these elements exist and to determine their properties, and this task requires a broad range of evidence not available to the child, including evidence from a variety of languages and evidence acquired by sustained empirical inquiry informed by complex theory construction.⁴⁶⁴

Of course, it might be that young children are very gifted linguists but not gifted physicists or musicians. All normal children *might* successfully hypothesise principles equivalent to UG and the parameter settings of their language. But then we would want to know why these gifts extend only to the domain of language, and why they are not available for second language acquisition.

In developing POS arguments for specific clusters of grammatical principles, the details of grammatical theory matter. A prerequisite for working out whether grammar acquisition is an instance of learning from PLD is a theory about the principles that are acquired. But the general form of the problem of language acquisition is simple. Given a theory of what is acquired, and considering the child as an empty box, we can ask what would have to be inside the box such that a child could output their grammar on the basis of PLD as input. Insofar as a theory of grammar acquisition requires us to attribute grammatical principles to the box beyond GPLM, there will be POS arguments to motivate nativist theories.

The candidate Gs for POS arguments are abundant. But I'm going to focus on two illustrative examples of binding principles and Yes/No question formation.

Principle A of binding theory states that a reflexive must be bound by a local antecedent, ruling in cases like (67) and ruling out cases like (68).⁴⁶⁵

⁴⁶⁴ Chomsky (1988) p.91

⁴⁶⁵ Segal (forthcoming) notes that Principle A may not be a good candidate for a UG principle as it appears to be violated by Mandarin, and perhaps Chinese dialects generally. Segal also notes that there are apparent counterexamples in English like "Phoebe saw Joey's picture of herself" (which

(67) John_i likes himself_i.

(68) *John_i thinks that [Mary likes himself_i].

Principle B states that a pronoun is not bound by a local antecedent. It rules in cases like (69) and rules out cases like (70).

(69) John_i said that [Mary likes him_i]

(70) *Mary said that [John_i likes him_i]

Principle C says that a referring expression must be free. It rules in cases like (71) and rules out cases like (72). This prohibits a pronoun or reflexive being structurally higher than its antecedent.

(71) She_i likes Mary_j.

(72) *She_i likes Mary_i.

Crain *et al* have shown that by five years children are effectively adults with respect to these principles, ruling out, for example, phenomena such as backwards anaphora in (73) which is blocked by Principle C.⁴⁶⁶

(73) *He_i ran while the guard_i shouted.

may be of borderline acceptability). These “picture reflexives” may be marginal as counterexamples to UG but there are further cases of contrastives like “Bill can’t imagine why Mary would want anyone other than himself”. Even here judgements may be marginal and the sentence may sound better with “him”. Moreover, whether these cases are genuine counterexamples will depend on their proper analyses. See Kayne (2002) for further discussion and see Boeckx (2006) pp.105-109 for a discussion of reconstruction effects that once seemed to violate Principle A.

⁴⁶⁶ See Crain and Mckee (1985), and Crain and Thornton (1998). A growing body of experimental evidence indicates that children have substantial knowledge of the three principles of binding theory; they can distinguish between reflexives and non-reflexives and know the local domain in which binding conditions apply. See Guasti (2002) ch.8 for discussion of some of the mistakes that children do make with respect to binding.

We can take the principles of binding theory to be the G in POS1, a subset of the grammatical principles that all normal children acquire.

Experimental studies also show that children of a young age have near flawless grasp of the structure-dependent rules involved in Yes/No question formation.⁴⁶⁷ Yes/no questions are questions of the following form:

(74) Is Gareth gone?

(75) Will you pass the salt?

These questions are clearly related to the following declarative sentences.

(76) Gareth is gone.

(77) You will pass the salt.

Such questions appear in the data to which young children are exposed. So we might think it is a simple task to work out the rules for forming such questions. But notice that from the related declarative sentences in (76) and (77), there are any number of rules that could be used to form the Yes/No question in (74) and (75). To mention just three, the rule could be (i) swap the first two words around, or (ii) swap the first verbal element with the first noun phrase, or (iii) front the auxiliary to the main verb. The correct rule is the third one. If the child were to adopt (i) then they should form the question (79) from (78).

(78) The field is full of grass.

(79) *Field the is full of grass?

Alternatively, if they were to adopt (ii) then they should form the question in (81) from (80).

(80) The man who is shouting is angry.

(81) *Is the man who shouting is angry?

⁴⁶⁷ See Crain and Nakayama (1987).

However, Crain and Nakayama conducted a series of experiments and showed that the errors expected under these hypotheses never occurred. This is true for children as young as three years and two months. All the children used the third and most complicated rule to produce the correct results, as in (82) and (83).

(82) Is the field full of grass?

(83) Is the man who is shouting angry.

So, there are two candidates for G in POS1: binding and Yes/No question formation. In each case, as Laurence and Margolis note, “The correct set of principles isn’t simple or natural in any pretheoretic sense. This means that the empiricist learner has to rely on there being sufficient environmental information to guide her through the vast number of competing sets of principles.”⁴⁶⁸ That the actual principles are not simple or natural in a pretheoretical sense is true of grammatical principles and their interaction across the board. As Chomsky pointed out with respect to empty categories, no one would claim that they are the most natural generalisations to arrive at upon presentation of linguistic data. This point becomes clearer as we move beyond relatively construction-specific rules (like the auxiliary inversion rule I described above) to more general and explanatory, but highly grammar-specific principles not suggested by simple analyses (such as the theory of movement that explains auxiliary inversion and a host of other phenomena).⁴⁶⁹

I’ll split the evidence that nativists draw upon to defend the key premise POS3 into issues concerning the nature of learning and of PLD.

Learning

The first issue I want to consider is what learning theories imply about the errors that we observe and, in particular, the errors that we do *not* observe in children’s

⁴⁶⁸ Laurence and Margolis (2001) p.222

⁴⁶⁹ These more general principles unify a range of phenomena but work in complex interaction with one another. So there is more chance of a child learner discarding a correct principle when they produce an incorrect construction as they hypothesise the other principles and their mode of interaction.

linguistic behaviour. Children exemplify competence with the binding principles but, tellingly, they don't ever try out certain more simple but mistaken hypotheses about binding. Children never try out a principle according to which there is no binding, so that pronouns and reflexives are never bound linguistically. Equally, children never try out the principle that binding is possible throughout a structure, independently of local domains. And they never adopt structure-independent rules for binding based on, say, word order. What strikes the nativist is that if children are learning from PLD then there should be *some* stages of exposure to the PLD at which at least *some* children take binding to be determined by alternative principles. But, on the contrary, they take binding to be determined by three principles involving hierarchical structure and the notion of locality or domain. In the case of binding, as with a very broad range of grammatical phenomena, children do not experiment with what might be the most natural generalisations at a given stage in grammar acquisition.

This is why nativists insist upon the inclusion of all normal children acquiring G 'unerringly' in POS1. The question is why children should converge so uniformly and unerringly on the correct structure-dependent binding principles rather than alternatives such as no grammatical principle, permissive grammatical principles or principles based on word order. If they are generalising from PLD as they are exposed to it, then we would expect them to try out some incorrect principles.⁴⁷⁰

The nativist's claim is not that GPLM are incapable of learning any principles. Rather, the nativist claims that there is a critical issue about why children would unerringly acquire such structure-dependent principles. Consider (84) and (85).

(84) Anne asked whether Mary_i wanted to shake hands with herself_i.

(85) Anne asked who Mary_i wanted ϵ to shake hands with herself_{i/j}.

The referential dependency of *herself* on *Mary* in (84) but not (85) is explained by Principle A. But consider how hard it would be to induce the referential dependency

⁴⁷⁰ Laurence and Margolis (2001) and Segal (forthcoming) p.6 press this point.

of *herself* on *Mary* in (84) and the failure of referential dependency in (85) from the PLD. From the perspective of a learner, strings like (84) have misleading properties for determining whether there is referential dependency in (85), and vice-versa. In (85) the verb *to shake* has a phonetically null subject bound by *who* (of which it is a copy). Hence, in (85) *herself* could be substituted for *himself* but not in (84). The nativist insists on ‘unerring’ in POS1 because there is no evidence that we ever try out other simpler principles with respect to complex structures like (84) and (85), or ever interpret the latter (85) with *Mary* binding *herself*.

The evidence is that mistakes are very rare relative to what we would expect if the child were using GPLM to develop generalisations and correct the erroneous generalisations that would inevitably result on the basis of only partial exposure to PLD. Children make very few systematic errors in acquiring principles of great subtlety.

This presents a difficulty for learning theories on which the child generalises from the growing corpus because the child should frequently come across sentences like (86) and generalise about sentences like (87) on that basis.

(86) Mary wanted to play games by herself.

(87) Jim wondered who Mary wanted to play games by herself.

Sentence (87) looks as if it has (86) in it. But it doesn’t because (87) is actually structured like (85) with the movement and the empty category which is the subject of the infinitival clause. From the perspective of a learner, the occurrence of structures like (85) and (87) revealing of the subtler principles ought, at least sometimes, to be counted as anomalous by the child. This is especially probable when the relevant sentences are statistically rare in the corpus, as we should expect them to be during at least *some* stages of grammar acquisition for *some* children.

Wanna-contraction provides a clear illustration of why nativists claim that misleading properties of the PLD would, at least some of the time, lead GPLM into error.⁴⁷¹ The contraction in (88) but not (89) is permissible.

⁴⁷¹ See Crain and Pietroski (2001) for discussion.

- (88) Who does Arnold want to make breakfast for?
(88a) Who does Arnold *wanna* make breakfast for?
(89) Who does Arnold want to make breakfast?
(89a) *Who does Arnold *wanna* make breakfast?

The explanation of these facts about *wanna*-contraction is that *wh*-questions are formed by the movement of the *wh*-phrase to the position at which they are pronounced. An empty category is left behind as a record of the position from which the *wh*-phrase has moved. This empty category is in object position in (88) which is interpreted much like *Arnold wants to make breakfast for WHO?* But the empty category has been left in the subject position of the clause in (89) which is interpreted just like *Arnold wants WHO to make breakfast?* Given their knowledge of the underlying representations, linguists offer the hypothesis that the empty categories left behind by subject extractions block *wanna*-contraction. The generalisation that subject extractions block *wanna*-contraction was obscured from linguists by the fact that the contractions are permissible in structures involving object extraction. Until a wide range of data was considered systematically, and the theory of movement and empty categories had been developed, linguists failed to see how the two phenomena were related. The child learner would be in a similar position when confronted with contractions like (88) to (88a). Such contractions should constitute, *prima facie* at least, evidence that runs counter to the right hypothesis that contraction is impermissible in (89). But the child doesn't make the mistakes we would expect if they were generalising from such cases.⁴⁷²

Children do sometimes try out alternative, mistaken options concerning a grammatical principle on their way to mature competence, some of which nativists have explained as options evidenced in other UG languages. A noticeable feature of such mistakes is that they do not exhibit principles that are less subtle than the mature principles. But it might be claimed that the very existence of mistakes

⁴⁷² Experiments by Thornton (1990) and (1996) confirm children's early adherence to the constraint. In a group of 21 children, mean age 4.3, 57% of the time the child chose to contract on the permissible (88) but less than 10% for the impermissible (89). Thornton's (1990) elicited production study on 12 children aged 2-4 evidenced 100% absence of contraction over the *wh*-trace.

provides some support for learning theories, according to which the child has made an incorrect generalisation. This is not so. If the acquired grammar is learnt then it must not be learnt using GPLM that would lead to mistaken but incorrigible generalisations.⁴⁷³ The alternative to nativist explanations of why children do not generalise from (91) entailing (90), to (93) entailing (92) is that children generalise from the growing PLD and correct their rules so as to block the entailment.

(90) John ate.

(91) John ate a fish.

(92) John is too clever to catch.

(93) John is too clever to catch a fish.

Sentence (90) means that John ate something. Looking at (90) and (91) one might naturally conclude that if a verb has both transitive and intransitive forms and appears without an argument in object position, then it is interpreted as having an implicit arbitrary object of the appropriate category. But if you applied that to (92) it would get the interpretation that John is too clever to catch something, which is wrong because the sentence means that John is too clever for one to catch him. Standard forms of reasoning would find a rule that applies to the simple cases and then apply it to the more complex ones. But that's not what children do.⁴⁷⁴ Children never make the mistake of understanding (92) such that it is entailed by (93). But suppose they did. The GPLM would have to be such as to correct the entailment despite the statistical preponderance of examples like the entailment of (90) by (91).

Grammatical principles are not only acquired unerringly, they are also acquired by *all normal children*, who all acquire grammars of the same complexity. If the child is solving a learning problem then it must be insensitive to not only variation in the PLD, but also to variation in general intelligence and background culture. Belying the subtlety of the acquired principles, grammar acquisition *is* highly insensitive to general intelligence. It is far less elastic with respect to

⁴⁷³ It is worth noting, however, that nativists owe us a detailed account of how children witch back an incorrectly set parameter where that parameter results in a sub-set language.

⁴⁷⁴ The example is Chomsky's (1986) and the explication from Segal (forthcoming).

intelligence than, say, the child's scientific, mathematical and musical abilities. In the most extreme cases, children with William's syndrome have intact grammatical competence and severe intellectual impairment. Yet, despite considerable latitude in intelligence, children all acquire a grammar characterised by principles of the same depth and complexity.

Moreover, the 'syntax spurt' is not correlated with the development in other areas of knowledge that would be consistent with a rapid improvement in GPLM. Laurence and Margolis endorse this point:

We are all extremely impressed if a two-year-old figures out how to put the square blocks in the square holes and the round blocks in the round holes. Yet somehow by this age children are managing to cope with the extraordinarily difficult task of learning language. If the empiricist is right, we are to believe that children do both of these things using the very same domain neutral intellectual resources.⁴⁷⁵

Further, any special levels of attention and coaching that are bestowed on only *some* children in the relevant window of development can't be the fundamental explanation of why *all* children acquire a complex grammar. This is true even if we assume that such special attention and coaching are effective aids to learning in the cases in which they do occur. What strikes the nativist is that a learning explanation would have to appeal to processes that *all* children employ.

Nativist explanations predict the unerring nature of acquisition and its uniformity across all normal children by appealing to innate grammatical principles that constrain all children's development of a grammar. Learning theories do less well on these counts. They predict errors in generalisations and a series of corrective hypotheses as part of a learning process. And we do not observe such sensitivity of grammar acquisition and the acquired competence to differences in general intelligence and aids to learning.

Learning theories also struggle with the *structure-dependence* of grammatical principles. Learning theories involve the child relying on cues in the PLD concerning word order, morphology and semantic plausibility.⁴⁷⁶ According to

⁴⁷⁵ Laurence and Margolis (2001) p.237

⁴⁷⁶ See, for example, Macwhinney and Bates (1989).

their frequency in the PLD, the learner is hypothesised to assign different probabilities to lexical items following one another, and from these resources the child is supposed to build his grammatical constituents and principles. But such statistical generalisations struggle to capture the structure-dependent relations that hold amongst grammatical constituents. Consider the following possibilities for forming Yes/No questions.

(94) The man is sad.

(94a) Is the man sad?

(95) The man who is beating a donkey is mean.

(95a) *Is the man who _ beating a donkey is mean?

(95b) Is the man who is beating a donkey_ mean?

(96) The man who is beating a donkey whilst he is shouting and who is incredibly red-faced is mean.

(96a) *Is the man who beating a donkey whilst he is shouting and who is incredibly red-faced is mean?

(96b) *Is the man who is beating a donkey whilst he shouting and who is incredibly red-faced is mean?

(96c) *Is the man who is beating a donkey whilst he is shouting and who incredibly red-faced is mean?

(96d) Is the man who is beating a donkey whilst he is shouting and who is incredibly red-faced mean?

The structure-dependent rule requires the child to front the auxiliary to the main verb in each case. Fronting the correct auxiliary requires the child to see its relation with the subject NP and main verb and ignore the embedded relative clause. This is more complicated than the structure-independent rule that says simply *front the first auxiliary*, which makes no appeal to verbal domains. This rule itself involves a grammatical category of verbal structure that would have to be acquired. An even simpler rule would be to search the declarative until the first “is” (or “can”) is found

and move that to the front. This latter rule is completely structure-independent in that it appeals purely to morphology and linear order.⁴⁷⁷

The structure-independent rule would cover (94) but fail to cover (95) and (96), leading to ungrammaticality because the auxiliary that has been moved is not from the main clause but rather one embedded in a relative clause. Pairs of declaratives and questions of the form (94) and (94a) are a part of most children's PLD. But the declarative and question pairing for forms like (95) might not be part of the PLD, and this is quite likely for the declarative and question pairs for sentences like (96) which involve two relative clauses. Yet on the basis of their exposure to many pairs like (94) and (94a) children never form the structure-independent rule, they go straight for the structure-dependent rule.⁴⁷⁸ Children never make the mistakes forming questions from sentences like (95), or even (96), that an attention to patterns of word order would suggest: natural generalisations would suggest structure-independent rules leading to (95a) or (96a). And there is no reason that the statistical learner wouldn't formulate structure-independent rules first, generating (96b) or (96c).⁴⁷⁹ Even if the child were confronted with (95b) and (96d) we should ask why they were not counted as anomalous on the basis of the preponderance of structures like (94a).

It is difficult to explain why GPLM, without the antecedent information on phrasal structure, would support the child's tendency to immediate plump for

⁴⁷⁷ Though Crain (1991) notes the presence of errors in child question formation, Crain and Nakayama's data do not attest to a single error of the sort that suggests a structure-independent rule, which is striking. Crain and Nakayama (1987) investigated 3-5 year olds to see whether they made the sorts of errors that would be expected if they were entertaining structure-independent principles. None of the thirty children made a single error of the sort associated with the structure-independent hypotheses. This is strong evidence that the children are not entertaining such principles before ruling them out with further PLD.

⁴⁷⁸ Even amongst the structure-dependent possibilities there may be an abundance of alternatives that associate the movement with a different structural variable. The child also has to discard the principle that it is optional which auxiliary is moved, which, as Laurence and Margolis (2001 p.229) point out is "compatible with any sentence that the child is likely to hear, and languages do seem to contain optional phenomena of this sort."

⁴⁷⁹ See Stromswold (1999 p.361) who points out the difficulties of distinguishing these auxiliaries from lexical verbs.

structure-dependence, applying it to long distance dependencies in forming questions from (95) and (96). If the child is innately constrained to develop the categories and structural principles that underlie the structure-dependent auxiliary inversion rule then we can explain why they are capable of dealing with both the monoclausal case of (94), as well as cases like (95) and (96) with embedded relative clauses.

The last point I'd like to make about learning is that we have independent reasons to doubt that children acquiring language perform the task of a GPLM. Simple GPLM can form generalisations that a child cannot. For instance, seven year old children cannot learn the structure-independent rule *drop the first four words of every sentence*. But learning machines can learn such rules easily and humans can easily grasp structure-independent patterns outside the domain of language. If a child of seven cannot cope with learning and applying a range of simple structure-independent procedures (but is highly proficient with structure-dependent grammatical rules), this casts serious doubt on an account of grammar acquisition according to which the child uses structure-independent grammatical rules as the basis on which to learn structure-dependent grammatical principles, and then expunges the structure-independent rules. Though learning machines can learn some properties of the languages that children can acquire, this may be as far as the comparison goes. Such learning machines struggle to learn principles that children effortlessly acquire and are capable of learning things that children cannot.⁴⁸⁰

⁴⁸⁰ In the context of what children can do that machines can't, we might also consider what children can do, that aphasics with language impairment can't. There are aphasics with extreme limits to their grammatical abilities who can perform normally in other cognitive domains (see Varley 1998). As these aphasics are otherwise unimpaired, according to learning theories they ought to be able to relearn their grammar using GPLM. But they can't (see Veletti-Glass et al 1973). This suggests that children acquire grammars with the aid of innate grammar-specific principles. For similar conclusions on dissociations see van der Lely and Stollwerck (1996), van der Lely et al. (1998) and (2004). This work suggests that there are severe grammatical impairments (affecting 3-6 people per 1000) that are domain-specific.

PLD

There have been several recent critiques of POS arguments claiming that they rely on an unduly pessimistic assessment of PLD.⁴⁸¹ Whilst all parties agree that PLD provides a finite and individually variable set of utterances, a common theme amongst critics of nativism is that nativists have not given PLD sufficiently careful and thorough assessment. Critics take this as evidence that nativists are likely to have overlooked significant properties of the PLD. Critics, quite reasonably, wonder how nativists can be confident that children do not learn the abstract principles of grammar if they don't know precisely what is in the PLD that could serve as a basis for learning. It would significantly weaken the POS argument if it only went through on dogmatic assumptions about PLD.

Nativists do care about investigating PLD, for instance, when investigating what sorts of mistakes children make. But it is true that nativist theories have focused more on the details of the grammars that are acquired and less on the nature of the PLD that a child might be exposed to. This focus might be taken to suggest that nativists are prepared to grant opponents very generous assumptions about PLD. For nativists do not think that their arguments are contingent on scouring PLD to eliminate learning theories on a piecemeal basis.

POS arguments do not generally rely on an inference from *no* data relevant for a particular construction turning up in a child's PLD. The argument that Pullum and Scholz consider is that children *never* come across *any* evidence in their PLD that would be relevant to ascertaining the correct rule for a particular construction.⁴⁸² Though it is wrong to attribute such an argument to nativists, two aspects of Pullum and Scholz's case against this argument are noteworthy. Firstly, Pullum and Scholz identify evidence that certain constructions do occur in corpuses. But Pullum and Scholz never consider how the occurrence of particular constructions would serve as evidence for the principles that generate such constructions. This may be because, as Crain and Pietroski claim, the construction rules that Pullum and Scholz seek evidence for are fairly superficial and

⁴⁸¹ See Pullum and Scholz (2002) and Cowie (1999). But see also Chouiard and Clark (2003) and Foraker et al (forthcoming).

⁴⁸² Pullum and Scholz (2002)

construction specific.⁴⁸³ For instance, their rules covering NPIs do not appeal to the underlying c-command relations that explain *why* NPIs are licensed. These principles connect the explanation of NPIs with a host of other phenomena, including binding principles. Secondly, Pullum and Scholz's case against the limited argument they present, involves claims about what is available in the child's PLD based on corpuses that are inappropriate. Their claims are based on broadsheet journalism and classics of 19th century English literature. They base this selection on the unwarranted assumption "that research on one kind of text will tell us at least something about the syntactic structures likely to be found in others."⁴⁸⁴ But one might be sceptical that such material serves as a guide to the properties which all children are exposed to or could utilise as a basis for learning. For instance, Pullum and Scholz appeal to the presence of question formation across more than one relative clause in an Oscar Wilde play as evidence that young children are typically exposed to such constructions.

Though some nativist arguments do appeal to principles that occur so early in a child's development that there is arguably no corresponding data available in the PLD, such cases are taken as *evidence* for nativism, not as the situation that must obtain for each UG principle for POS to obtain.⁴⁸⁵ Nativists do not think their arguments vulnerable to foraging in PLD because they do not think that learning constitutes an explanation of the facts about grammar acquisition, as I described above, but also because of quite broad properties of PLD considered as a basis for such learning.

POS arguments do not rely on any special claims about the data each child might encounter. But POS arguments do rely on the claim that if a child were acquiring its grammar by learning from PLD, without the aid of an innate set of grammatical principles, then their PLD would have to be "absurdly rich" in ways that it is not.⁴⁸⁶ For a child to learn the specific grammar it acquires, the child would need data on each choice which would serve to differentiate its target

⁴⁸³ Crain and Pietroski (2002) pp.9-10

⁴⁸⁴ Pullum and Scholz (2002) p.23

⁴⁸⁵ See Crain and Pietroski (2001) for a discussion of cross-linguistically attested constraints that emerge so early that there is no widespread data in the PLD and arguably no data at all.

⁴⁸⁶ Collins (2003)

language from closely related languages it might acquire. The child would also need to eliminate choices that lead to no human language.

If the child requires information rich enough to eliminate *every* possible alternative for the binding principles, for Yes/No questions and for *wanna*-contraction at each stage of exposure to PLD; then the nativist might plausibly claim that children are not exposed to such rich data on each grammatical construction with which they acquire competence. But equally importantly (as we saw with binding, question formation and contraction) the PLD serves to cast doubt on the correct hypotheses. Apart from the presence of perfectly grammatical expressions that suggest the wrong rules, it can also be the case that, as Chomsky notes, some of the “observed speech consists of fragments and deviant expressions of a variety of sorts.”⁴⁸⁷ Yet every normal child develops competence with all the constructions “even though the primary linguistic data that he uses as a basis for this may...be deficient in various respects.”⁴⁸⁸

If we combine the uniformity of the grammars children acquire with the variation in the quantity and content of their PLD, we can see why the POS argument does not turn on whether specific constructions do or do not happen to occur in a child’s PLD, for:

[W]e might not know much about the PLD, but it is patently not uniform in the sense of every child’s PLD including every potentially falsifying conclusion up to the correct principle.⁴⁸⁹

When we consider the uniformity amongst populations of speakers it seems that PLD not only needs to be extremely rich, in order to reveal to the child the very abstract principles, but also highly uniform from child to child. Given the large variation in the PLD on which each child’s hypotheses would be based, learning theories predict variation in the mature rules, simply because the generalisations are based on different data. Though there is limited grammatical variation, even among speakers in a single speech community, it is highly restricted. Indeed, as Collins

⁴⁸⁷ Chomsky (1965) p.200 fn.14

⁴⁸⁸ *ibid.*

⁴⁸⁹ Collins (2003) pp.12-3

highlights, according to recent minimalist theories all variation is morphologically driven due to differences in lexical features, whilst the grammatical operations and interface conditions are univocal.

Building a case that broad features of PLD serve to support POS arguments, the nativist can appeal to striking cases of grammar acquisition where any relevant PLD seem to be almost entirely absent. Two such cases are that of Creoles acquired by children brought up in Pidgin speaking communities, and that of sign languages acquired in isolation from other signers. These are both instances of grammar acquisition where none of the principles that organise a mature grammatical competence are present in the child's PLD.

A Pidgin is a system of verbal communication that develops when speakers of a number of different languages are brought together, and have to get along using isolated words and phrases without the full grammatical structure of their native language. In Pidgin-speaking communities there is no UG language present in the PLD with the sort of grammatical structure possessed by the individual languages of the adult participants in the Pidgin. But children brought up in Pidgin speaking communities do not acquire the Pidgin, or their parents' native languages, but rather a Creole. These children are surpassing the Pidgin and "inventing a new natural language, a Creole, for which they effectively have *no model at all*."⁴⁹⁰ A Creole is a parameterised UG language with a full generative grammar. Similarly, there are deaf children brought up in the absence of other signers who acquire sign languages of similar grammatical complexity to the spoken languages acquired by their peers.⁴⁹¹ These cases suggest a preponderant role for grammatical principles emerging independently of shaping effects from experience. In these cases, grammar acquisition doesn't require much of the environment beyond decent physical conditions, absence of trauma, and some linguistic stimulation. Nativists argue that this is borne out when we consider the nature of PLD in more mundane cases.

⁴⁹⁰ Laurence and Margolis (2001) p.243

⁴⁹¹ See Feldman, Goldin-Meadow and Gleitman (1978)

Early inquiries into PLD suggested a lack of information (explicit or implicit) about the *ungrammatical* expressions of natural languages.⁴⁹² Learning theories rely on information being made available, explicitly or implicitly, to *all* children about the ill-constructed, non-expressions in the PLD. Accordingly, some nativists have pressed the point that even if idiosyncracies of a child's PLD tell them about the grammatical expressions of their language, then learning theories have to find corresponding information in the PLD to tell them which expressions within the PLD are ungrammatical, hence not a part of their target language. Unless such information is available, then the possibility that all of the utterances within the PLD are grammatical should at least be open to the child. But this possibility is not available. Nativists, therefore, challenge learning theorists to provide evidence that there is sufficient information amongst PLD for effectively ruling out the ungrammatical expressions, and not *an absence of negative evidence*.

The issue concerning negative evidence can be developed in the following way. If the child learner were conservative in their generalisations, admitting into their grammar principles that generate *only* those expressions encountered in the PLD, then the child would massively *undergenerate* with respect to the sentences of their target language. So, if the child is to produce and understand a wide range of novel utterances in the target language, then they must project beyond the PLD. But then we should expect the child, at least some of the time, to *overgenerate* and produce a deviant superset of the expressions generated by the target grammar. Given that the child does, in fact, acquire the target grammar, any overgeneration must be corrected for. A major challenge for learning theories of grammar acquisition is then to explain how children recover from such attested grammatical errors as "I makes it with water" and "And fill the little sugars up in the bowl".⁴⁹³

⁴⁹² Though more recent work by Chouinard and Clark (2003) serves at least to cast doubt on some of these early claims about the absence of negative evidence.

⁴⁹³ The examples are taken from Marcus (1993). As Marcus points out (1993 p.80, also Collins 2003) it is a mistake to see nativist arguments as depending on a lack of negative evidence. For nativists claim that the many plausible errors, which negative evidence would be required to eliminate, never in fact occur. The learning theorist has a problem with eliminating these plausible errors, but even if perfect negative evidence is available, there would still be an argument for nativism on the basis that the natural errors never occur, as has been well-documented with binding and question formation

So, we need to consider what would happen if the child really did overgenerate on the basis of their individual PLD. Nativists claim that, where the errors are systematic, they are the result of the child switching their parameter settings so as to generate these errors which are corrected on the basis of internal mechanisms. According to learning theories, the child uses *negative* evidence to eliminate the incorrect principles of their intermediate states. But nativists claim that while PLD is an impoverished source of positive evidence for the child learner, it is barely a source of negative evidence at all. The natural assumption is that all normal children receive an abundance of negative data as part of their PLD in the form of correction from adult speakers. But this assumption is false. Many of the world's children receive no explicit correction at all.⁴⁹⁴ Yet, they all acquire a grammar.

Moreover, the focus and effect of parental correction is far from clear. Brown and Hanlon carried out two analyses of hundreds of hours of tape of English children interacting with their parents, in order to determine whether parents provide feedback contingent on children's grammatical errors. In the first analysis, they examined whether parents understand their children more easily if a child's question is grammatical. Parents failed to understand about as many grammatical questions as ungrammatical ones. Parental replies indicating understanding were equally likely following grammatical and ungrammatical speech. In their second analysis, Brown and Hanlon examined parental replies indicating approval and disapproval. Again, there was no relation between parental reply types and child grammaticality. Brown and Hanlon found not "even a shred of evidence that approval and disapproval are contingent on syntactic correctness."⁴⁹⁵ The bases they did find for disapproval were always semantic or phonological. The very rare approval or disapproval of grammatical form suggests that this is not the force

(see Crain and Mckee 1985, Crain and Nakayama 1987, Crain 1991, Crain and Thornton 1998 and Crain and Pietroski 2001).

⁴⁹⁴ Slobin (1972) found that children are not corrected for errors in many of the societies his group studied. Chomsky (1965 p.200-1 fn.14) made this a part of POS arguments for nativism: "It seems clear that many children acquire first or second languages quite successfully even though no special care is taken to teach them and no special attention is given to their progress."

⁴⁹⁵ Brown and Hanlon (1970) p.47

driving children towards their mature grammar. Brown and Hanlon concluded that “In general, the results provide no support for the notion that there is a communication pressure favouring mature constructions.”⁴⁹⁶

Pinker found that where they exist, both parental sensitivity to grammatical errors and corrective behaviour are “noisy, indiscriminate and inconsistent from child to child and age to age.”⁴⁹⁷ Moreover, he found that children are persistent in their errors and remarkably insensitive to the little correction received. Bowerman found little negative evidence available that was relevant to the specific cases of overgeneration that occur.⁴⁹⁸

Some researchers have disputed that there is a lack of negative evidence.⁴⁹⁹ These studies report evidence from the distribution of certain patterns of discourse between parents and children. As such they provide at least a basis on which to argue that children make use of negative feedback to acquire their grammar. Cowie argues that this negative feedback provides a mechanism by which children could hone a correct grammar.⁵⁰⁰

The kind of evidence which these researchers investigated, and which Cowie claims could serve as a basis for learning grammar, is *implicit negative evidence*. This includes a range of different parental behaviours such as repetitions, questions and recasting that might influence the child but not *explicit* information on ungrammatical expressions. For this evidence to provide a basis for acquisition, it must be ubiquitous and the process by which it is utilised must not rely on any special intelligence, attention or memory that is not possessed by all normal children. Cowie claims that there is an abundance of such implicit negative evidence in the PLD, which could be used to acquire the grammatical principles and eliminate non-expressions.

Cowie cites studies which show that mothers of two year-olds in some cultures repeated and corrected their child’s ill-formed utterances 20% of the time

⁴⁹⁶ Brown and Hanlon (1970) p.45

⁴⁹⁷ Pinker (1990) p.217

⁴⁹⁸ Bowerman (1988)

⁴⁹⁹ Bohannon & Stanowicz (1988), Chouinard and Clark (2003), Demetras, Post and Snow (1986), Hirsh-Pasek, Treiman and Schniederman (1984), Morgan and Travis (1989).

⁵⁰⁰ Cowie (1999) ch.8

but repeated well-formed utterances only 12% of the time. These studies form the basis of Cowie's challenge to Brown and Hanlon's conclusion. But as 20% correction means that 80% of mistakes go uncorrected, Cowie's idea is that children are sensitive to the difference between the 20% of ill-formed utterances repeated and the 12% of well-formed utterances repeated, and that they are capable of keeping track and determining what aspects of their utterances a parent is responding to. However, the studies Cowie cites are not generally considered to provide evidence of the child's employment of such negative evidence.

As Marcus details, such implicit negative feedback as exists is *noisy*: the differences in parental responses are statistical rather than categorical. Neither *complete* feedback where parents provide a corrective signal for all and only ungrammatical utterances, nor *partial* feedback where parents provide corrective feedback after only ungrammatical utterances actually exists. Parents who provide feedback provide each type of response after both grammatical and ungrammatical sentences though in different proportions. Therefore, the child would have to determine the status of responses. For each response, the child has to work out whether it is a response to grammaticality or ungrammaticality. Marcus cites Penner's study which found that parents "repeated correct utterances slightly more frequently than incorrect child utterances" but that this pattern "was not consistent for all parents."⁵⁰¹ Any mechanism that the child employs to effect such discriminations must not appeal to the child's prior knowledge of the grammatical sentences of their language, for this is precisely the knowledge the child is trying to acquire: "The child cannot simply record how many of her ungrammatical versus grammatical sentences elicit instances of some reply."⁵⁰²

Given the low percentage of errors that are followed by feedback, one might wonder whether the feedback correlates with recovery from error. Marcus' study of negative evidence concludes that where noisy feedback is available, it is too weak to act as the mechanism for acquiring grammatical principles: "a child would have

⁵⁰¹ Penner (1987)

⁵⁰² Marcus (1993) p.65

to repeat a given sentence verbatim at least 85 times to decide with reasonable certainty that it is ungrammatical.”⁵⁰³

Further, Marcus concludes that no feedback, even of the noisy variety, is provided to all children, or at all ages. So not only is the noisy negative evidence very weak but it also only appears in some mothers with young children. It was not, for example, found in all the pairs studied by Bohannon and Stanowicz, and is most typical of the Western middle classes.⁵⁰⁴ Even if negative evidence is fed back to *some* children on an expression-by-expression basis, and we had some evidence of their sensitivity to it, this could not serve to explain the uniformity of grammar acquisition across *all* normal children.

Cowie misses this point arguing, that it doesn't matter that implicit negative evidence is not available to all children, as this does not imply that it is not used by *some* children.⁵⁰⁵ She claims that to ignore the role of implicit negative evidence would be too heavy-handed because she thinks that for all we know some children might make use of feedback even if others don't receive it or make use of it.

There are two problems with Cowie's response. First of all, implicit negative evidence was being proposed as the *crucial* mechanism for children to correct errors and arrive at the mature competence. Secondly, the claim that children do exploit different forms of implicit negative evidence is not supported by the existence of very weak negative evidence in some cultures. There is no basis for the claim that Cowie uses to supplement this reasoning when she suggests that where one form of negative evidence isn't available then another probably is.⁵⁰⁶ That most normal children acquire grammar without such evidence is suggestive that those who do have it do not need it and this is substantiated by Marcus' results.

Even amongst those children who receive noisy feedback, it is not provided for all types of errors. Noisy feedback is available to some children in a fraction of the circumstances in which it would be needed. Moreover, Marcus argues that noisy feedback may be partly an “artefact of defining parental reply categories relative to

⁵⁰³ Marcus (1993) p.53. Marcus' study actually used the same data from Hirsh-Pasek et al that Cowie (1999) appeals to as a basis for learning.

⁵⁰⁴ Bohannon and Stanowicz (1988)

⁵⁰⁵ Cowie (1999) p.232

⁵⁰⁶ Cowie (1999) p.231

the child's utterance...because nearly all parental speech is grammatical, exact repetitions...necessarily follow more of children's grammatical utterances than their ungrammatical utterances."⁵⁰⁷ Marcus concludes that there is no evidence that noisy feedback is required for grammar acquisition.

Marcus also highlights independent evidence that there is no correlation between provision of negative evidence and speed of grammar acquisition.⁵⁰⁸ At most, Cowie's argument could suggest that there might be data in some children's PLD to eliminate some false hypotheses. It's not part of POS arguments to deny this. What nativists deny is that PLD is sufficiently uniform to always contain enough crucial data, and that children in fact make use of it.

One proposed substitute for negative evidence is *Motherese*.⁵⁰⁹ Motherese is simplified, caretaker speech with exaggerated intonation and distinctive prosody that might make structure more perspicuous, used with young children. Whilst Motherese certainly makes word boundaries more perspicuous it is unclear that it makes phrasal categories so.⁵¹⁰ And similar issues apply to Motherese as I raised about implicit negative evidence. There is evidence in *some* cultures that adults simplify their speech and use cues such as rising intonation that might make structure more perspicuous. But there is also evidence that in some cultures parents do not talk directly to their children much at all.⁵¹¹ Of course, sometimes Motherese will contain grammatical errors precisely because it seeks to simplify and convey a message rather than maintain grammatical structure. The existence of Motherese for some children doesn't mean that all or most of what those children are exposed to will have the properties of Motherese. There is no evidence that differential exposure to Motherese is correlated with different rates of grammar acquisition.⁵¹²

Cowie thinks that once we come to a more moderate evaluation of the PLD this will help us see how grammatical principles could be honed using implicit negative evidence, but the evidence she has provided for this view is insubstantial.

⁵⁰⁷ Marcus (1993) p.53

⁵⁰⁸ Newport, Gleitman and Gleitman (1977)

⁵⁰⁹ Cowie (1999 p.190) appeals to Motherese.

⁵¹⁰ See Pinker (1987)

⁵¹¹ Heath (1982)

⁵¹² Newport, Gleitman and Gleitman (1977)

However, motivating her view is a comparison between our knowledge of curries and our knowledge of grammar. Though the child has to project from PLD in highly specific ways and find materials amongst that data to correct overgeneration, Cowie claims there is no special issue here. On exposure to a few examples, and with no clear errors en route, all normal children will come to know what curries are, reliably classifying dishes as curries or not. How does the child's knowledge extend beyond the *primary curry data* so that the child counts in all and only the curries? Learning about curries is possible, Cowie suggests, because there is a vast amount of implicit negative evidence available from which to generalise. The child might pick up on the fact that burgers are generally called "burgers" and not "curries". This suggests to the child that burgers are not curries. And the same goes for pizzas, stir-fries and casseroles. This is *Cowie's Curry Argument*:

Just as there are many sources of negative evidence in the data concerning curries, so there must be substantial sources of negative evidence in the data concerning language.⁵¹³

Cowie thinks the "poverty" in POS an exaggeration in light of the curry argument. But of course the nativist about grammar never denied that people can learn things on the basis of implicit evidence. The question was always about how a child on exposure to PLD acquires the grammatical principles of binding, movement, empty categories and so forth. It is the abstractness, specificity and uniformity of what is acquired that impresses nativists because the principles seem so grammar-specific and so arbitrary from the perspective of a learner.

The curry argument suggests nothing about grammar acquisition beyond the fact that children do learn to discriminate between things. It does not engage *any* of the special issues about grammar acquisition that I have outlined: (i) the unerring acquisition of complex and abstract principles, (ii) acquisition by all normal children, (iii) acquisition irrespective of general intelligence, (iv) acquisition ignoring the most natural generalisations, (v) acquisition that involves grasping structure-dependence, (vi) the facts about what learners can do in other domains, (vii) Pidgins and Creoles, or (viii) the sorts of positive and negative evidence that

⁵¹³ Cowie (2004) p.222

would be required to learn grammatical principles. To set the POS apart from the general problem of making an inductive inference, one has to note not only that there are innumerable alternative hypotheses from the learner's point of view, but also that the correct hypotheses that children readily acquire are so abstract, arbitrary and languages specific - never the most simple or natural hypotheses that might to be selected by a learner - and are acquired under special conditions. Therefore, Cowie's curry argument doesn't present much of a challenge to the POS argument that I have outlined.

To meet explanatory adequacy we require a psychological theory of grammar acquisition. The best available theories of grammar acquisition are nativist theories which do not presuppose non-psychological grammatical properties but rather appeal to an innate psychological system of grammatical principles and parameters. The non-psychological stimulus (PLD) with which children are presented appears not to suggest the UG principles and the target parameter settings, but simpler and incorrect generalisations that children do not adopt. According to nativists, non-psychological properties act as a trigger for the child's development along a path determined by UG and its limited parameters. (Or on Yang's nativist model, the child imposes their innate grammatical categorisations on properties of input stimuli so as to shift probabilities across choices within the innate grammatical schematism.) The nativist holds out for an account of parameter setting but his opponent, the learning theorist, holds out for a theory of child learning that we have good reasons to think will not be forthcoming. As far as we can currently tell, in meeting explanatory adequacy generative grammarians do not focus upon, or presuppose, non-psychological grammatical properties.

5.4.3 Conventions

Devitt argues that external grammatical properties, established by convention, are required to explain language acquisition in the following way:⁵¹⁴

⁵¹⁴ Devitt (2006) p.181

(Pi) Language acquisition requires that the open parameters on the grammatical principles are set.

(Pii) Parameter setting depends on grammatical conventions.

Therefore,

(C) Grammatical conventions are required to explain language acquisition.

If sound, Devitt's argument for a role for grammatical conventions in meeting explanatory adequacy would provide some support for the non-psychological conception of generative grammar. The required grammatical conventions would constitute a non-psychological domain of grammatical properties. In my view, (Pii) is highly doubtful. Setting the parameters on grammatical principles does not require the child to be sensitive to grammatical conventions. Hence, (Pii) receives no support from the P&P model on which children have their parameters set in grammar acquisition. So we do not require grammatical conventions to explain grammar acquisition, and Devitt has not carved out an explanatory role for non-psychological grammatical properties within generative grammar.

We can see this if we consider that such conventions are required by neither the triggering model of parameter setting nor a statistical inference model. The triggering model requires such causal influence from external stimuli as is required to set predetermined paths of development in motion but clearly does not presuppose grammatical conventions. The statistical inference model requires such regularity in the external stimuli as the child depends upon to regularly impose the categories of their innate grammatical schematism. It may also be true that the triggering model requires regularities depending on what level of exposure to linguistic stimuli is taken to constitute "regularity". As Collins argues, if "regularity" means merely "exposure to the relevant stimuli within the relevant time period" then it is trivially true.⁵¹⁵ But, however specified, mere regularities in the external stimuli do not equate to grammatical conventions.

Devitt's main argument for the claim that parameter setting requires conventions is that a child's parameter settings are rarely eccentric with respect to

⁵¹⁵ Collins (2008a)

those of their linguistic community.⁵¹⁶ This is true enough. But Devitt infers from this that to set the parameters for the relevant grammatical constructions non-eccentrically, the child must pick up on grammatical conventions that are prevalent in his speech community. On the basis that children are born with some innate grammatical principles but not born to speak English or Japanese, Devitt concludes that the grammatical properties of speakers' languages are "largely conventional (although partly innate)"⁵¹⁷:

Acquiring a language is almost entirely a matter of moving, under the causal influence of primary linguistic data that are (performance errors aside) instances of local linguistic conventions, from an innate "initial state" of readiness for language to a "final state" of participation in those conventions.⁵¹⁸

That acquiring a grammar is "a matter of moving under the causal influence of primary linguistic data" to a final state of grammatical competence is unobjectionable. What is missing from Devitt's argument, as Collins has argued, is any explanation of why the child getting their parameters set in the same way as members of their community requires the child to be sensitive to grammatical conventions within that community.⁵¹⁹

Members of the population in which the child grows up have their parameters set in the same way. We can explain the child getting their parameters set on the basis that there will be regularities in the speech patterns of members of this population. Such regularities in their speech patterns will serve to determine future parameter settings: "Children exposed to speakers who have their parameters

⁵¹⁶ Devitt (2006) p.181

⁵¹⁷ Devitt (2006a) p.562. It is a mistake to suppose, as Devitt does, that because children are not born to speak a *particular* UG language, this is a reason to think some of the grammatical principles of their language are not innately specified and determined by conventions. Variation is perfectly consistent with the grammatical principles of languages being innately specified and a process of selection amongst the limited parameter settings.

⁵¹⁸ Devitt (2006) p.181

⁵¹⁹ Collins (2008a)

set to X, Y, Z will have their parameters set to X, Y, Z. Mere exposure to regularities will do the job.”⁵²⁰

According to Devitt’s view of conventions, a convention is a regularity where “there is some sort of mutual understanding.”⁵²¹ But why should we expect the setting of *grammatical* parameters to involve such “mutual understanding”? There does not appear to be any requirement for the child and his population to be engaged in any “mutual understanding” over grammatical matters, only that the population exposes the child to sufficient linguistic material. Hence, we lack a reason to think that parameter setting requires the child’s sensitivity to grammatical convention.

But there are further issues with regard to Devitt’s suggestion that parameter setting involves co-ordination over grammatical conventions, beyond their seeming dispensability. As the principles that characterise speakers’ grammatical competences are highly subtle and sophisticated, it is unclear in precisely what sense there might be a “mutual understanding” between speakers and children. So, we would want some explanation of the manner in which the relevant population, let alone the child in the early stages of grammar acquisition, are abiding by such regularities as the principles determine with “mutual understanding”. The explanation seems to be individual and cognitive rather than to do with shared understandings. Of course, Devitt must be supposing that the mutual understanding is tacit rather than conscious. But he will not want to weaken the notion of “mutual understanding” so much that external regularities determined by the individual competence systems could do the job. Devitt acknowledges the “difficulty of coming up with a satisfactory account” of grammatical convention, but says this “should not shake our conviction that there are such conventions.”⁵²²

But much more fundamentally, it is unclear that grammar acquisition requires the child’s presence in a community as might co-ordinate over conventions or even within a population that includes more members than just its mother. As Collins notes: “A child could acquire a language from its mother alone,

⁵²⁰ Collins (2008a)

⁵²¹ Devitt (2006) p.180

⁵²² *ibid.*

independently of whether the mother was according with a set of conventions or not.”⁵²³ Suppose that the child’s mother *is* acting according to conventions, though the child only interacts with its mother. We would want to know how the presence of conventions impacted on the child, and what explanatory or causal role the conventions were supposed to play in the child acquiring their grammar.

Chomsky argues that the only sense in which such public conventionality is a requirement on grammar acquisition that is actually supported by the facts, is the sense in which children in complete isolation do not acquire a grammar. Chomsky suggests that: “Presumably some interaction is necessary (though no one really knows, since isolation imposes extreme psychic trauma.)”⁵²⁴ But Chomsky claims that there are well-attested cases of grammar acquisition by a few children in the absence of prior relevant communal experience. Goldin-Meadow *et al* have studied children who are congenitally deaf so cannot participate in the spoken language of their population, and neither are they exposed to sign-language by the communities in which they grow up. The children develop signed languages with grammatical structure, which Goldin-Meadow describes as “the resilient properties of language”.⁵²⁵

If the conventionality of grammar resides only in the requirement that children not be isolated from other language users then “conventionality” is a requirement on grammar acquisition along with “having a sensory system with at least limited functioning and not too much brain damage, also essential for language acquisition.”⁵²⁶ But the requirement for other language users, in itself, is clearly not a requirement for grammatical conventions and so does not embody an appeal to non-psychological grammatical properties.

⁵²³ Collins (unpublished ms. c) p.25

⁵²⁴ Chomsky (2003b) p.314

⁵²⁵ Goldin-Meadow and Feldman (1977), Goldin-Meadow and Mylander (1990), and Goldin-Meadow (2005).

⁵²⁶ Chomsky (2003b) p.314

5.5 Communication

Part of the motivation for Devitt's commitment to grammatical conventions is his view that they are required to explain communication. Devitt claims that communication would be mysterious without the existence of external grammatical conventions and non-psychological grammatical properties via which speakers coordinate.⁵²⁷ Compare this with Chomsky's view of the significance of external linguistic properties in explaining communication:

Suppose we postulate that corresponding to an element 'a' of phonetic form there is an external object '*a' that 'a' selects as its *phonetic value*; thus, the element [ba] in Jones' I-language picks out some entity [*ba], 'shared' with Smith if there is a counterpart in his I-language. Communication could then be described in terms of such (partially) shared entities, which are easy enough to construct ... one could defend such a view, though no one does, because it's clear we are just spinning wheels.⁵²⁸

Chomsky thinks that if one considers the possible role such entities could perform in a putative theory of linguistic communication their definition and postulation serves no explanatory end. Given that Devitt thinks external grammatical properties are theoretically indispensable to whatever account of communication linguists propose, why does Chomsky claim that such properties are 'wheel-spinning' from the linguist's perspective?

As discussed in Chapter Three, explaining linguistic communication takes us well beyond the resources of generative grammars.⁵²⁹ Both parties agree that no one has a firm grip how linguistic communication works. But suppose we accept, as seems reasonable, that generative grammars are part of a much larger project of explaining linguistic communication.⁵³⁰ Devitt has to make a case that there will be

⁵²⁷ See Devitt (2006, 2007, 2008 p.220-5)

⁵²⁸ Chomsky (2000) p.129

⁵²⁹ See Breheny (2006) for a discussion of the psychology of communication. Although less is known about our pragmatic and other communicative abilities than about core linguistic competence, it is an area of very intense research. See Sperber and Wilson (1986) for an influential account.

⁵³⁰ It should be noted, however, that depending on how the idea of communication is to be clarified, it may be that linguistic communication is blind to important features of grammatical structure, such as structural case and the other uninterpretable features described in the Appendix.

indispensable non-psychological grammatical posits in this theory, appealing to grammatical structures beyond the psychological systems engaged in language.

The best available accounts of linguistic communication, assume the integration of an internalised grammar with an ensemble of pragmatic systems. And given that speakers and hearers employ grammatical structure in producing and comprehending language, it seems reasonable to assume with Fodor that:

[T]he internal representation of the grammar...is causally implicated in communication exchanges between speakers and hearers in so far as these exchanges are mediated by the use of the language they share.⁵³¹

Despite our relative ignorance, Devitt claims that whatever account is offered will require non-psychological grammatical properties to account for speakers and hearers mapping the same structured meanings onto the same sounds. And he thinks that conventions will be required to explain the mutual alignment of speakers and hearers with these non-psychological properties. But, even ignoring the latter claim about conventions, this former claim seems like a remarkably strong claim to defend.

Collins has responded to Devitt's claims by arguing that there is no mystery to be cleared up about communication in the absence of non-psychological grammatical properties. There is a perfectly coherent and non-mysterious account of the role of grammar in linguistic communication which makes no such appeal to non-psychological grammatical properties. Though communication may be a complicated and messy affair, we might attribute such grammatical stability as there is in communication to speakers and hearers success in mapping sufficiently similar grammatical structure onto external signals that are sufficiently similar to prompt the mutual mappings (with similarity admitting of degree).⁵³²

This stability is ultimately due to speaker-hearers having their parameters set in the same way. On this account there is no appeal to non-psychological grammatical properties or attendant grammatical conventions, and no apparent explanatory reason to conceive the external signals as possessing grammatical

⁵³¹ Fodor (1985) p.149

⁵³² See Collins (2006) pp.486-7, (2008) pp.29-31

properties. The stability in communication, at least in so far as grammar enters into the explanation, is explained by the similar mappings.⁵³³

We can understand Devitt's argument from communication as presenting a challenge to the psychological conception: how do we explain linguistic communication without non-psychological grammatical properties produced by speakers and in which hearers partake? This challenge is not a very powerful one. This is in part because of the possibility that Collins highlights according to which what is essential to such grammatical exchanges is only that the participants have similar psychological structure. What mediates the psychological structures for grammar in communication might be a set of cues and prompts to which the mind/brain is attuned. Whatever external materials mediate between speakers and hearers, we need some explanation of how speaker-hearers orientate themselves with respect to them. A psychological theory would explain such orientation by appealing to resources that everyone is committed to: the I-language.

But there is another reason why Devitt's challenge is not very powerful. Though Devitt does not embed his challenge in any particular account of linguistic communication, one such influential account that has been offered - the relevance-theoretic framework - does not seem to appeal to grammatical externalia. Its general perspective is one from which:

Communication is a process involving two information-processing devices. One device modifies the physical environment of the other. As a result the second device constructs representations similar to the representations already stored in the first device ... a language can be seen as a code which pairs semantic and phonetic representations.⁵³⁴

Within the relevance theoretic framework, linguistic exchanges involve a process of grammatical decoding.⁵³⁵ This is a matter of the grammatical system decoding

⁵³³ More generally, Collins is sceptical that communication between A and B will require some third thing C to which they both stand in a relation. I should make it clear at this point that my issue is solely about *grammatical* properties. I am making no claims whatsoever that pertain to the different issues that might arise with respect to semantic properties, such as sense and reference.

⁵³⁴ Sperber and Wilson (1996) p.461

⁵³⁵ Sperber and Wilson (1996)

linguistic symbols and assigning them structural interpretations. Relevance theorists then argue that communication involves much more than grammatical decoding of linguistic symbols, appealing to inferential processes that draw upon environmental information. What relevance theorists want to know is how a stimulus can bring about the required similarity in speaker's thoughts. But the grammatical part of their answer is that there is an internal grammatical system pairing external signals with grammatical forms, enabling the two information processing devices to engage in linguistic exchanges.

Cognitive pragmatics aims to explain how the gaps between grammatical form and the thoughts we communicate are filled by inferences that draw on environmental information. These theories make an explanatory commitment to information located in the environment upon which communicators draw as part of an inferential process. But it is unclear why they would appeal to non-psychological grammatical properties. Such theories seem to constitute just the sorts of theories that Devitt claims ought not to work. Moreover, it is unclear that any air of *mystery* resides with these theories, despite the apparent fact that they make no discernible explanatory commitment to non-psychological grammatical properties.

Devitt has not provided us with good reason to think that non-psychological grammatical properties will be theoretically indispensable to theories of linguistic communication. Therefore, Devitt has still provided no reason to think that generative grammars appeal to non-psychological grammatical properties. Linguistics targets psychological properties; in short, linguistics is a part of psychology.

6. Conclusion

In this thesis I considered the subject matter of linguistic theory and the question of whether linguistics is a part of psychology. I argued that linguistics is a part of psychology. In Chapter One, I outlined the psychological conception of generative grammar and Devitt's argument against that conception, in which he argues that generative grammars and theories of grammatical competence are distinct theories. I then considered a range of non-psychological conceptions of language as a topic of inquiry.

In Chapter Two, I considered the explanatory goals of generative grammar (§2.1). In §2.2 I argued that generative grammars must aim to be more than observationally adequate if they are to reveal linguistic structure, they must aim to be descriptively adequate: assigning the correct structural descriptions to sentences, indicating how they are understood. In §2.3 I argued further that generative grammars should aim to meet explanatory adequacy: determining the actual descriptively adequate grammar for speakers' languages on the basis of a hypothesis about grammar acquisition. In §2.4 I defended explanatory adequacy against two counter-arguments from Devitt and Katz. The conclusion of the chapter is that generative grammar has psychological goals.

In Chapter Three, I distinguished the grammarian's notion of grammatical competence from the ordinary notion of competence (§3.1). I then argued that in order to determine what the grammatical properties of a speaker's language are, grammarians need to make an explanatory commitment to a psychological distinction between a system of grammatical competence and the independent factors that enter into linguistic performances (§3.2). The commitment to this distinction is crucial to the argument in the final chapter that the psychological properties of the grammatical system are explanatorily indispensable in grammatical theory. I then distinguished competence theories from processing theories (§3.3) and argued that Katz fails to provide a cogent argument that the relevant grammatical facts outstrip the facts about grammatical competence (§3.4).

In summary, generative grammar is committed to the competence-performance distinction.

In Chapter Four I explained the central evidential role of linguistic intuitions in grammatical theory (§4.1). I then defended an orthodox model of linguistic intuitions as evidence for psychological hypotheses about grammatical competence (§4.2). So there is a defensible model available of how the evidence bears on psychological hypotheses. I demonstrated the inadequacy of Devitt's view that linguistic intuitions are speakers' theoretical judgements (§4.3) and considered an alternative observational model of linguistic intuition that might gel with a non-psychological conception of generative grammar (§4.4).

In the final chapter, I argued that grammatical hypotheses are best interpreted as psychological hypotheses concerning properties of the grammatical competence system. I distinguished this explanatory argument from supervenience arguments (§5.1). I then suggested that non-psychological conceptions of generative grammar which locate grammatical properties in the physical environment might be made consistent with grammatical results (§5.2). But I argued that the required psychological theory of competence meets the explanatory goals of generative grammar (§5.3, §5.4). The psychological conception then fulfils the requirements on a conception of generative grammar on a more parsimonious basis than non-psychological conceptions. This argument cuts against each of the non-psychological conceptions described in Chapter One. Linguistics targets psychological properties and not non-psychological properties; in short, linguistics is a part of psychology.

The more recent goal of moving *beyond explanatory adequacy* that I do not discuss in the thesis (see my Appendix), is contingent on defending the goal of explanatory adequacy as I have undertaken to do. These moves beyond explanatory adequacy more deeply embed generative grammar into the biological sciences. Consideration of moves beyond explanatory adequacy only strengthens my conclusion that linguistics is amongst the psychological sciences, offering an abstract characterisation of the structures of the mind/brain, to be integrated with studies of the brain and its evolution.

Appendix. Beyond Explanatory Adequacy

Chomsky now argues that it is an explanatory goal of linguistics, though a nascent one, to go *beyond explanatory adequacy* and to seek to explain why the human language faculty takes the form that it does.

Grammatical theories constructed within the P&P framework provide a route to resolving the tension between descriptive and explanatory adequacy. The tension resides in the fact that the languages different populations speak exhibit a variety of complex grammatical patterns but are acquired by children on the basis of a uniform endowment for language. So to meet descriptive adequacy grammatical theories must accommodate this variety and complexity in the possible languages, but to meet explanatory adequacy they must explain how a child is able to acquire any of these languages on the basis of the uniform endowment. This is the empirical problem to which P&P is the proposed solution.

P&P theories postulate a set of innate grammatical principles and a limited process of selecting amongst predetermined parameter settings on those principles during a course of experience. The scope and interaction of these parameter settings is hypothesised to account for the observed variation and complexity in the world's languages. The grammatical properties of particular languages might then be explained by being reduced to the uniform parameterised principles of UG and a residue of parameter setting. Hence, if the P&P approach is successful then the goals of descriptive and explanatory adequacy could be met. A recent development in grammatical theory is the idea that there may be further conditions on adequate grammatical theories that look beyond descriptive and explanatory adequacy, and require grammatical theories to offer some explanatory perspective on the form that UG takes.

To bring out this distinctively new explanatory goal, one might consider a Martian scientist arriving on earth. Suppose the Martian is presented with a theory of generative grammar that recursively enumerates all the hierarchical structures and other grammatical properties of the possible languages, and includes a correct hypothesis about the UG on the basis of which children acquire particular grammars. Even with such a theory at his disposal, the Martian might question *why* humans

acquire grammars with these properties. The Martian's question would be beyond the scope of an EAG that tells us why humans acquire the particular grammars they do on the basis of a theory of UG. The Martian's question presses on why the language faculty takes the initial form that it does; on why it is these UG principles and not others that characterise possible human languages. Humans develop a special system of lexical features and recursive principles that relate linguistic structures to other cognitive systems. But we could easily imagine that the grammatical system worked according to different universal principles.

Whilst there are clearly coherent issues to be pursued here, one might wonder whether the form of UG is amenable to further explanation by the grammarian, or whether his explanations stop there at the level of explanatory adequacy: characterising UG and the options it allows before handing the job over to other sciences. It is not immediately clear how the issue of *why* UG takes the form that it does could be elucidated by grammatical theory. Questions about the grammatical structures that speakers know and about the universal grammatical principles that serve as a basis for their acquisition of such knowledge, both have a very overt grammatical component. But our new question about the form UG takes seems to be a question about biology and the evolution of the language faculty. How then could grammatical theory help to resolve such questions?

HCF argue that if theories of why UG takes the form it does, are to progress then there is an outstanding task for the grammarian.⁵³⁶ The primary task of generative grammar thus far has been to provide a clear explanation of the computational requirements on the languages that we know: "The computational system must (i) construct an infinite array of internal expressions from the finite resources of the conceptual-intentional system, and (ii) provide the means to externalize and interpret them at the sensory-motor end."⁵³⁷ HCF suggest that in order to understand why UG takes the form that it does, we need grammatical theories of the computational principles that can be brought to bear on how and why they have come about.

⁵³⁶ HCF (2002), Fitch, Chomsky and Hauser (2005) and Chomsky, Hauser and Fitch (2005).

⁵³⁷ HCF (2002) p.1578

1. Questions of Perfection

HCF offer a vision of how this explanatory goal might be pursued. It centres on the notion of *perfection*, or *optimality*. Consider the following passage from HCF:

We may now ask to what extent the computational system is *optimal*, meeting natural conditions of efficient computation...To the extent that this can be established, we will be able to go beyond the (extremely difficult and still distant) accomplishment of finding the principles of the faculty of language, to an understanding of why the faculty follows these particular principles are not others. We could then understand *why* languages of a certain kind are attainable, whereas other imaginable languages are impossible to learn and sustain. Such progress would...open the door to a greatly simplified and empirically more tractable evolutionary approach to the faculty of language.⁵³⁸

The questions that HCF raise concern the efficiency or “optimality” of the grammatical system. The guiding idea, connecting such questions to questions about the evolution of the system, is that if we could get some perspective on how perfect, or imperfect, a solution the distinctive form of the language faculty is for achieving some end then this will be revealing of why it has attained such a form.

Whether the language faculty is, in some sense, a “perfect” system depends on how well or badly the grammatical system is designed for what it does, and “no matter how well or badly, to answer that question you have to add something: designed for what?”⁵³⁹ One can easily imagine criteria according to which the design of the grammatical system is far from perfect. For instance, the core recursive operations are less than perfectly adapted to the linguistic performance systems with their memory limitations and the like; they generate many unusable structures such as garden paths and multiple self-embeddings. One might also think of structural ambiguity as a kind of imperfection; if one were thinking about the grammatical system from the broader perspective of systems involved in communication. Such phenomena provide support for the view that, understood in very general terms, the grammatical system is not perfectly designed for use.

⁵³⁸ HCF (2002) p.1578 my italics

⁵³⁹ Chomsky (2002) p.104

But the thesis that HCF suggest we explore is not that the language faculty is a perfect system in such a broad sense. The thesis they suggest concerns the narrow faculty's interaction with other systems involved in the cognition of language, and whether the faculty might be optimally designed to meet certain conditions imposed by the other cognitive systems with which it interacts. HCF note that the grammatical system has to "report" to other components of the mind that make use of linguistic structures. This is often described as the grammatical system *interfacing* with other cognitive systems, in particular systems for articulating and perceiving speech and systems for thought and conceptualisation, or sometimes as the system *integrating* sound and meaning. The representations that the faculty makes available at the interface are called *PF* (roughly, a phonetic form) on the sound side, and *LF* (roughly, a logical form) on the meaning side.⁵⁴⁰ The language faculty has to interact with these two systems; otherwise the structures it generates could not serve to pair uttered and heard sounds with linguistic meanings.

HCF argue that the relevant question about perfection is whether FLN is perfectly designed for interfacing with these two independent systems; for integrating these representations of sound and linguistic form. FLN is "perfect", by this measure, if it is the most economical possible system for connecting representations that interface with the sensory-motor systems and the conceptual-intentional systems.

2. The Minimalist Program

So the research program that HCF recommend is that grammarians seek to determine to what extent grammatical design is characterised by the most economical and efficient principles, making the computation that FLN has to perform as simple and smooth as possible. Grammatical work undertaken in this

⁵⁴⁰ The existence of LF as a distinct level of representation is a matter of empirical dispute. Some recent theories dispense with this level of representation and appeal to cyclical derivations in which the faculty reports to the conceptual-intentional system at several stages (or even each stage) of the derivation with no single level of interface.

direction is called the *Minimalist Program* (MP).⁵⁴¹ MP investigates the strongest hypothesis that can be envisaged: namely, that the language faculty is a perfectly designed system for reporting to the external systems of sound and meaning. It is an open empirical question how close to the truth the strong hypothesis might be. But spurred on by some initial successes, linguists working within MP think even if the guiding hypothesis turns out to be false, much will be learnt about grammar by its falsification and, looking further ahead, the falsification of its weaker variants.

There are many different theories within MP. MP is really a methodological program guided by a very broad hypothesis, rather than a particular theory. It focuses on minimising the computational operations and constraints within grammatical theories to reflect the simplicity in principles and the efficiency in their operation within the grammatical system.

The methodological commitment of MP is to postulate only what is required of the grammatical system by *virtual conceptual necessity* (VCN). One VCN is that the grammatical system pairs sounds and linguistic forms. We could perhaps imagine that the core linguistic system did not achieve this task; hence it is only a *virtual* conceptual necessity. VCN prescribes that the grammarian attributes the grammatical system only such structure as is necessary to achieving this very general task. MP could be seen as a specialised application of Ockham's Razor; first narrowing down the focus to what is required of the language system and then working out what is absolutely required by a system that fulfils this function.

To take an important example of VCN at work within MP, consider that generative grammars once included the levels of deep structure, or D-structure, and rules mapping this level into a level of surface structure, or S-structure, and later the levels of LF and PF too. LF, for example, was postulated to meet such requirements as that the linguistic system disambiguates "Some student solved every problem" with operations of quantifier raising performed after surface structure had been reached. But none of these levels, apart from perhaps LF and PF, are required by

⁵⁴¹ See Chomsky (1995). Boeckx (2006 p.83) describes three pillars upon which he thinks MP stands: (i) the empirical emergence of economy considerations on grammatical derivations and representations, (ii) a commitment to just what is required of the grammar by virtual conceptual necessity, and (iii) an intensified search for unity and symmetry amongst grammatical operations and representations.

VCN, so minimalists have effectively junked them and tried to capture all their effects with more general operations and just the two levels of representation that are required by VCN. PF and LF seem to be required by VCN because if the grammatical system is to relate sounds and meanings it requires an interface with the perceptual-articulatory and with the conceptual-intentional systems. But neither deep structure, nor surface structure, nor operations such as the raising operation from surface structure to LF, are required by VCN. So minimalists look to account for the phenomena without appealing to such intervening grammatical levels and specialised principles.⁵⁴²

So by VCN the system integrates the conceptual-intentional information required for meaning and thought with the phonetic information required to speak and perceive language. To do so, the faculty encodes a set of grammatical features, along with recursive computational operations and the interface conditions governing the forms that can be accessed by the interfacing systems. In grammar acquisition, the grammatical features are organised into lexical items stored in the lexicon.

By VCN, the grammatical system must bring the features of one lexical item into combination with others according to recursive principles. To effect this combination, minimalist theories postulate a single combinatorial operation called *Merge* which joins two grammatical constituents together. But *Merge* is not the only operation that minimalists think the grammatical system requires. The system also requires an operation to displace lexical items from one place in a structure to another. This operation is called *Move*.⁵⁴³ Movement is required because it seems to be a fact that all natural languages involve constituents being interpreted in different places from those in which they are heard or pronounced. So the grammatical system must include some operation for moving the items from the position at

⁵⁴² However, VCN allows for the possibility that there may be many levels of representation, so long as these are all *interface levels* and not internal to FLN. See Boeckx (2006) p.75 for discussion.

⁵⁴³ According to the copy theory of movement adopted by most minimalists, this displacement leaves a copy of the displaced item/s behind in the position from which the item is moved. The copy theory of movement is more in keeping with minimalism than previous theories of empty categories, such as the trace theory. This is because it does not involve importing further items into the derivation, such as traces, that are not determined by the features of the lexical items in the numeration.

which they are interpreted, to the position at which they are heard. The generalised Movement operation allows any constituent to be moved anywhere but only subject to further general economy constraints on the computation such as *least effort* – do no more computation than is required – and *last resort* – only move something as a matter of last resort.⁵⁴⁴

To take the former least effort conditions as an example, within minimalist grammatical theory, these least effort considerations are stringently applied as economy principles on derivations. Chomsky describes the organisation of the derivation in the following way:

the language L thus generates three sets of computations: the set D of derivations, a subset D_c of convergent derivations of D, and a subset D_a of admissible derivations of D. FI [the principle of Full Interpretation] determines D_c, and the economy conditions [conditions of shortest movements] select D_a.⁵⁴⁵

The convergent derivations are those that meet the interface conditions, in particular they are fully interpretable at the interface in a sense that I will shortly elucidate. The principle of least effort in grammatical theory restricts the grammar to the most economical derivations, involving the least Movement required to derive structures that meet the interface conditions. Suppose we had two possible derivations for an expression, consisting of the same lexical items and generated by the same computational operations of Merge and Move. If they can both satisfy the conditions at the sound-meaning interface, then they are compared with regard to computational effort, and the one that requires more effort is discarded. To take a concrete case, the structures in (97) and (98) involve the same lexical items, and they get the same interpretation:

⁵⁴⁴ Whenever both Merge and Move are applicable, Merge, the fundamental structure-building operation, pre-empts Move to satisfy computational needs. Movement is a “last resort” operation: there is no “free”, in the sense of truly optional, movement. Every bit of movement within a structure must be motivated by the computation deriving a legible structure at the interface. See Boeckx (2006 pp.67-70) for discussion of last resort grammatical operations including do-support (*John left to Did John leave?*) and resumption (the improvement of a structure by insertion of a dummy pronoun, filling the gap left by illicit Movement out of an island).

⁵⁴⁵ Chomsky (1995) p.220

(97) *What did you persuade who to buy <what>?

(98) Who did you persuade <who> to buy what?

The comparison between (97) and (98) serves to rule (97) out, as (98) requires less computational effort to front <who> the shorter distance than (97) does to front <what> the longer distance from the starting structure *You persuaded who to buy what*. Equally, moving less material is always counted more economical than moving more material an equal distance. So within grammatical theory the way that least effort considerations work is by assembling possible derivations and choosing one according to economy criteria.⁵⁴⁶ Minimalist explanation involves different concepts from explanatory adequacy, for there is no reason why explanatory adequacy couldn't, in principle, be met by a system that did not meet the minimalist desiderata.

At first blush, the very existence of Movement within the grammatical system might appear to be an imperfection with respect to sound-meaning integration. One might think that a perfect system, one involving less computational operations and less computational effort, would not need the Movement operation at all. The minimalist strategy is to try and show that Movement, take together with generalised economy conditions on computation, is somehow a perfect solution to the *conflicting* demands imposed by the sound and meaning interfaces. If this could be supported, then rather than providing a clear case against minimalism, the existence of such constrained Movement would be evidence for the minimalist thesis.

Generally within MP grammarians seek to identify imperfections in the way that the computational systems pairs sounds and meanings including Movement operations but also redundant lexical features and limits on the admissible combinations of lexical items. The grammarian then tries to trace these apparent

⁵⁴⁶ These properties of *wh*-movement can be explained by the more general principle of Relativised Minimality (Rizzi 1990) which says that if you try to front an element X of type Y to a position Z, you cannot do this if there is an element W of type Y that is in between X and Z. See Boeckx (2006 pp.103-105) for discussion. See also Carston (2000) for a discussion of the rather different way that least effort considerations apply in generative grammar and relevance-theoretic pragmatics.

imperfections to the competing demands of the systems with which FLN interfaces. The aim is to show that the apparent imperfections are *only apparent*, rather reflecting perfection in the broader accomplishment of integrating representations of sound and linguistic form.⁵⁴⁷

To get a clearer idea, we can see how a simplified minimalist derivation might look. Minimalist derivations begin with a *numeration*: a selection of a set of lexical items from the lexicon. The derivation begins by Merging two lexical items, creating the most deeply embedded structural unit. The resulting Merged item is then combined by second Merge with another lexical item to create a larger phrasal unit.⁵⁴⁸ In tree-diagrammatic terms, the Merge operation extends the tree upwards and leftwards. Movement alters the structures resulting from Merge by displacing Merged elements to higher structural positions in the tree. There is then a point of “spell out” for sound, PF, and meaning, LF.⁵⁴⁹

The conditions that LF and PF representations must meet are determined by properties of cognitive systems external to the faculty that it must interface with. A derivation that is legible to the external systems is said to *converge* and one that is not legible to the external systems is said to *crash*. The possible convergent derivations that begin from the same numeration are compared in terms of the number of steps and the length of the required Movements they involve, with the less economical derivations being rejected in favour of a single maximally efficient derivation.

Lexical items are composed of features and the driving force behind minimalist derivations is that some of these features need to be checked off before the grammatical system reports at the interfaces, otherwise the derivation will crash. It is this feature checking that sanctions both Merge and Move operations.

In what sense do some of the features on the lexical items in the numeration need to be checked off? Minimalist theories assume a principle of *Full Interpretation*, according to which every grammatical feature visible at the

⁵⁴⁷ See Chomsky (2000a) for discussion.

⁵⁴⁸ For an introduction to building grammatical structure in minimalist theories see Adger (2003) pp.104-52

⁵⁴⁹ Or on some theories there are cycles with individual reports to the conceptual-intentional system rather than a single spell-out at LF.

interfaces must be interpretable by either the conceptual-intentional systems or the perceptual-articulatory systems. Focusing on the meaning side, Full Interpretation says:

Full Interpretation: The structure to which the semantic interface rules apply contains no uninterpretable features.⁵⁵⁰

So if the derivation involves uninterpretable features at any point then they must be got rid of by the point at which LF interfaces with the conceptual-intentional system. Otherwise, the derivation will crash. Once the uninterpretable features have been checked they can *delete*. This is the *Checking Requirement*.⁵⁵¹ The uninterpretable features can be checked by being Merged with a lexical item that bears a matching interpretable feature. This is called *Checking under Sisterhood*.⁵⁵²

Consider how the grammatical system then determines the structure of (99).⁵⁵³

(99) Bill sleeps

The system begins by enumerating the lexical items in (99), namely the bundles of grammatical features that constitute “Bill”, “sleep”, and an item responsible for tense which is pronounced as the *s* suffixed to *sleep*. This gives us the numeration in (100), with some of the features on *Bill* and *tense* that are operative in the derivation subscripted.

⁵⁵⁰ Adger (2003) p.85. It is noteworthy that the external systems and their requirements are not well-understood, far less well understood than the core grammar. But as Chomsky suggests (2004 p.396) “progress in understanding them goes hand-in-hand with progress in understanding the language system that interacts with them.”

⁵⁵¹ Adger (2003) p.85

⁵⁵² Adger (2003) p.85. Checking under sisterhood is a particular instance, pertaining to categorical selection features, of the more general principle of checking under c-command as occurs with relations of Agreement.

⁵⁵³ I borrow the example and its presentation from Collins (2007) and Longworth (forthcoming).

(100) [Bill {+3rd Per, +Sing Num}, sleep, tense {+Pres, -Per, -Num, -EPP}]

In broad terms, the system has to structure the elements so as to determine who, or what, did what (and perhaps, to whom). This is determined through *theta assignment*, where verbs (and other predicates) assign thematic roles to arguments. Theta assignment takes place within verbal domains, in this case through the assignment of a theta role to *Bill* by *sleep* via the combinatorial operation Merge.⁵⁵⁴ The system thus derives (101) and so is potentially in a position to make available to the conceptual-intentional systems that it was *Bill* that was the theme of the condition encoded in *sleep*.

(101) [VP Bill {+3RD Per, + Sing Num} sleep]

The next step mandated by the numeration is the merge of *tense* with (101), delivering (102) which adds information about tense, +Pres.

(102) [T tense {+Pres, -Per, -Num, -EPP} [VP Bill {+3rd Per, +Sing Num} sleep]]

However this is not sufficient for the derived structure to be legible to the conceptual-intentional system at the meaning interface. Some of the subscripted features on *tense*, those marked +, are interpretable by the conceptual-intentional systems. But others, marked -, are uninterpretable. There are grammatical person and number features involved in agreement that add no further information for the interpretive systems, as with the person and number features on *tense*. One other such feature that does not carry any information for interpretation is the -EPP feature, the “extended projection principle”, which serves to explain the movement of arguments like *Bill* into subject position. We can represent elimination (or valuation) through checking by underlining the eliminated uninterpretable feature.

In our example, the elimination of uninterpretable features involves at least two steps. First the uninterpretable person feature $-_{\text{Pers}}$ and number $-_{\text{Num}}$ on *tense* are

⁵⁵⁴ This can easily be stated in feature checking terms, where *sleep* carries an uninterpretable (-N) feature that needs checking by a (+N) item like *Bill*.

eliminated through Merge with the VP that includes *Bill*. The uninterpretable person and number features necessary for agreement are checked against the interpretable person and number features *Bill* carries giving us (103).

(103) [T *tense* {+Pres, -Per, -Num, -EPP} [VP *Bill* {+3rd Per, +Sing Num} *sleep*]]

Next, the remaining uninterpretable feature on *tense*, the -EPP feature, is eliminated through copying of *Bill*, and the movement of *Bill* into subject position. This is represented by enclosing *Bill* within <,>. *Bill* is pronounced at the start of the sentence but assigned its thematic role in the verbal domain, thereby satisfying conditions imposed by both sound and meaning. The derivation, as far as our simplified model goes, ends with (104).⁵⁵⁵

(104) [TP *Bill* {+3rd Per, +Sing Num} [T *tense* {+Pres, -3rd Per, -Sing Num, -EPP} [VP <*Bill* {+3rd Per, +Sing Num}> *sleep*]]]

Since all uninterpretable features have been eliminated, the derivation converges at the interface with the conceptual-intentional system. This is the grammatical contribution to a structure being acceptable to a speaker. Comparative unacceptability is hypothesised to be a function of, amongst other things, the number and type of uninterpretable features remaining that make the derivation crash at interface.

Feature-driven derivation is brought together with P&P in the following way. Universal principles govern the operations Merge and Move, legibility conditions at the interfaces, and the availability, and co-tenability of grammatical features (both within the lexical items and within the system as a whole). The principles are uniform with parametric variation over the feature bundles that are fixed in acquisition. Variations at the level of phrasal structure are hypothesised to be ultimately due to small variations in feature bundling which ramify through Merge and Move.

⁵⁵⁵ This simplified derivation does not include the level of structure that Adger (2003 pp.126-142) claims is added by little v.

One feature specification typically considered uninterpretable is *structural case* (nominative and accusative case). An element bearing nominative case in English can play any thematic role. The nominative *he* is the agent in (105), the benefactor in (106), the experiencer in (107), the patient or theme in (108). Moreover, the nominative can even bear no thematic role at all as with *There* in (109).

(105) He invited Mary

(106) He got the prize

(107) He saw Mary

(108) He was invited/seen by John

(109) There was a snowstorm

Accusative case is equally blind to interpreted thematic properties. Unlike nominative and accusative, other types of case, *inherent cases*, are linked to specific thematic interpretations. For example, in languages with rich case systems, an argument marked with locative case designates a location. But nominative and accusative appear to be thematically irrelevant: in this sense, they are considered uninterpretable.

Another feature specification which is considered uninterpretable is the grammatical specification of person, number, and gender (and other analogous specifications such as the class specification in some languages) which appear on *predicates* in many languages. The specification of gender, number and person on noun phrases has obvious interpretive import, but the specifications on verbs and their auxiliaries are redundant, uninterpretable features. The external systems interpreting linguistic structures will certainly want to know if the sentence is talking about one or many individuals, but the reiteration of this information on verbs does not add anything of interpretive value.

Chomsky suggests that such uninterpretable features are part of the explanation of the ubiquitous displacement property of natural languages which Moves elements from one place in a structure to a higher position in the structure.⁵⁵⁶

⁵⁵⁶ See Chomsky (1995, 2002)

Lexical items are very often articulated in one position in a structure but interpreted as if they are somewhere else. For example, the displaced subject of passive constructions is interpreted as if it is in object position, in a local relation with the verb that assigns it its thematic role. The uninterpretable features may be part of the mechanism for implementing displacement. If uninterpretable features and displacement turned out to be an optimal mechanism then minimalists could deny that they are grammatical imperfections.

Some parts of morphology have clear interpretive significance and hence are not imperfections from the minimalist perspective. Plurality on nouns, for example, is not an imperfection as it looks to have clear interpretative import for the conceptual-intentional systems. Chomsky says:

You want to distinguish singular from plural, the outside systems want to know about that. So, in fact, plurality on nouns is rather like different words: just as you have “table” and “chair”, you have singular and plural, and there are sensible reasons why plural should be an inflection and “chair” shouldn’t. Namely, everything has to be singular or plural, but not everything has to be a chair or not a chair. So there are plausible reasons why some part of morphology should be there.⁵⁵⁷

Minimalists wonder why, for example, is there singular or plural feature on elements for tense like *s* in (99)? As the singularity is already encoded on the noun *Bill*, there seems to be no interpretive reason to have it on other lexical items too. As inflections on verbs and adjectives are redundant to interpretation, they seem to be an imperfection. A grammarian interested in pursuing MP will ask why they are there. And, in this case of inflection, Chomsky thinks that “There is at least a plausible [minimalist] suggestion: they are there as perhaps an optimal method of implementing something else that must be there, namely dislocation.”⁵⁵⁸

How do minimalists seek to explain inflectional features as optimal methods for implementing dislocation? For Movement, three features are required: one to identify the target of Movement and the kind of expression that can Move to it, one that identifies the expression to be Moved and one that determines whether the

⁵⁵⁷ Chomsky (2002) p.111

⁵⁵⁸ Chomsky (2002) p.113

target has the required extra position or not. In our example (99), the thing that is Moved, *Bill*, is identified by its structural case, the target *tense* is identified by its uninterpretable features of agreement and the extra position is identified by the $-EPP$ feature. EPP marks that “here’s a position to which you can dislocate to”⁵⁵⁹. So, roughly, we have three requirements on Movement, and the uninterpretable inflectional features that are there to implement each requirement. If such a hypothesis can be sustained, then it might be that inflectional morphology, beyond the interpretable parts like singularity and plurality on nouns, is not an imperfection after all but rather an optimal way of doing something that the faculty needs to do, namely, Movement. The hypothesis that Movement is implemented by these uninterpretable features is supported where we find evidence of a robust connection between the inflectional richness of a language and the amount of Movement that it implements.⁵⁶⁰

3. Minimalism and the Evolution of Language

How could minimalist hypotheses about the grammatical system bear on the evolution of language and why UG takes the form it does? Suppose that MP is on the right track and FLN comprises bundles of lexical features, along with core recursive operations, and there are externally imposed conditions which FLN meets in a maximally efficient way. On this minimalist picture, much of the complexity in grammatical structure is traceable not to complexity in the recursive operations of FLN but rather to the lexicon and, in particular, conditions imposed on the derivation by the peripheral components of FLB. This has some interesting consequences for theories of why FLN takes the form that it does.

It makes apparent the possibility that the underlying computational operations of FLN may have a different evolutionary explanation to the complex peripheral systems of FLB. On the basis of this distinction between FLN and FLB, HCF offer the speculative hypothesis that whilst “most, if not all, of FLB, is based

⁵⁵⁹ Chomsky (2002) pp.115-6

⁵⁶⁰ See Belletti and Rizzi (2002) p.33 and references cited there for discussion.

on mechanisms shared with nonhuman animals” in contrast “FLN – the computational mechanism of recursion – is recently evolved and unique to our species.”⁵⁶¹ HCF suggest that the mechanisms of FLB are instanced, though in a far less sophisticated form, elsewhere in the animal kingdom. By contrast, they argue that grammatical recursion is species specific.

But HCF claim that if the core grammatical system involves only this simple recursive device, then this may have the “interesting effect of nullifying the argument from design, and thus rendering the status of FLN as an adaptation open to question.”⁵⁶² Arguments from design, for a biological system being an adaptation, are premised on the fact that the only known biological mechanism capable of generating complex structure is natural selection. But according to minimalist theories of the grammatical system, FLN itself does not exhibit complex structure; it includes only the recursive operations which act on the properties of lexical items to derive structures legible at the interfaces. If FLN itself does not have complex structure then there is no such immediate argument for it being the result of natural selection, where external pressures make successive selections amongst complex and differentiated structures. Of course, if this line of reasoning is correct, then conversely the failure of MP might suggest complexity to the core grammatical system and might lend some support to the hypothesis that the system is the result of progressive selections.

HCF are sceptical of the assumption that the evolution of FLN can be explained as a matter of natural selection. It is their view that “minor modifications to this foundational system alone seem inadequate to generate the fundamental difference – discrete infinity – between language and all known forms of animal communication.”⁵⁶³ As they see it, there is:

[L]ittle reason to believe either that FLN can be anatomised into many independent but interacting traits, each with its own independent evolutionary history, or that each of these

⁵⁶¹ HCF (2002) p.1573

⁵⁶² *ibid.*

⁵⁶³ HCF (2002) p.1574

traits could have been shaped by natural selection, given their tenuous connection to communicative efficacy.⁵⁶⁴

And here we can see how grammatical theories about the structure of the core grammatical system may begin to extend beyond explanatory adequacy and get some purchase in arguments about why the language faculty takes the form it does.⁵⁶⁵

In evaluating how grammatical theory could bear on issues of language evolution, Chomsky is concerned to highlight the range of factors involved in the design of any biological system; in particular, environmental influence, genetic factors and more general physical constraints (constraints that are language or even organism independent).⁵⁶⁶ Chomsky thinks that the first of these factors often been assumed as the explanation of the faculty's form, but that the last has often been ignored:

The theory of *evolution*, not necessarily the workings of natural selection; and surely not these alone, since, trivially, they operate within a physical "channel", the effects of which are to be discovered, not stipulated.⁵⁶⁷

The workings of natural selection will be most to the fore where an evolved system shows evidence of tinkering in achieving intricate structure. So the hypothesis that the core grammatical system does not display the characteristic of intricate structure could be directly relevant to the prospects of such an explanation. It is a plus point of MP that the falsification of its guiding hypothesis would actually suggest something about the evolution of language. To support the broader claim that grammatical theory seeks integration with a theory of language evolution, it need not be the case that minimalism is true: MP would be a way of pursuing this integration even if minimalism turns out to be false. This seems to me to be the perspective of HCF. They say:

⁵⁶⁴ *ibid.*

⁵⁶⁵ For an opposing view, see Pinker and Jackendoff (2005).

⁵⁶⁶ Chomsky (2005)

⁵⁶⁷ Chomsky (2002) p.47

It has been a useful research guide to formulate the Strong Minimalist Thesis (SMT), which holds that language is a “best possible” solution to the problem of linking SM [sensory-motor interface] and CI [conceptual-intentional interface]. In these terms, the task of MP is to clarify the notions that enter into the SMT and to determine how closely the thesis can be approached. Insofar as this can be achieved, the traditional concerns of identifying the specific features of FL are advanced, and the study of its evolution rendered more feasible.⁵⁶⁸

Chomsky has speculated that the special characteristics of FLN that minimalist theories suggest might be *spandrels*: by-products of pre-existing physical constraints rather than products of natural selection. He is keen to point out that psychological capacities, like all biological functions, operate within a physical channel. Psychological capacities, such as recursion, are implemented in neural tissue in the brain. Hence, they are constrained by biophysical principles, physical principles that govern all biological development. Chomsky speculates that such physical constraints on design might turn out to be important, maybe more important than adaptive function, amongst the plurality of factors involved in shaping the grammatical system. Of course, it is an open empirical question whether this speculation will prove prescient.

The point is not that FLN, forming part of FLB as it does, is not adaptive. In allowing us to produce an endless variety of meaningful expressions and to communicate an endless variety of thoughts, FLN is clearly of adaptive value. The question is whether the particular structure of FLN is *the product of adaptations*, acted on by natural selection. As Collins characterises Chomsky’s view, he is suggesting is that:

We should be ‘pluralists’ about evolution, which amounts to no more than taking into account all of the factors that make for organic design, not just environmental contingencies and mutation: physical constraints, developmental processes, and adaptive pressures...such a pluralism is not optional, for no course of natural selection operates in a vacuum; it must

⁵⁶⁸ Chomsky, Hauser and Fitch (2005) p.2

proceed along a physico-chemical channel and must factor in developmental constraints that are only just being revealed.⁵⁶⁹

And Chomsky is keen to point out that no one hopes to explain cell division, the presence of the Fibonacci sequence in plant structure, or other simple mathematical sequences in nature, in terms of their being selected for.

What these attempts to move beyond explanatory adequacy suggest is that the psychological conception of the goals of grammatical theory, which I defended in sections (§2.2 and §2.3), offers the prospect of unification, ultimately with the core biological sciences. This is a very attractive feature of that conception.

⁵⁶⁹ Collins (unpublished ms. b) p.4

Bibliography

- Adger, D. (2003) *Core Syntax: A Minimalist Approach*, Oxford University Press, Oxford.
- Antony, L. M. and Hornstein, N. (eds.) (2003) *Chomsky and his Critics*, Blackwell Publishing, Oxford.
- Baker, M. (2001) *The Atoms of Language: The Mind's Hidden Rules of Grammar*, Basic Books, New York.
- Barber, A. (ed.) (2000) *The Epistemology of Language*, Oxford University Press, Oxford.
- Beattie, J. (1788) *The Theory of Language*, London Press, London.
- Belletti, A. and Rizzi, L. (2002) 'Editor's Introduction' to N. Chomsky *On Nature and Language*, Cambridge University Press, Cambridge.
- Bernecker, S. and Dretske, F. (2000) *Knowledge: Readings in Contemporary Epistemology*, Oxford University Press, Oxford.
- Bloomfield, L. (1936) 'Language or Ideas', *Language*, Volume 12, pp.89 – 95
- Boeckx, C. (2006) *Linguistic Minimalism: Origins, Concepts, Methods and Aims*, Oxford University Press, Oxford.
- Bohannon, J.N. and Stanowicz, L. (1988) 'The Issue of Negative Evidence: Adult Responses to Children's Language Errors', *Developmental Psychology*, Volume 24, pp.684 - 689
- Bowerman, M. F. (1988) 'The "no negative evidence" problem: How do children avoid constructing overly general grammar?', in *Explaining Language Universals* (ed.) J. A. Hawkins, Basil Blackwell, Oxford.
- Breheny, R. (2006) 'Communication and Folk psychology', *Mind & Language*, Volume 21, Number 1, pp.74-107
- Burge, T. (1979) 'Individualism and the Mental', *Midwest Studies in Philosophy*, Volume IV, University of Minnesota Press, Minneapolis pp.73-121
- _____ (1986) 'Individualism and Psychology' *The Philosophical Review*, Volume 95, pp.3-45

- _____ (1988) 'Individualism and Self-Knowledge', reprinted in S. Bernecker and F. Dretske (ed.) 2000
- Burton-Roberts, N., Carr, P. And Docherty, G. (ed.) (2003) *Phonological Knowledge* Oxford University Press, Oxford.
- Carruthers, P. & Boucher, J. (eds.) (1998) *Language and Thought: Interdisciplinary themes*, Cambridge University Press, Cambridge.
- Carston, R. (2000) 'The Relationship between Generative Grammar and (Relevance-Theoretic) Pragmatics', *Language and Communication*, Volume 20, pp.87-103
- Cauman, L., Levi, I., Parsons, C. & Schwartz, R. (ed.) (1983) *How Many Questions? Essays in Honor of Sidney Morgenbesser*, Indiannapolis, Hackett Publishing Co.
- Chomsky, N. (1955) *The Logical Structure of Linguistic Theory*, reprinted by Plenum Press, New York, 1975.
- _____ (1957) *Syntactic Structures*, Mouton, The Hague.
- _____ (1965) *Aspects of the Theory of Syntax*, M.I.T. Press, Cambridge MA.
- _____ (1972) *Language and Mind*, Harcourt Brace Janovich, New York.
- _____ (1975) *Reflections on Language*, Pantheon Books, New York.
- _____ (1975a) 'Quine's Empirical Assumptions', in D. Davidson and J. Hintikka (ed.) 1975
- _____ (1976) 'On the Biological Basis of Language Capacities' in R. Reiber (ed.) 1976
- _____ (1979) *Language and Responsibility*, Pantheon Books, New York.
- _____ (1980) *Rules and Representations*, reprinted by Columbia University Press, New York, 2004.
- _____ (1981) *Lectures on Government and Binding*, Foris, Dordrecht.
- _____ (1986) *Knowledge of Language: Its Nature, Origin and Use*, Praeger, New York.
- _____ (1987) 'Replies to George and Brody', *Mind & Language*, Volume 2, Issue 2, pp.178-97.
- _____ (1988) *Language and Problems of Knowledge*, The M.I.T. Press, Cambridge MA

- _____ (1990) 'Language and Problems of Knowledge' A.P. Martinich (ed.) 1990 pp.581-98.
- _____ (1995) *The Minimalist Program*, The M.I.T Press, Cambridge MA.
- _____ (1995a) 'Language and Nature', *Mind*, Volume 104, pp.1-61
- _____ (2000) *New Horizons in the Study of Language and Mind*, Cambridge University Press, Cambridge.
- _____ (2000a) 'Minimalist Inquiries: The Framework' in R. Martin, D. Michaels and J. Uriagareka (eds.) 2000.
- _____ (2001) 'Derivation by Phase' in M. Kenstowicz, (ed.), M.I.T. Press, Cambridge MA.
- _____ (2002) *On Nature and Language*, Cambridge University Press, Cambridge.
- _____ (2003) 'Reply to Egan' in L. Antony and N. Hornstein (eds.) 2003
- _____ (2003a) 'Reply to Ludlow' in L. Antony and N. Hornstein (eds.) 2003
- _____ (2003b) 'Reply to Millikan' in L. Antony and N. Hornstein (eds.) 2003
- _____ (2003c) 'Reply to Pietroski' in L. Antony and N. Hornstein (eds.) 2003
- _____ (2003d) 'Reply to Rey' in L. Antony and N. Hornstein (eds.) 2003
- _____ (2004) 'Language and Mind: Current Thoughts on Ancient Problems', in L. Jenkins (ed.), Elsevier, London, pp.379-405.
- _____ (2005) 'Three Factors in Language Design', *Linguistic Inquiry*, Volume 36, Number 1, pp.1-22
- Chomsky, N. and Halle, M. (1968) *The Sound Pattern of English*, Harper and Row, New York.
- Chomsky, N., Hauser, M. D, and Fitch, T. W. (2005) *Appendix. The Minimalist Program*, available at:
<http://www.wjh.harvard.edu/~mnkylab/publications/recent/EvolAppendix.pdf>
- Chomsky, N. and Katz, J. (1974) 'What the Linguist is Talking About', *The Journal of Philosophy*, Volume 71, Number 12, pp.347-367
- Chouinard, M. and Clark, E. (2003) 'Adult reformulations of child errors as negative evidence', *Journal of Child Language*, Volume 30, pp.637-69
- Collins, J. (2003) 'Cowie on the Poverty of Stimulus', *Synthese*, Volume 136, pp.159-90

- _____ (2004) 'Faculty Disputes', *Mind & Language*, Volume 19, Number 5, pp.503-33.
- _____ (2006) 'Between a Rock and a Hard Place: A Dialogue of on the Methodology and Philosophy of Linguistics.' *Croatian Journal of Philosophy*, Volume 6, pp.471-505.
- _____ (2007) 'Review of Michael Devitt's *Ignorance of Language*', *Mind*, Volume 116, pp.416-423.
- _____ (2008) 'Explanation as the Mark of *Psychological Reality*' paper presented to the London Institute of Philosophy
- _____ (2008a) 'Knowledge of Language Redux', *Croatian Journal of Philosophy*, Volume VIII, Number 22.
- _____ (2008b) A Note on Conventions and Unvoiced Syntax, *Croatian Journal of Philosophy*, Volume VIII, Number 23.
- _____ (unpublished ms.) 'Critical Study of Ray Jackendoff's *Foundations of Language: Brain, Meaning, Grammar, Evolution*'
- _____ (unpublished ms. a) 'Devitt's Dodging'
- _____ (unpublished ms. b) 'Review of Wolfram Hinzen's *Mind Design and Minimal Syntax*'
- _____ (unpublished ms. c) 'The Perils of Content'
- Cowie, F. (1999) *What's Within: Nativism Reconsidered*, Oxford University Press, Oxford.
- Crain, S. (1991) 'Language Acquisition in the absence of Experience', *The Behavioural and Brain Sciences*, Volume 4, pp.597-650.
- Crain, S. and McKee, C. (1985) 'Acquisition of Structural Constraints on Anaphora', Paper for the 10th Annual Boston Conference on Language Development.
- Crain, S. and Nakayama, M. (1987) 'Structure Dependence in Grammar Formation', *Language*, Volume 63, pp.522-543.
- Crain, S. and Pietroski, P. (2001) 'Nature, Nurture and Universal Grammar', *Linguistics and Philosophy*, Volume 24, pp.139-86
- _____ (2002) 'Why Language Acquisition is a Snap', *The Linguistic Review*, Volume 19, pp.163-83

- Crain, S., and Thornton, R. (1998) *Investigations in Universal Grammar: A Guide to Experiments in the Acquisition of Syntax and Semantics*, The M.I.T. Press, Cambridge MA.
- Curtiss, S. (1988) 'Abnormal Language Acquisition and the Modularity of Language', in F. Newmeyer (ed.) 1988, pp. 96-116
- Davidson, D. and Harman, G. (eds.) (1972) *Semantics of Natural Language*, Humanities Press, New York.
- Davidson, D. and Hintikka, J. (eds.) (1975) *Words and Objections*, D. Reidel Publishing, Boston.
- Demetras, M.J. and Post, K.N. (1985) 'Negative Feedback in Mother-Child Dialogues', Paper presented at the Society for Research in Child Development, Toronto.
- Devitt, M. (2006) *Ignorance of Language*, Oxford University Press, Oxford.
- _____ (2006a) 'Defending *Ignorance of Language*: Responses to the Dubrovnik Papers', *Croatian Journal of Philosophy*, Volume VI, pp.573-609
- _____ (2006b) 'Intuitions in Linguistics', *The British Journal for the Philosophy of Science*, Volume 57, pp.481-513.
- _____ (2008) 'Explanation and Reality in Linguistics', *Croatian Journal of Philosophy*, Volume VIII, pp.203-231.
- _____ (2008a) 'Response to Collins' *Note on Conventions and Unvoiced Syntax*', *Croatian Journal of Philosophy*, Volume VIII, pp.249-55.
- _____ (unpublished ms.) 'Dodging the Argument on the Subject Matter of Grammar: A Response to Collins and Slezak, available online at: <http://web.gc.cuny.edu/philosophy/people/devitt/LingExchange-heading.htm>
- _____ (unpublished ms. a) 'Response to Collins' Review of *Ignorance of Language*'
- _____ (unpublished ms. b) 'Response to Pietroski's *Think of the Children*'
- Devitt, M. and Sterelny, K. (1989) 'Linguistics: What's Wrong with the "Right View"', *Philosophical Perspectives*, 3, pp.497-531
- _____ (1999) *Language and Reality: An Introduction to the Philosophy of Language*, 2nd Edition, Blackwell Publishers, Oxford.

- Dummett, M. (1978) 'What Do I know When I know a Language?' reprinted in Dummett 1993.
- _____ (1989) 'Language and Communication', in A. George (ed.) 1989.
- _____ (1993) *The Seas of Language*, Oxford University Press, Oxford.
- Egan, F. (1995) 'Computation and Content', *The Philosophical Review*, Volume 104, pp.181-203
- _____ (2003) 'Naturalistic Inquiry: Where does Mental Representation Fit in?', in L. Antony and N. Hornstein (eds.) 2003
- Emmorey, K. and H. Lane (eds.) (2000) *The Signs of Language Revisited: An Anthology in Honour of Ursula Bellugi and Edward Klima*, Lawrence Erlbaum, New Jersey.
- Epstein, S. D. and Seely, T. D. (eds.) (2002) *Derivation and Explanation in the Minimalist Program*, Blackwell Publishing, Oxford.
- Fiengo, R. – (2003) 'Linguistic Intuitions', *Philosophical Forum*, Volume 34, pp.253-266.
- Fitch, Hauser and Chomsky (2005) – 'The Evolution of the Language Faculty: Clarifications and Implications', *Cognition*, 97, pp.179-210
- Fodor, J. A. (1974) 'Special Sciences and the disunity of Science as a working hypothesis', *Synthese*, Volume 28, pp.77-115
- _____ (1980) 'Methodological Solipsism Considered as a Research Strategy in Cognitive Psychology', *The Behavioural and Brain Sciences*, Volume 3, pp.63-73.
- _____ (1983) *The Modularity of Mind*, The M.I.T. Press, Cambridge MA.
- _____ (1985) 'Some notes on what linguistics is about', reprinted in Katz (ed.) 1985.
- _____ (2001) 'Doing Without *What's Within*: Fiona Cowie's Critique of Nativism', *Mind*, Volume 110, pp.99-148
- Fodor, J. A., Bever, T. & Garrett (1974) *The Psychology of Language*, McGraw-Hill Books, New York.
- Foraker, S., Regier, T., Khetarpal, N., Perfors, A., & Tenenbaum, J. (forthcoming). 'Indirect evidence and the poverty of the stimulus: The case of anaphoric "one."', *Cognitive Science*

- Frank, A. (1997) 'Quantum Honey Bees', *Discover*, pp.80-7
- Franks, B. (1995) 'On Explanation in the Cognitive Sciences: Competence, Idealisation and the Failure of the Classical Cascade', *The British Journal of the Philosophy of Science*, 46, pp.475-502
- Frege, G. (1956) 'The Thought: A Logical Inquiry', *Mind*, Volume LXV, pp.289-311
- _____ (1967) *The Basic Laws of Arithmetic*, University of California Press, Berkeley.
- Garrett, M. (1990) 'Sentence Processing', in D. N. Osherson and E. E. Smith (eds.) 1990
- Geirsson, H. And Losonsky, M. (eds.) (1996) *Readings in Language and Mind*, Blackwell Publishers, Oxford.
- George, A. (ed.) (1989) *Reflections on Chomsky*, Blackwell Publishers, Oxford
- _____ (1996) 'Katz Astray', *Mind & Language*, Volume 11, Number 3, pp.295-305
- Gibson, E. and Wexler, K. (1994) 'Triggers', *Linguistic Inquiry*, 25, pp.355-407
- Goldin-Meadow, S. (2005) *The Resilience of Language: What gesture creation in language can tell us about how all children learn language*, Psychology Press, New York.
- Goldin-Meadow, S. and Feldman, H. (1977) 'The Development of Language-Like Communication without a Language Model', *Science*, 197, pp.401-3
- Goldin-Meadow, S. and Mylander, C. (1990) 'The Role of Parental Input in the Development of a Morphological System', *Journal of Child Language*, Volume 17, pp.527-63
- Guasti, M. (2002) *Language Acquisition: The Growth of Grammar*, The M.I.T. Press, Cambridge MA.
- Haegeman, L. (1992) *Introduction to Government and Binding Theory*, 2nd Edition, Blackwell Publishers, Oxford.
- Hahn, E. and Schlipp, P. A. (eds.) (1986) *The Philosophy of W. V. Quine*, Open Court, La Salle.
- Harman, G. (ed.) (1974) *On Noam Chomsky: Critical Essays*, Anchor Press/Doubleday, New York.

- Harris, J. and Lindsay, G. (2003) 'Vowel Patterns in Mind and Sound' in N. Burton-Roberts, P. Carr and G. Docherty (ed.) 2003
- Harris, Z. (1985) *Distributional Structure*, originally printed in *Word*, 1954, Volume 10, pp.775-793, reprinted in Katz (ed.) 1985.
- Hauser, M. D., Chomsky, N. and T. W. Fitch (2002) 'The Faculty of Language: What is it, Who has it, and How did it Evolve?' *Science*, 298, pp.1569-79
- Heath, S. B. (1982) *Ways with Words: Language, Life, and Work in Communities and Classrooms*, Cambridge University Press, Cambridge.
- Hempel, C. G. (1965) *Aspects of Scientific Explanation*, Free Press, New York.
- Higginbotham, J. (1983) 'Is Grammar Psychological?' in L. Cauman, I. Levi, C. Parsons and R. Schwartz (ed.) 1983, pp. 170-9.
- Hinzen, W. (2006) *Mind Design and Minimal Syntax*, Oxford University Press, Oxford.
- Hirsh-Pasek, Treiman ad Schneiderman (1984) 'Brown and Hanlon Revisited: Mothers' sensitivity to ungrammatical forms', *Journal of Child Language*, Volume 11, pp.81-88
- Hjelmslev, L. (1985) 'Structural Analysis of Language' reprinted in Katz (ed.) (1985)
- Horn, L. (1988) 'Pragmatic Theory' in F. Newmeyer (ed.) 1988 pp.113-45.
- Husserl, E. (1970) *Logical investigations*, Humanities Press, New York.
- Jackendoff, R. (2002) *Foundations of Language: Brain, Meaning, Grammar, Evolution*, Oxford University Press, Oxford.
- Jacobson, P. (1999) 'Towards a variable-free semantics', *Linguistics and Philosophy*, Volume 22, pp.117-184
- Jenkins, L. (ed.) (2004) *Variation and Universal in Bilingualism*, Elsevier, London.
- Kalderon, M. (2007) 'Color Pluralism', *Philosophical Review*, Volume 116, pp.563-601
- Katz, J. (1964) 'Mentalism in Linguistics', *Language*, Volume 40, pp.124-137
- _____ (1981) *Language and Other Abstract Objects*, Rowman and Littlefield, New Jersey.
- _____ (1985) 'An Outline of Platonist Grammar' in Katz (ed.) 1985.
- _____ (ed.) (1985) *The Philosophy of Linguistics*, Oxford University Press, Oxford.

- _____ (1996) 'The Unfinished Chomskyan Revolution', *Mind & Language*, Volume 11, pp.270-94.
- Kayne, R. (2000) *Parameters and Universals*, Oxford University Press, Oxford.
- _____ (2002) 'Pronouns and their Antecedents', in S.D. Epstein and T.D. Seely (eds.) 2002.
- Keats, John – 'Ode to a Nightingale', available at:
http://www.everypoet.com/Archive/Poetry/John_Keats/keats_ode_to_a_nightingale.htm
- Kenstowicz, M. (ed.) (2001) *Ken Hale: A Life in Language*, The M.I.T. Press, Cambridge MA.
- Lappin, S. and Schieber, S. (2007) 'Machine Learning as a Guide to Universal Grammar', *Journal of Linguistics*, 43, pp.393-427.
- Laurence, S. and Margolis, E. (2001) 'The Poverty of Stimulus Argument', *The British Journal for the Philosophy of Science*, 52, pp.217-276.
- Lees, Robert B. (1957) 'Review of *Syntactic Structures*', *Language*, Volume 33, Number 3, pp.375-408.
- Lepore, E. and Pylyshyn, Z. (ed.) (1999) *What is Cognitive Science?*, Blackwell Publishers, Oxford.
- Lepore, E. and Smith, B. C. (ed.) (2006) *The Oxford Handbook to the Philosophy of Language*, Oxford University Press, Oxford.
- Lewis, D. (1983) 'Language and Languages', in Martinich (ed.) 1990.
- _____ (1992) 'Meaning without Use: A reply to Hawthorne', *Australasian Journal of Philosophy*, Volume 70, Issue 1, pp.106-110.
- Longworth, G. (2007) 'Conflicting Grammatical Appearances', *Croatian Journal of Philosophy: Philosophy of Linguistics*, Volume VII, Number 21.
- _____ (forthcoming) 'Ignorance of Linguistics: A Note on Michael Devitt's *Ignorance of Language*', *Croatian Journal of Philosophy*.
- _____ (unpublished ms.) 'Locating Grammatical Properties'
- Ludlow, P. (2003) 'Referential Semantics for I-languages?', in L. Antony and Hornstein, N (eds.) 2003 pp.140-161.
- _____ (forthcoming) *Philosophical Issues in Generative Linguistics*
- Marcus, G. (1993) 'Negative Evidence in Language Acquisition', *Cognition*, 46, pp.53-85

- Martin, R., Michaels, D., and Uriagereka, J. (eds.) (2000) *Step by Step – Essays in Minimalist Syntax in Honor of Howard Lasnik*, The M.I.T. Press, Cambridge MA.
- Martinich, A. P. (ed.) (1990) *The Philosophy of Language*, 4th edition, Oxford University Press, Oxford.
- Matthews, R. (2006) ‘Knowledge of Language and Linguistic Competence’, *Philosophical Issues, 16: Philosophy of Language*, pp.200-220.
- McDowell, J. (1984) ‘Wittgenstein on Following a Rule’ in A. Miller and C. Wright (ed.)
- _____ (1998) *Meaning, Knowledge and Reality*, Harvard University Press, Cambridge MA.
- _____ (1998a) ‘Anti-realist Semantics and the Epistemology of Understanding’, in McDowell 1998.
- McGilvray, J. (ed.) (2005) – *The Cambridge Companion to Chomsky*
- MacWhinney, B. (ed.) (1987) *Mechanisms of Language Acquisition*, Erlbaum, New Jersey.
- MacWhinney, B. and Bates, E. (1989) *The Cross-linguistic Study of Sentence Processing*, Cambridge University Press, New York.
- Miller, G. (1962) ‘Some Psychological Studies of Grammar’, *American Psychologist*, Volume 17, pp.748-62
- Miller, A. and Wright, C. (ed.) (2002) *Rule-Following and Meaning*, Acumen Press.
- Morgan, J. and Travis, L. (1989) ‘Limits on Negative Information on Language Learning’, *Journal of Child Language*, Volume 16, pp.531-52
- Neeleman, A. and van de Koot, H. (2008) ‘Theoretical Validity and Psychological Reality of the Grammatical Code’, available at:
http://www.phon.ucl.ac.uk/home/hans/grammatical_code.pdf
- Newmeyer, F. (ed.), (1988) *Linguistics: the Cambridge Survey*, Volumes I and II, Cambridge University Press, Cambridge.
- Newport, E., Gleitman, H., & Gleitman, L. (1977) ‘Mother I’d rather do it myself: some effects and non-effects of maternal speech-style’ in Snow and Ferguson (eds.)
- Nowak, M.A. *et al* (2001) ‘Evolution of Universal Grammar’, *Science*, 291, pp.114-118

- Osherson, D. N., and Smith, E. E. (eds.) (1990) *An Invitation to Cognitive Science*, Vol. 1: Language, The M.I.T. Press, Cambridge MA.
- Patterson, S. (1998) 'Competence and the Classical Cascade: A Reply to Franks', *The British Journal for the Philosophy of Science*, Volume 49, pp.625-636
- Peirce, C.S. (1958) *Collected Papers of Charles Sanders Peirce*, Volume 4, Harvard University Press, Cambridge MA.
- Penner (1987) 'Parental Responses to Grammatical and Ungrammatical Child Utterances', *Child Development*, 58, pp.376 -384.
- Pettito, L. (2000) 'On the Biological Foundations of Human Language' in K. Emmorey and H. Lane (eds.) 2000 pp.447-471.
- _____ (2005) 'How the Brain Begets Language', in McGilvray (ed.) 2005, pp. 84-101
- Pinker, S. (1987) 'The Bootstrapping problem in language acquisition', in B. MacWhinney (ed.) 1987
- _____ (1990) 'Language Acquisition' in D.N. Osherson and H. Lasnik (ed.) 1990, pp.107-33.
- _____ (1999) *Words and Rules*, Weidenfeld & Nicolson, London.
- Pinker, S. and Jackendoff, R. (2005) 'The Faculty of Language: What's special about it?' *Cognition*, 95, pp.201-236
- Prince, E. (1988) 'Discourse Analysis: a part of the study of linguistic competence' in F. Newmeyer (ed.) 1988, pp.164-182
- Pullum, G. K. and Scholz, B. (2002) 'Empirical Assessment of Stimulus Poverty Arguments', *The Linguistic Review*, Volume 19, pp.9-50.
- Putnam, H. (1975) 'Philosophy of Logic' in *Mathematics, Matter and Method*, Cambridge University Press, Cambridge.
- Quine, W.V. (1953) 'Two Dogmas of Empiricism', in *From a Logical Point of View*, Harvard University Press, Cambridge MA.
- _____ (1960) *Word and Object*, The M.I.T. Press, Cambridge MA.
- _____ (1969) 'Epistemology Naturalised', in *Ontological Relativity and Other Essays*, Columbia University Press, New York.
- _____ (1972) 'Methodological Reflections on Current Linguistic Theory', in Davidson and Harman (eds.)
- _____ (1975) 'Reply to Chomsky', in Davidson and Hintikka (ed.)

- _____ (1985) 'The Problem of Meaning in Linguistics' reprinted in Katz (ed.) 1985
- _____ (1986) 'Reply to Harman', in Hahn, E and Schlipp, P. A. (eds.)
- _____ (1987) 'The Indeterminacy of Translation Again', *The Journal of Philosophy*, Volume 84, Number 1, pp.5-9.
- _____ (1990) *The Pursuit of Truth*, Harvard University Press, Cambridge MA.
- Radford, A. (1988) *Transformational Grammar: A First Course*, Cambridge University Press, Cambridge.
- Rey, G. (2000) 'Intentional Content and a Chomskian Linguistics' in A. Barber (ed.) 2000.
- _____ (2003) "Chomsky, Intentionality, and a CRTT" in L. Antony and N. Hornstein (ed.) 2003
- _____ (2006) 'Conventions, Intuitions and Linguistic Inexistents: A reply to Devitt', *Croatian Journal of Philosophy*, Volume VI.
- Rieber, R. W. (ed.) (1976) *The Neuropsychology of Language*, New York, Plenum Press, New York.
- Rizzi, L. (1990) *Relativized Minimality*, The M.I.T. Press, Cambridge MA.
- Ryle, G. (1949) *The Concept of Mind* Hutchinson, London, (Reprinted 2000, Penguin Books, London).
- Safir, K. (2004) *The Syntax of (In)dependence*, Linguistic Inquiry Monograph, 44, The M.I.T. Press, Cambridge MA
- Samuels, R. (2002) 'Nativism in Cognitive Science', *Mind & Language*, Volume 17, pp.233-265
- Sapir, E. (1921) *Language: An Introduction to the Study of Speech*, Harcourt Brace, New York.
- Schiffer, S. (2006) 'Two Perspectives on Knowledge of Language', *Philosophical Issues*, Philosophy of Language, Volume 16, pp.275-87
- Searle, J. (1974) 'Chomsky's revolution in linguistics' in G. Harman (ed.) 1974
- _____ (1992) *The Rediscovery of the Mind*, The M.I.T. Press, Cambridge MA.
- Segal, G. (1989) 'Seeing What is Not There', *The Philosophical Review*, Volume XCVIII, Number 2, pp.189-214
- _____ (2000) *A Slim Book about Narrow Content*, The M.I.T. Press, Cambridge MA

- _____ (forthcoming) 'Poverty of Stimulus Arguments Concerning Language and Folk Psychology'
- Skinner, B. F. (1938) *The Behavior of Organisms*, Appleton-Century-Crofts, New York.
- _____ (1957) *Verbal Behavior*, Appleton-Century-Crofts, New York.
- Slezak, P. (unpublished ms.) 'Linguistic Explanation and "Psychological Reality"'
Online at:
<http://hist-phil.arts.unsw.edu.au/staff/staff.php?first=Peter&last=Slezak>
- _____ (unpublished ms. a) 'Devitt's *Ignorance of Language*'
- Slobin, D. (1972) 'Children and Language: They learn the same way all around the world', *Psychology Today*, Volume 6, pp.71-82.
- Smith, B. C. (2006) 'What I know when I know a Language' in E. Lepore and B. C. Smith (ed.)
- _____ (2006a) 'Why we still need knowledge of language', *Croatian Journal of Philosophy*, Volume VI.
- _____ (2008) 'What Remains of Our Knowledge of Language? Reply to Collins', *Croatian Journal of Philosophy*, Volume VIII, Number 22.
- _____ (unpublished ms.) 'The Nature of "Linguistic Reality": A Response to Devitt's *Dodging*'
- Smith, N. (2000) *Foreward* to N. Chomsky 2000
- Snow, C. and Ferguson, C. (eds.) (1977) *Talking to Children: Language Input and Acquisition*, Cambridge University Press, Cambridge.
- Sperber, D. and Wilson, D. (1986) *Relevance: Communication and Cognition*, Blackwell, Oxford.
- _____ (1996) 'Précis of *Relevance: Communication and Cognition*', in H. Geirsson and M. Losonsky (eds.)
- Soames, S. (1984) 'Linguistics and Psychology', *Linguistics and Philosophy*, Volume 7, pp.155-79
- _____ (1985) 'Semantics and Psychology' in Katz (ed.) 1985
- Stanton, R. (ed.) (2006) *Contemporary Debates in Cognitive Science*, Blackwell, Oxford.
- Stanley, J. (2000) 'Context and Logical Form', *Linguistics and Philosophy*, Volume 23, pp.391-44

- Stich, S. (1985) 'Grammar, Psychology and Indeterminacy' in Katz (ed.) 1985
- Strawson, P. (1979) 'Universals', *Midwest Studies in Philosophy*, Volume 4, Issue 1, pp.3-10
- Stromswold, K. (1999) 'Cognitive and Neural Aspects of Language Acquisition', in E. Lepore and Z. Pylyshyn (ed.), pp.356-400.
- Travis, C. (2002) 'Frege's Target', in *Logic, Thought and Language*, Royal Institute of Philosophy Supplements, Number 51.
- Van der Lely, H.K.J, Rosen, S., & Adlard, A. (2004) 'Grammatical Language Impairment and the specificity of cognitive domains: Relations between auditory and language abilities', *Cognition*, Volume 94, pp.167-183
- Van der Lely, H.K.J, Rosen, S., & McClelland, A. (1998) 'Evidence for a grammar-specific deficit in children'. *Current Biology*, Volume 8, pp.1253-1258
- Van der Lely, H. K. J and Stollwerck, I. (1996) 'A grammatical specific language impairment in children: An Autosomal Dominant Inheritance?', *Brain and Language*, Volume 52, pp.484 - 504
- Van Fraassen, B. C. (1980) *The Scientific Image*, Clarendon Press, Oxford.
- Varley, R. (1998) 'Aphasic Language, aphasic thought: propositional thought in an apropositional aphasic' in Carruthers, P. & Boucher, J. (eds.)
- Veletti-Glass, A., Gazzaniga, M., & Premack, D. (1973) 'Artificial Language Training in Global Aphasics' *Neuropsychologica*, Volume 11, pp.95-104
- Wallman, J. (1992) *Aping Language*, Cambridge University Press, Cambridge.
- Wiggins, D. (1997) 'Languages as Social Objects', *Philosophy*, Volume 72, pp.499-524.
- Williamson, T. (2006) 'Can Cognition be Factorised in Internal and External Components?' in Stainton (ed.) 2006
- Wittgenstein, L. (2001) *Philosophical Investigations*, 3rd edition, originally published 1953, Basil Blackwell, Oxford.
- Wright, C. (1986) *Realism, Meaning and Truth*, Basil Blackwell, Oxford.
- Yang, C. (2002) *Knowledge and Learning in Natural Language*, Oxford University Press, Oxford.
- _____ (2004) 'Universal Grammar, Statistics or both?', *Trends in Cognitive Science*, Volume 8, Number 10, pp.451-456