

Vagueness, Inconsistency, and Semantic Competence

BPhil Thesis

Will Davies
Balliol College, University of Oxford

Word Count: 29,998

ABSTRACT: I defend Cognitive Inconsistentism, a novel explanation of vagueness qua susceptibility to sorites paradoxes. The view is contrasted with the Inconsistency Theories of Michael Dummett and Matti Eklund. I argue that these latter views are committed to unattractive theories of linguistic understanding. I propose an alternative Chomskyan, cognitive theory of linguistic understanding. I employ this cognitive framework to argue that speakers mentally represent a similarity constraint on conceptual classification, which is causally and explanatorily implicated in language acquisition and use. This cognitive similarity constraint causally explains our linguistic intuitions with respect to sorites paradoxes.

Acknowledgements:

Thanks to an audience at the Ockham Society in Oxford for comments on early material from the thesis. Thanks also to Kim Plunkett for discussion of scientific issues. I express a deep gratitude to Martin Davies, who provided comments of significant range and depth on later versions of this work. I owe a quite different debt to Gabriel, for her love, patience, and teacakes. I reserve my greatest thanks for Tim Williamson, for his generous and incisive supervision, and for his continued support and encouragement.

Contents

0. Introduction	5
1. Vagueness and Inconsistency	6
1.1 The Nature of Vagueness	6
1.2 The Problems of Vagueness	9
1.3 Inconsistency Theories	13
1.3.1 Incoherentism	14
1.3.2 Eklund's Inconsistency Theory	21
1.3.3 Cognitive Inconsistentism	26
2. Linguistic Understanding: Abilities and Knowledge	28
2.1 Introduction	28
2.2 The Major Divide in Theories of Understanding	28
2.3 The Role of Intention in Language Acquisition	33
2.4 The Separation of Competence and Performance	37
2.5 Language Acquisition and Poverty of Stimulus	40
2.6 Problems with Chomsky	45
2.6.1 The Impropriety of 'Knowledge'	45
2.6.2 The Objectivity of Language	48
2.7 Eklund's Inferential Role Semantics	57
3. Vagueness, Similarity, and Semantic Competence	59
3.1 Introduction	59
3.2 The Inadequacy of Classic Arguments for Tolerance	59
3.3 Respecting Centre-Focus: Exemplars and Prototypes	65
3.3.1 Prototypes and Compositionality	80
3.4 The Similarity Constraint and Vagueness	82
3.4.1 Similarity and Poverty of the Stimulus	87
3.4.2 Dissociations between PRO and SIM	98
3.4.3 Explaining Borderline Cases	100
3.5 The Need for a Cognitive Similarity Constraint	101
4. Conclusions	105
Bibliography	106

0. Introduction

Why is so much of our language vague? What are the origins of vagueness? Why are we so susceptible to vagueness-related sorites paradoxes? In this thesis I shall propose a novel explanation of the source of vagueness *qua* sorites susceptibility: *Cognitive Inconsistentism*. I shall argue that our understanding of language, specifically our semantic competence, partly consists in mentally representing an inconsistent principle called the ‘similarity constraint’. The similarity constraint explains important aspects of language acquisition and use, and key features of conceptual classification. Importantly, the constraint also causally explains our linguistic intuitions with respect to the major premise of sorites paradoxes.

Cognitive Inconsistentism is inspired by the semantic Incoherency theory of Michael Dummett, and by Matti Eklund’s recent attempts to revive Dummett’s approach. I shall argue that both Dummett and Eklund put forward unattractive theories of linguistic understanding, and that neither view satisfactorily explains the pervasiveness and robustness of our susceptibility to the sorites. I shall argue that we have overwhelming independent reason to favour a Chomskyan, cognitive theory of linguistic understanding. Against the background of such a cognitive conception of our semantic competence, I shall demonstrate that the similarity constraint is causally and explanatorily implicated in our linguistic and conceptual processes. The resulting explanation of vagueness in natural language is both empirically viable and philosophically satisfying.

1. Vagueness and Inconsistency

1.1 The Nature of Vagueness

What is vagueness? Which features are characteristic of vague expressions? These particular questions have received a significant amount of attention in recent years, spawning a cottage industry in characterising vagueness.¹ In this section, I do not propose to *define* 'vague', but merely to provide a working characterisation of the phenomenon of vagueness. I shall outline two key features in this regard: 1. the existence of borderline cases, and 2. sorites susceptibility. While both features ultimately must be explained by a theory of vagueness, I propose initially to focus on the notion of sorites susceptibility.

It has perhaps been typical in philosophy to characterise vagueness as the existence of borderline cases for natural language expressions.² A working definition of a borderline case *b*, with respect to a predicate *F*, is that *b* lies somewhere between clear cases of *F* and clear cases of *not-F*, although it is unclear and undecidable, for a normal, competent language user *S*, whether *b* is in, or out, of the extension of *F*, no matter how far *S* extends her evidence, within her capacity, with respect to *b*.³ We could say that borderline cases offer a *prima facie* basis for doubting the semantic principle of bivalence, that every sentence is either true or false, and with that its logical counterpart, the validity of excluded middle, that for all well-formed

¹ (Greenough, 2003), (Smith N. J., 2005), (Weatherson, Forthcoming).

² (Peirce, 1902), (Black, 1939), (Fine, 1975), (Williamson, 1994), (Tye, 1995).

³ (Greenough, 2003, p. 245), describes a similar *indeterminist* conception of vagueness *qua* borderline cases. I wish to avoid complications here, however, concerning *definiteness* and *determinacy* operators. For discussion see (Wright, 1987, pp. 229-41) and (Williamson, 1994, pp. 149-64).

formulas w , $w \vee \neg w$ is true.⁴ As an example, take Jones, whose hair has noticeably thinned, from his temple to his crown, but who retains a layer of coverage, fairly evenly distributed across that area. It is neither clear that Jones is bald, nor that he is not bald. Given my capacities, no matter how far I extend my evidence with respect to Jones's hair distribution, the meaning of 'bald', expert opinion, and so on, it will remain unclear and undecidable for me whether Jones is bald or not-bald. The case gives us at least a *prima facie* basis for doubting that 'Jones is bald' is true or false in this case.⁵

The use of borderline cases in characterising vagueness must be exercised with some caution, however. It appears that we can artificially generate expressions that are not vague, yet admit of borderline cases. Mark Sainsbury has provided the example of *child**, which is defined as being true of people less than 16, and false of people greater than or equal to 18 years of age.⁶ *Child** intuitively has borderline cases of people who are 17, lying between the extension and anti-extension of the concept, for whom it is unclear and undecidable whether the term applies to them or not. But *child** seems quite different from standard examples of vague terms, such as 'bald', 'red', and so on. Perhaps our theory of vagueness will eventually deliver an account of the difference between *child** and vague terms *proper*.⁷ This may well involve refinement or clarification of the notion of a borderline case. Pretheoretically,

⁴ Future contingents, for example, also provide *prima facie* reason to doubt bivalence; but they are not borderline cases. Future contingents do not lie in the zone between clear cases of the extension and anti-extension of a concept.

⁵ The foregoing characterisation of borderline cases adverts to *predicates*; but predicates are not the only vague expressions in natural language. Transitive and intransitive verbs, ('loves', 'moves'), adverbs, ('quickly'), prepositions, ('on'), and perhaps more contentiously, proper names, ('Gordon Brown'), all intuitively admit of vagueness. Kit Fine goes so far as to say, (Fine, 1975, p. 120): 'any type of expression capable of meaning is also capable of being vague'.

⁶ (Sainsbury, 1991, p. 173).

⁷ Patrick Greenough, (Greenough, 2003, pp. 245-48), suggests that the difference is that *child** does not suffer from *higher order* vagueness. I remain neutral here on the issue of higher order vagueness.

however, we should simply be aware that borderline cases, as described here, provide no more than a working characterisation of the phenomenon of vagueness.

Our second candidate characterisation of vagueness was *sorites susceptibility*. A term F is *sorites susceptible* just in case a sorites paradox can be constructed for F , with premises that seem intuitively true to ordinary, competent speakers, but a conclusion that is clearly false. Consider the following example with respect to ‘rich’:

1. A person with £1,000,000,000.00 is rich.
2. For all n , if a person with £ n is rich, a person with £ $[n - 0.01]$ is rich.
3. Therefore, a person with £0.01 (1p) is rich.

Both premises 1 and 2 seem intuitively acceptable. The *minor* premise 1 makes an indubitable assertion about a clear case of richness, in any ordinary context. The *major* premise 2 seems equally acceptable. Denying that 2 true, at least assuming classical logic, would be logically equivalent to claiming the existence of sharp boundaries, the existence of some n for which a person with £ n is rich, but a person with £ $[n - 0.01]$ is *not* rich. It clearly forms no part of our ordinary understanding that ‘rich’ has sharp boundaries. Combined with some uncontroversial logic,⁸ however, premises 1 and 2 entail the abhorrent conclusion that a person with 1p is rich. The term ‘rich’ is sorites susceptible: the term ‘rich’ is vague.

As with borderline cases, we should be wary of overstating the significance of sorites susceptibility. Just as borderline cases proved insufficient for characterising the phenomenon of vagueness, sorites susceptibility seems not to be necessary. This

⁸ Specifically, the rules of $\forall$ -elimination and *modus ponens*.

is precisely what one would expect, given the fickleness of linguistic intuitions. It is certainly conceivable, for instance, that there could exist a community of fully competent speakers who use vague terms in almost exactly the same way as us, except that they have an inherent distrust of any sentence involving universal quantification. Perhaps this is a community of radical sceptics about induction, who are taught from an early age to expect generalisations to fail somewhere along the line. This community lacks the intuition that the major premise of the sorites is true, but would still consider vague all the terms that we do.⁹ Again, sorites susceptibility is not intended to provide more than a working characterisation of vagueness. I propose to focus on the explanandum of sorites susceptibility in what follows. As we shall see in Section 3.4.3, however, my proposed explanation of sorites susceptibility extends naturally to borderline cases.

1.2 The Problems of Vagueness

I have opted to focus on vagueness as characterised by *susceptibility to sorites paradoxes*. Philosophers tend to see paradoxes as in need of resolution. With this tendency in mind, a *default response* to the paradox has been that the major premise must be either *false*, or at least *not true*. The default response will block the absurd conclusion of the sorites, thus *solving* the paradox; but this move consequently raises a host of troubling questions.¹⁰ The *semantic/logical* question asks, if the major premise is not true, then must its classical negation, i.e. sharp boundaries, be true?

⁹ Perhaps such a community would be susceptible to the ‘forced march’ sorites, on which speakers are presented with a chain of pair-wise similar cases, as opposed to a universally quantified claim. But we could redescribe the thought experiment so that the community is also averse to concentrating on more than one thing at a time, and so on for other candidate forms of the sorites.

¹⁰ I here follow (Graff, 2000, p. 50).

And if not, then what revision to classical logic or semantics is required? The *epistemological* question asks, if the major premise is not true, why are we unable to identify which instance(s) is (are) not true? Finally, the *psychological* question asks, if the major premise is indeed not true, why are we so inclined to accept it as true? The question could be rephrased, *why* are vague terms sorites susceptible? These three questions typify the *problems of vagueness*.

Variations on the default response have proven by far the most popular in theorising about vagueness. The tendency has been to attempt to *solve* the sorites, and to provide some combined semantic, logical, or epistemic account as to a) where the major premise goes wrong, and b) why we cannot identify or know where it goes wrong. To the extent that such accounts have tried to explain why the sorites possesses allure for us, and thus to tackle the psychological question, this often seems like something of an afterthought, an attempt to extend the semantic or epistemic theory to the psychological domain.¹¹

To take an illustrative example, consider Supervaluationism.

Supervaluationism is a semantic theory about the interpretation of vague expressions. Working within the model theoretic paradigm, Supervaluationism holds that vague predicates are not truth-evaluable relative to single assignments of semantic values, but are rather ‘super-valuated’ over a range of *precisifications*, or arbitrary assignments of sharply bounded extensions. Supervaluationist Kit Fine has two suggestions as to why we are inclined to accept the major premise of the sorites

¹¹ Crispin Wright (Wright, 1987, pp. 211, fn. 3) does not disguise his discontent: ‘Writers have been content to devise more or less *ad hoc* semantics in whose terms the major premises [of the sorites] fail of strict truth without doing anything to disclose *why* their plausibility is specious, still less confront the intuitively powerful arguments on their behalf’.

as true.¹² The first is that the falsifying instance of the major premise is *arbitrary*. This arbitrariness is an artefact of the formal semantic model employed by Supervaluationism. The arbitrariness pertains to *precisification*, which is a technical notion, relating to the provision of a model theoretical specification space. But why should such a technical semantic notion explain the *psychological* fact that we are so inclined to accept the major premise? No connection with actual linguistic cognition or understanding is forthcoming.

Fine's second suggestion is that the falsifying instance of the major premise *shifts* for each complete specification, or precisification.¹³ This shiftiness might serve as an explanation of the *epistemological* issue, as to why we cannot *identify* which instances are not true. Adapting this response to suit the psychological question, it might be suggested that we judge the major premise to be true *because* we cannot identify any particular instance of the generalisation as *not* true. This purported explanation will not suffice, however, as we clearly do dissent from universal generalisations for which no particular falsifying instance is identifiable. Let us define a predicate 'primedog' to mean 'the *n*th dog born since midnight (GMT) January 1st 1901, where *n* is any prime number'. Now 'every dog is a primedog' is clearly false, but no person could identify any *instance* of a dog that is not a primedog. Such examples abound; there is no easy move between answers to the epistemological and psychological questions.

Tim Williamson's Epistemicism, however, might be taken to provide just such a move between the epistemological and psychological types of explanation. In

¹² (Fine, 1975, p. 139).

¹³ That is to say, no particular instance is *super*-false, but several instances may be neither true nor false.

response to the epistemological question, Williamson claims that we cannot *identify* the falsifying instance of the major premise, because we cannot *know* where such a boundary could lie. We cannot *know* where the boundary of *F* lies, because we cannot *reliably judge* where it lies. Even if we were to judge – correctly, let us suppose – that the boundary occurred between *n* and *n+1*, the meaning of *F* could have been ever so slightly different, without our knowing it, so as to determine the boundary to lie between *n-1* and *n*.¹⁴ Williamson claims that this epistemic account can provide an answer to the psychological question. Given the unknowability of such boundary cases, it is thus *unimaginable* for us that the major premise of the sorites could be falsified. This feature purportedly explains the allure of that premise.

While offering an improvement on Fine's Supervaluationist proposal, Williamson's suggestion will not suffice as an answer to the psychological question. In short, this is because Williamson's explanation of the allure of the sorites does not have a psychological *basis*. If one tracks Williamson's explanation back far enough, we arrive not at *psychological* facts, but normative epistemic principles. To recall, Williamson's suggestion is that the allure the major premise is explained by our being unable to imagine the falsity that premise. Our being unable to imagine its falsity in turn is explained by our lack of *knowledge* of the falsifying instance. But our lack of knowledge is explained not with reference to any aspects of individual cognition, but rather with reference to the nearby possibility in which the boundary of *F* is unknowably different. This is a perfectly reasonable account of the *epistemic* aspects of vagueness, of course. But it does not explain what it is about our *actual psychology* that makes us susceptible to sorites paradoxes. Unless some further

¹⁴ For more on Williamson's 'margin for error principles', see (Williamson, 1994, pp. 226-30).

explanation is given as to how Williamson's epistemic principles are implicated in our actual cognitive processes, his account will not satisfy our desire for a psychological explanation of vagueness.¹⁵

So far, we have seen that two popular variations on the default response, Supervaluationism and Epistemicism, lack satisfactory explanations as to the *source* of sorites susceptibility, and thus the *origin* of vagueness.¹⁶ The force of the psychological question therefore remains: *why* do ordinary, competent speakers judge the major premise of the sorites to be true?¹⁷ In the next section, I shall introduce three *Inconsistency Theories*, which take the psychological question as their primary focus.

1.3 Inconsistency Theories

Why are vague terms sorites susceptible? What explains the allure of the major premise of the sorites? Inconsistency Theories, at a very high level of description, claim that we are 'sometimes led to accept untrue and even jointly inconsistent propositions by virtue of our semantic competence'.¹⁸ Given an absolutely minimal gloss, 'semantic competence' with respect to a language *L* is whatever state, or condition, is attained and possessed by people who *understand the expressions in L*.¹⁹

An Inconsistency Theory for vagueness, then, claims two things: first, that the

¹⁵ For illuminating recent discussion on the relationship between 'constitutive' and 'normative' theories of content, see (Greenberg, 2005).

¹⁶ For further criticisms of this nature, see (Graff, 2000, pp. 51-54).

¹⁷ One might question whether an explanation of the source of vagueness, which is often assumed to be a semantic phenomenon, should advert to actual human psychology *at all*. Haven't we learned from Frege that semantics and psychology ought to be kept separate? I here ask the reader to withhold judgement until Chapter 2, when I shall defend a cognitive, psychologistic account of linguistic understanding.

¹⁸ (Eklund, 2002, p. 251).

¹⁹ People may be more or less competent, of course, in respect of how *many* of *L*'s expressions are understood, but also in how *much* about these expressions is understood.

meanings of vague expressions are in some way inconsistent; and second, that it is by virtue of understanding these meanings that we are inclined to accept the major premise of the sorites.

As it stands, of course, this serves as no more than a schematic, promissory response to the psychological question. Inconsistency Theories owe us an account as to A) what understanding vague expressions, i.e. semantic competence, *consists in*, B) which aspects of the meanings of vague expressions are inconsistent, and C) why possessing such semantic competence should explain our assent to the major premise of the sorites. Briefly put, my overall aim in this thesis is to critically evaluate two candidate accounts of these issues, and to propose a novel third account. Let us now discuss the first candidate account: *Incoherentism*.

1.3.1 Incoherentism

The earliest form of Inconsistency Theory, ‘Incoherentism’, was due to Michael Dummett.²⁰ In summary form, Incoherentism provides the following responses to issues A-C:

A1. Linguistic understanding consists in possessing a set of practical abilities, which are governed by a set of semantic rules.

B1. The *semantic rules* governing our use of vague expressions are inconsistent.

²⁰ See also (Wheeler, 1979), and (Unger, 1979). My characterisation of Dummett’s views will draw on Crispin Wright’s (Wright, 1976, 1987) critical discussions of Incoherentism.

C1. Complying with these rules, and thus being competent with respect to vague expressions, *requires* one to assent to the major premise of the sorites.

Let us first consider point A1. On this picture, a semantic rule for a predicate dictates the conditions for its correct use, under which that predicate may justifiably be attributed to an object. Semantic rules are part of the *meaning* of expressions, specified in what Dummett would call a *theory of sense*. A theory of sense is to be distinguished from a theory of *reference* for the language.²¹ A theory of reference contains axioms, which assign references – or ‘semantic values’, such as objects and functions, or properties – to all primitive terms in the language. Together with a finite number of recursive rules for the combination of primitive terms, the semantic value – a *truth* value – for any number of complex sentences can be determined within the theory. A theory of sense, in contrast, represents that which is *known* by competent speakers of the language.²² Dummett claims, however, that such knowledge is properly conceived as ‘practical knowledge, knowledge *how* to speak the language’.²³ A theory of sense, insofar as it is given in terms of rules and propositions, is merely a ‘theoretical model of a practical ability’,²⁴ which proceeds ‘by correlating specific practical abilities of the speaker to certain propositions of the theory’.²⁵ Connecting this back to A1, our abilities to use language are ‘governed’ by

²¹ (Dummett, 1993, p. 40).

²² This is captured in Dummett’s iconic phrase, ‘a theory of meaning [sense] is a theory of understanding’. (Dummett, 1981, p. 92). I shall use the term ‘sense’ in a fairly minimal way, simply to distinguish what is understood by speakers from the semantic values assigned to expressions on interpretation. I do not commit to any explicitly Fregean or truth-conditional interpretation of the notion.

²³ (Dummett, 1993, p. 36), italics added.

²⁴ (*Ibid*, p. 15).

²⁵ (*Ibid*, p. 40).

semantic rules, in the sense that a theoretical representation of those abilities would consist in a specification of those rules.²⁶

In Chapter 2, I shall critically discuss Dummett's 'abilities view' of linguistic understanding at some length. The foregoing will remain somewhat unclear until then. My present purpose, however, is to illuminate the connection between this conception of understanding and vagueness *qua* sorites susceptibility. The crucial further claim made by Incoherentism is that the semantic rules governing our use of vague expressions include *tolerance principles*. The notion of tolerance requires explication. In the words of Crispin Wright,

[A term] *F* is *tolerant* with respect to ϕ if there is some positive degree of change in respect of ϕ insufficient ever to affect the justice with which *F* applies to a particular case.²⁷

This suggests that tolerance ought to be analysed with respect to an ordered triple, $\langle F, P, \epsilon \rangle$, in which we have a term *F*, a property *P* that forms the basis for the application of *F*,²⁸ and some arbitrarily small change in value ϵ with respect to *P*. Taking *P*(*x*) to denote the value of *P* for *x*, let us state the following tolerance principle:

$$(TP) \forall x, \forall y, \text{ if } |P(x) - P(y)| \leq \epsilon, \text{ then } Fx \leftrightarrow Fy$$

²⁶ Compare: in diving to catch a fish, a Kingfisher has to calculate the adjustment in angle required to account for the refraction of light at the water's surface. There may be a trigonometric rule 'governing' the Kingfisher's diving practices. But we would only say that the Kingfisher 'knows' that rule, in the sense that it possesses a diving ability which could be *theoretical represented* by that rule.

²⁷ (Wright, 1976, pp. 156-7).

²⁸ The property *P* need not be that which is signified by *F*, although it might be. In the case of *F*: is rich, and *P*: the amount of money a person has, it is eminently plausible that these properties are distinct. In contrast, perhaps *F*: is phenomenally red both denotes, and is applied on the basis of, the property of *phenomenal redness*.

Instantiating TP for F : is rich, P : the amount of money a person has, and $\varepsilon = 1p$, TP dictates that any two people differing in respect of at most 1p in monetary wealth are either both rich, or both not rich. The crucial Incoherentist claim, then, is that this instance of TP is a semantic rule for the vague expression 'rich'. Relating this to point C1, it should be clear that complying with this instance of TP would require one to assent to the major premise of the sorites for 'rich'. Vagueness *qua* sorites susceptibility thus arises as a result of speakers exercising a normal understanding of the semantic rules governing vague expressions.

We might wonder, however, what reason there is to think that TP *is* a semantic rule for our language. Or to phrase the question in Dummett's terms, what practical ability is possessed by competent users of vague terms, of which TP is supposed to be a propositional model? Dummett's argument for the status of TP as a semantic rule focuses on the somewhat limited class of so-called 'observational terms'. I shall take issue with this narrow focus in Chapter 3. Leaving that issue aside for now, for Dummett the ability in question is to *apply such terms simply on the basis of looking*. Our understanding of a colour term such as 'red' would on this view consist in the ability, say, to produce 'red' when presented with certain shades of colour, and not with others. This ability would necessarily be constrained by our ordinary powers of perception. Ordinary perception, however, is subject to indiscriminable differences in colour: there might be two shades a and b , which are in fact of slightly different hue, but not *discriminably* so. If 'red' is to be applied purely on the basis of ordinary perceptual observation, Dummett notes, its use 'is to

be governed by the principle that, if I cannot discern any difference between the colour of *a* and the colour of *b*, and I have characterised *a* as red, then I am bound to accept a characterisation of *b* as red'.²⁹ If our understanding of 'red' consists in the ability to apply the term simply on the basis of looking, then the rules governing its use must include TP.

If Dummett is correct, and TP is indeed a semantic rule, this implies that the rules governing our use of language are *inconsistent*, as per B1, issuing in contradictory instructions for our use of the term. Given a sorites series from red to orange, our tolerance principle for *F*: is red would dictate, if applied repeatedly, that the last shade, which is in fact *orange*, is red. The result generalises: if anything is red, then anything that can be connected to it by a sorites series is also dictated to be red. Given the continuity of the colour spectrum, of course, it follows that *every* shade is red. But this plainly contradicts other semantic rules, for example, that there exists something that is *not* red. Accepting TP as a semantic rule seems a quick road to semantic chaos.

A deeply troubling line of thought follows from such chaos. In response to the apparent semantic incoherence of 'red', Dummett says that 'phenomenal qualities', the properties designated by observational colour terms, are 'beyond rescue'.³⁰ Dummett here seems to be following Frege in concluding that such expressions are *semantically defective*, in the sense that they stand for no genuine properties.³¹ The reasoning is as follows. 'Red' and 'blue' are *antonyms*, and so must designate distinct

²⁹ (Dummett, 1975, p. 114). Dummett naturally relates this back to Nelson Goodman's discussion of the non-transitivity of phenomenal indiscriminability, (Goodman, 1951).

³⁰ (Dummett, 1975, p. 116).

³¹ Frege says that vague expressions lack *Bedeutung*, e.g. (Frege, 1979, pp. 122, 155, 179).

properties, if they are to designate any properties at all. The properties assigned as semantic values for ‘red’ and ‘blue’ must satisfy all semantic rules for these terms, including TP. But instantiating TP for F : is red, and G : is blue, by the foregoing sorites reasoning we could yield the conclusions that every shade is red *and* that every shade is blue. ‘Red’ and ‘blue’ thus apply to every shade, and designate no distinct properties. Ergo, ‘red’ and ‘blue’ designate no properties at all. This constitutes *nihilism* about the semantic content, or reference, of vague expressions.³²

The envisaged route to nihilism might strike us as too quick, however.

Consider an instantiation of TP for F : is red, and P : observed colour:

(TP1) $\forall x, \forall y$, if the difference in observed colour between x and y is less than or equal to ϵ , then x is red $\leftrightarrow y$ is red.

TP1 would imply that the *reference* of ‘is red’ includes every colour. Making the disquotational assumption, that ‘ x is red’ is true iff x is red, the right hand side of TP1 would imply that ‘ x is red’ is true iff ‘ y is red’ is true. Under sorites circumstances, then, TP1 could be used to deduce that ‘is red’ is *true* of every colour, and thus that the reference of the expression is all encompassing. And it is here that the threat of nihilism looms large. Call this reading of TP ‘*reference tolerance*’.

³² For a clear overview of the relationship between inconsistency and nihilism, see (Williamson, 1994, pp. 165-84).

We should recall, however, that Incoherentism was made as a claim about the *meaning* of vague expressions, at the level of a theory of *sense*, not reference.³³ TP was conceived as a semantic rule governing the use of vague expressions, and part of the *knowledge* possessed by competent speakers about the meanings of vague expressions. TP was *not* billed as an axiom in a theory of reference. We might therefore reformulate TP, letting *J* serve as an operator for ‘justifiably predicates’,

$$(TP') \forall x, \forall y, \text{ if } |P(x) - P(y)| \leq \epsilon, \text{ then } J(Fx) \leftrightarrow J(Fy)$$

TP' states that, where *x* and *y* differ by at most ϵ with respect to *P*, *F* is justifiably predicated of *x* iff *F* is justifiably predicated of *y*. Here, ‘justifiably predicated of *x*’ is not to be understood as meaning ‘is true of *x*’. TP' is conceived as part of speakers’ *understanding*, governing their *use* of *F*, not the term’s reference.³⁴ The operative notion of ‘justification’ is therefore that which is possessed by speakers in virtue of their understanding *F*, and knowing how to use that term in accord with the semantic rules governing the term. Call this reading of TP' ‘*epistemic tolerance*’.³⁵

Instantiating on TP', rather than TP, we could still derive the conclusion that ‘red’ is justifiably predicated of every shade. But given that TP' represents speakers’ knowledge of the meaning, or understanding, of ‘red’, this possible derivation need not trouble us so much. For one, speakers do not always track the logical entailments of their epistemic commitments. This is especially true when those commitments are

³³ (Eklund, 2005, p. 34) notes that a so-called ‘representational theory of meaning’, on which ‘the meaning of an expression is what it stands for in the world’, might rule out the possibility of such a distinction from first principles. I shall assume that such ‘representational’ theories are to be avoided.

³⁴ The theory of understanding will thus not be *truth-conditional*. I cannot expand on this consequence here.

³⁵ For a somewhat different notion of epistemic tolerance, see (Greenough, 2003, pp. 256-64).

not *explicitly* known to the speaker, but rather only manifest in their practice, as Dummett believes, or as I shall suggest later, in their cognitive processes. Such understanding could thus still eventuate in a non-trivial, long-run consistent pattern of use for vague expressions. Furthermore, unlike being *true* of every shade, the relation of being justifiably predicated is *defeasible*, subject to competing or conflicting constraints on our use of language. This defeasible conception of tolerance, and its role in our understanding and cognition of vague terms and concepts will be a major theme in Chapter 3.

By insisting that tolerance be conceived at the level of sense, Incoherentism can avoid a descent to nihilism about the reference of vague expressions. By opening up a gap between semantic rules, construed as objects of speakers' knowledge, and the determination of reference, or assignment of semantic values, it is possible to resist the move from inconsistency in semantic rules, to inconsistency in reference. What is now needed, however, is a model of reference determination that can accommodate and 'neutralise' the inherent inconsistency at the level of sense, or understanding. As I shall now proceed to discuss, Matti Eklund has recently provided just such an account.

1.3.2 Eklund's Inconsistency Theory

In a series of recent papers, Matti Eklund has done much to revive the fortunes of Inconsistency Theories. In summary form, his responses to issues A-C are as follows:

A2. Linguistic understanding consists in accepting, or being disposed to accept as true or valid a set of meaning constitutive rules, principles, and inferences.

B2. The *meaning constitutive principles* (MCPs) for vague expressions are inconsistent.

C2. Our semantic competence with respect to vague expressions *disposes* us to assent to the major premise of the sorites.

Taking point A2, Eklund advocates inferential role semantics (IRS), claiming that semantic competence consists in *accepting*, or being *disposed to accept*, a set of MCPs for expressions of the language.³⁶ As with Dummett's views on linguistic understanding, we shall have to wait until Chapter 2 to discuss the problems with Eklund's IRS. Again, however, our primary current concern is to understand how Eklund's view of understanding connects with the issue of vagueness *qua* sorites susceptibility. Eklund claims that TP – what we called *reference tolerance* – is an MCP. By A2, this means that our understanding, or semantic competence, with respect to vague expressions partly *consists* in accepting, or being disposed to accept TP as true. Taking *D* as an operator for 'disposed to accept',

$$(TP'') D[\forall x, \forall y, \text{ if } |P(x) - P(y)| \leq \epsilon, \text{ then } Fx \leftrightarrow Fy]$$

As with TP', TP'' is a claim at the level of *sense*, as to what speakers' understanding of vague expressions consists in. Namely, in this instance, that understanding

³⁶ (Eklund, 2005, p. 41).

consists in dispositions to accept TP, an inconsistent principle. There is no theoretical commitment to the truth of TP as such, and thus no commitment to its absurd logical consequences. Our disposition to accept TP might not extend to dispositions to accept all of its logical consequences. Moreover, just as with the notion of ‘justifiably predicated’, linguistic dispositions are defeasible, and can be *overridden*. Eklund suggests that this might well occur with our disposition to accept TP, when we come to realise that the principle is inconsistent.³⁷

Eklund’s theory marks an advance from Incoherentism in that it combines a commitment to inconsistent semantic rules, or MCPs, at the level of sense, with consistent reference determination. Eklund claims that the semantic value for a vague expression is that which comes ‘*closest* to satisfying the meaning-constitutive principles associated with the expression’.³⁸ Semantic values for vague expressions are assigned in accord with three *acceptability constraints*. First, the extension and non-extension of the term should be (at least possibly) *non-empty*.³⁹ This respects the MCP for ‘red’, for instance, that *some* differences in shade are large enough to discount something as being red. Second, all assignments are *classical*. This respects the MCP that everything is either red or not red. Third, there are many acceptable assignments. Eklund employs two more general considerations:

Respect: ‘Other things equal, the semantic values of expressions are such as to make speaker judgements maximally correct.’

³⁷ (*Ibid*, p. 50).

³⁸ (*Ibid*, pp. 42-3).

³⁹ The qualification is intended to catch vague predicates that have no actual extension. I shall ignore the complication of expressions that *necessarily* have no extension.

Approximation: ‘When rendering speaker judgements “maximally correct”, different judgements are differently weighted.’⁴⁰

Eklund does not expand on the notions of ‘maximal correctness’, ‘weighting’ of speaker judgements, or the ‘approximation’ that is undertaken with respect to these weightings. These notions undoubtedly require further explication. However, Eklund would not be alone in considering the particular algorithm for determination of semantic values to be at best unclear to theorists of meaning, and perhaps at worst simply unknowable.⁴¹ The principles of respect and approximation, as they stand, do seem to be assumptions that are ‘nearly universally made’ in semantic theorising.⁴² With these constraints in play, Eklund suggests that the reference for a vague expression, such as ‘red’, would in fact be a genuine property, with a classical extension, and that the reference tolerance principle, TP, with respect to ‘red’ would therefore be strictly *false*. Nihilism about the semantic contents of vague expressions is thus averted.

Eklund’s anti-nihilist proposal has reinstated Inconsistency Theory as a viable explanation of vagueness. If the meanings of vague terms are indeed inherently inconsistent, and if semantic competence requires that we in some sense grasp these meanings, this would provide a powerful account as to why fully competent

⁴⁰ (Eklund, 2002, pp. 263-4).

⁴¹ On a related theme, see (Williamson, 1994, pp. 205-9). Interestingly, Eklund’s form of Inconsistency theory appears perfectly compatible with Epistemicism, at least with respect to the latter’s commitment to sharp boundaries.

⁴² (Eklund, 2002, p. 264).

speakers find the major premise of the sorites so alluring. But what reason do we have for thinking that TP *is* an MCP?

In offering support for this claim, Eklund revisits Dummett's argument from observational terms. Eklund says that 'it seems that "looks red" must be tolerant. For then differences in colour that cannot be detected by mere observation should not make a difference to the applicability of the predicate.'⁴³ As noted above, however, I shall argue in Chapter 3 that Dummett's focus on observational terms is too narrow. Eklund's espousal of Dummett's argument ultimately will provide no great support for his claim. Eklund does provide some further considerations in favour of TP's status as an MCP. He suggests that someone not 'taken in' by the sorites argument would be considered to have '*missed* something the rest of us *see*'.⁴⁴ He infers from this that tolerance plausibly is a genuine feature our ordinary understanding of vague terms. Eklund also suggests that *all* attempts to remove inconsistency from vague terms, such as the Supervaluationist and Epistemicist theories encountered in Section 1.2, seem counterintuitive in *some* way. He infers from this that we must thereby be attempting to eradicate something that it is truly part of the meaning of such expressions.⁴⁵

Neither of Eklund's additional arguments is particularly compelling, however. Both arguments essentially appeal to the common intuition that the major premise of the sorites seems *true*, which is precisely the phenomenon in need of explanation. It simply will not suffice to take these linguistic intuitions as primitive:

⁴³ (*Ibid*, p. 258).

⁴⁴ (Eklund, 2005, p. 41).

⁴⁵ (*Ibid*, p. 42).

a deeper explanation of their psychological provenance is required. Eklund seems to recognise the limitations of his theory, claiming that,

The real justification of the account will lie in the plausibility of what is eventually churned out by these considerations.⁴⁶

Eklund is thus greatly relying on his account constituting an inference to the best explanation of the allure of the major premise. The problem, of course, is that the account is far too schematic and promissory to offer a completely satisfying explanation of vagueness. Coupled with the lack of non-intuitive evidence for the claim that TP is an MCP, Eklund seems to be taking too much on faith. The market is therefore open, I feel, for a more substantive Inconsistency Theory, founded on more cogent philosophical argumentation and a wider range of empirical evidence. I hope to deliver on this promise in what is to follow.

1.3.3 Cognitive Inconsistentism

In summary, I shall be arguing for the following three claims:

A3. Linguistic understanding consists in cognising, or mentally representing, a set of syntactic, phonological, and semantic rules, principles, and constraints.

B3. Our semantic competence partly consists in the mental representation of inconsistent constraints.

⁴⁶ (*Ibid*).

C3. Our semantic competence *causally explains* our acceptance of the major premise of the sorites.

I shall defend claim A3 in Chapter 2, introducing a broadly Chomskyan, cognitive theory of linguistic understanding. The Chomskyan theory, as I shall discuss later, posits mentally represented rules and principles in order to explain various aspects of our linguistic capacity. Crucially, in Chapter 3, I shall argue that important features of language acquisition and use would be best explained by a mentally represented ‘similarity constraint’, operative on conceptual classification. There is evidence to suggest that this constraint forms part of our cognitive semantic competence. The similarity constraint is inconsistent in the sense that it comes into conflict with other constraints on our linguistic and conceptual processes. I shall argue that this similarity constraint causally explains our linguistic intuitions with respect to the major premise of the sorites. Unlike Dummett’s Incoherentism, on which our linguistic understanding is taken to consist in a set of practical abilities, Cognitive Inconsistentism thus appeals to features of our actual psychology in explaining our paradox-inducing understanding of vague expressions. Unlike Eklund’s Inconsistency Theory, on which semantic competence is taken partly to consist in our dispositions to accept tolerance principles, Cognitive Inconsistentism provides a deeper cognitive explanation of the competence underlying these aspects of linguistic performance. I now ask the reader to bear in mind these aims as I move on to discuss theories of linguistic understanding in more detail.

2. Linguistic Understanding: Abilities and Knowledge

2.1 Introduction

In Chapter 1, we encountered three Inconsistency Theories, on which vagueness *qua* sorites susceptibility is explained by aspects of our linguistic understanding. Michael Dummett's view, Incoherentism, held that linguistic understanding consists in possessing a set of *practical abilities*, or *knowledge how*. Matti Eklund's Inconsistency Theory held that understanding consists in *dispositions to accept* certain meaning constitutive principles. I shall argue that both of these views are mistaken. My first target will be Dummett's abilities conception of linguistic understanding. I shall argue that we have strong reason to reject Dummett's view in favour of a cognitive, broadly Chomskyan, account of linguistic understanding. I shall later use this Chomskyan account to demonstrate the shortfalls of Eklund's inferential role semantics. The Chomskyan account, I shall argue, also has the potential to offer a far better explanation of the allure of the major premise of the sorites. This latter claim will concern me in Chapter 3.

2.2 The Major Divide in Theories of Understanding

What is it to understand a language? What does linguistic understanding consist in? In this section I shall outline *the* major divide on these issues, between cognitive and non-cognitive theorists. I shall introduce an exemplar of each view, Chomsky on the cognitive side, and Dummett on the non-cognitive side. In section 2.3, I shall criticise Dummett's *abilities* based view of linguistic understanding.

Perhaps the most plausible cognitive account of linguistic understanding is that of Noam Chomsky. Chomsky explicates the concept of ‘knowledge of language’ as the ‘mastery and internal representation of a specific I-language’.⁴⁷ An I-language ‘consists of a computational procedure and a lexicon’.⁴⁸ Both the computational procedure and the lexicon encode ‘a system of rules and principles... that assign interpretations to utterances’.⁴⁹ The computational procedure is taken to generate *syntactic* structures for inputted utterances.⁵⁰ I shall be saying very little about syntax. Entries in the mental lexicon, according to Chomsky, encode syntactic, morphological, phonological, and semantic properties of terms. For instance, our entry for *tiger* includes information that ‘*tiger* is a noun, it is pronounced with an initial /t/; it means TIGER’.⁵¹ My main concern will eventually be with the acquisition and representation of *semantic* properties, on which more in Chapter 3. Chomsky holds that a speaker’s acquisition of an I-language is facilitated by an ‘initial state’, or ‘language acquisition device’, (LAD), which ‘takes experience as “input” and gives the [I-] language as an “output”’.⁵² The LAD is a genetically determined, innate state of the mind/brain, which Chomsky famously believes is more or less homogeneous across the human species.⁵³ On the I-language theory, the task of empirical linguistics is to make explicit both a) the I-language generating

⁴⁷ (Chomsky, 2000, p. 73).

⁴⁸ (Chomsky, 1995, p. 15). This corresponds to the *broad* sense of ‘faculty of language’, discussed in (Hauser, Chomsky, & Fitch, 2002). The *narrow* sense of ‘faculty of language’ adverts only to the computational system.

⁴⁹ (Chomsky, 2000, p. 69). Chomsky talks of interpreting *utterances*, which suggests that I-languages require auditory inputs. But I-languages evidently can also interpret inputs from visual, (written/sign language), senses.

⁵⁰ I have omitted discussion of the role of a ‘parser’ system in *applying* the rules and principles represented in the I-language to inputted utterances. See e.g. (*Ibid*, pp. 69-70).

⁵¹ (Smith N. V., 2004, p. 130).

⁵² (Chomsky, 2000, p. 4). The current best theory of the human LAD for *syntax* takes it to function by setting various *parameters*, whereby for each language that one could acquire, there exists a certain profile of parameters, to which the child will converge, given exposure to linguistic evidence.

⁵³ (Cowie, 2008) offers a comprehensive survey of the controversy surrounding Chomskyan nativism. Empiricists do not deny the role of innate endowment in language acquisition. The issue really is over the richness and domain-specificity of that innate endowment.

function encoded in the LAD, and b) the rules and principles – syntactic, phonetic, semantic – encoded in specific I-languages. In either case, linguistic theory is taken to be about human *psychology* and our mental representations of such rules and principles.

Chomsky famously distinguishes between *competence* and *performance*. A speaker's competence *consists* in her possession of an I-language, and thus in the mental representation of a computational procedure and a lexicon. Performance, in contrast, relates to the numerous manifestations of that competence, 'in articulation, interpretation, expression of beliefs and desires, referring, telling stories, and so on'. Competence is not completely divorced from performance: our I-languages, Chomsky claims, are integrated with 'performance systems', which 'make use of expressions generated by the I-language' in the production of actions. Nor is performance deemed irrelevant or uninteresting: it is simply an ensemble effect of a range of cognitive and behavioural systems, the details of which are a highly complex matter. The take-home point for current purposes is that Chomsky clearly distinguishes between the ability to *use*, or employ one's language in practice, on the one hand, from *mental representation* of language, or *competence*, on the other.⁵⁴

In contrast with the Chomskyan, cognitive approach, however, many philosophers have eschewed focusing on the psychology of individual language users. One way to view this prominent philosophical view is in relation to Wittgenstein's claims in the *Philosophical Investigations*, that we ought to 'try not to think of understanding as a "mental process" at all'.⁵⁵ Wittgenstein instead

⁵⁴ See (Chomsky, 2000, pp. 50-51).

⁵⁵ (Wittgenstein, 2001, §154). This passage is noted in (Dummett, 1993, p. 99).

commends the view that ‘to understand a language means to be master of a technique’, (§199). Understanding a language, on this view, is taken to consist in a set of *practical abilities*, not a system of mental representations. On this view, unlike Chomsky’s, there would be no conceptual distinction between understanding a language, or competence, and being able to use it, or performance: the latter is deemed to be explicative of the former.⁵⁶

As we saw in claim A1 of section 1.3.1, Michael Dummett has been a major proponent of this *abilities* view of linguistic understanding. In an illustrative passage, Dummett claims that ‘what a speaker knows, in knowing the language, is how to use the language to say things, i.e. to affect linguistic acts of various kinds’.⁵⁷ On the resulting conception, ‘that in which our understanding of the language consisted would lie open to view... in our use of the language, in our participation in a common practice.’⁵⁸ As noted above, Dummett does retain a role for theoretical, propositional models in *representing* these practical abilities. A theory of meaning, or *sense*, would be given in terms of semantic rules – including, as we saw, *tolerance principles* with respect to vague terms. But Dummett qualifies this point by saying,

A theory of meaning of this kind is not intended as a psychological hypothesis. Its function is solely to present an analysis of the complex skill which constitutes mastery of a language...; it is not concerned to describe any inner psychological mechanisms which may account for his having those abilities.⁵⁹

⁵⁶ Wittgenstein would not have liked the notion ‘explicative’.

⁵⁷ (Dummett, 1993, p. 38).

⁵⁸ (*Ibid*, p. 102).

⁵⁹ (*Ibid*, p. 38).

The theory of sense itself, *qua* semantic rules, simply ‘plays no genuine role in the account of what constitutes a speaker’s mastery of his language’.⁶⁰ This is all some distance from the Chomskyan view, on which a theory of meaning represents the *object* of a speaker’s knowledge of language, which in turn is revealed by empirical investigation into actual speakers’ cognitive states.

It would surely be wrong, however, to attribute to Dummett the view that the individual’s psychology has *nothing at all* to do with her understanding of language. By that same token, we would be wrong to think of ‘abilities’ in a purely behavioural sense. On any plausible view, our abilities to use language will be underpinned and supported by a complex range of psychological processes. As we saw in relation to the expression ‘red’, for instance, our observational use of this term clearly implies a substantial role for perception and memory. More generally, the claim that understanding consists in our *participation* in a practice, and *use* of language, imports overtly intentional idioms, themselves slave to psychological interpretation. In fact, this intentional aspect of understanding is a major feature of Dummett’s theory. He claims that ‘any adequate philosophical account of language must describe it as a rational activity on the part of creatures to whom can be ascribed *intention* and *purpose*’.⁶¹ Even abilities theorists, then, would allow that human psychology is integral to the explanation of linguistic understanding.

In deeming the abilities theorist ‘non-cognitive’, then, I do not mean to imply that mental state vocabulary will be completely absent from an account as to our use of language. Rather, the claim is that the mental states which *do* figure in that

⁶⁰ (*Ibid*, p. 17).

⁶¹ (*Ibid*, p. 104).

account are *not domain specific*. By that I mean that the mental processes purportedly underlying our linguistic abilities, related to perception, intention formation, coordination of behaviour, and so on, are not uniquely deployed in language processing. In contrast, on the Chomskyan view, our mind/brain represents *specifically linguistic* rules and principles, which are deployed in language acquisition and processing. This contrast will arise again later, in discussion of language acquisition in section 2.5.

In summary, the major divide in theories of linguistic understanding is between cognitive, and non-cognitive theorists. Cognitive theorists take linguistic understanding to consist in the internalisation, or mental representation, of a set of linguistic rules and principles. In contrast, non-cognitive theorists such as Dummett take understanding to consist in the complex set of abilities involved in our participation in a linguistic practice. In the next section, I shall present some challenges to Dummett's view of linguistic understanding. I shall then go on to discuss the comparative explanatory power of the cognitive conception.

2.3 The Role of Intention in Language Acquisition

On Dummett's view, understanding consists in our 'use of the language, in our participation in a common practice.' Dummett sees our ability to use language as a necessarily rational activity, undertaken only by beings to which we can ascribe *intentions* and *purposes*. The range of intentions and purposes, moreover, will be complex and iterated. In order truly to *participate* in a practice, speakers must not only intend to use language in conformity with others. They must also believe that

their use will be *recognised* by others as being in conformity, and recognised as being performed with that intention. Such social abilities would thus arguably presuppose some understanding of the mental states of others. In this section, I want to assess the prospects for such an account. My appraisal will be largely negative.

Dummett's view of understanding as inherently intentional seems incompatible with several apparent facts about language acquisition. On Dummett's view, if understanding consists in our intentional participation in linguistic practice, it seems plausible that our acquisition of understanding would crucially involve the *formation* of such intentions. A first criticism, however, is that intentions themselves possess meaningful contents. As such, it seems illegitimate to employ such attitudes in an account of our *acquisition* of an understanding of meaning. Let's assume for argument's sake that an intention to use an expression *F* in conformity with a practical convention *C*, consists in a belief-desire pair: a belief that *F* is used in conformity with *C*, and a desire to comply with *C*. The formation of such an intention, however, clearly requires some prior understanding as to what it means to use an expression in conformity with a convention, and what it means to comply with a convention. Dummett therefore owes us an account as to what *this* understanding consists in, which does not appeal to the notion of meaning.⁶²

Dummett portrays linguistic understanding as involving rational, purposive action on the part of speakers. A second criticism, however, is that language acquisition, at its most prolific stage, appears to be a perfectly unintentional phenomenon. In the words of Chomsky, 'it is something that happens to the child,

⁶² We can fairly assume that Dummett would not want to appeal to a prior, perhaps innate, representational system – a 'language of thought' – in which these pre-linguistic intentions are formulated.

not that the child does'.⁶³ This claim is supported by both commonsense reflection, and empirical evidence. First, with tongue in cheek, it bears mention that even the most obstinate, contrary, and uncooperative infants acquire language. Young infants do not need to be encouraged to 'play the language game', in the same sense that they are coaxed to follow social norms of politeness, cultural eating conventions, and so on.⁶⁴ It is sufficient for a child to learn a language that they simply be *exposed* to linguistic stimulation: no active participation on their part is required.

Second, *participation* in a practice requires a complex range of iterated propositional attitudes. These might include beliefs about other speakers' linguistic intentions, beliefs about their beliefs about one's own intentions, and so on. Empirical evidence suggests, however, that language acquisition reaches a fairly advanced stage before infants acquire an understanding of other people's mental states. Children generally start producing 'telegraphic speech' – utterances such as 'more cookie', or 'bring toy' – around the age of 18 months.⁶⁵ Although infants suppress subject terms, determiners, prepositions, and such like, even at this stage word order errors are extremely rare, suggesting that the infants possess an early understanding of syntactic structure.⁶⁶ This supposition is supported by the fact that such children seem to *comprehend* fully grammatical sentences better than telegraphic sentences.⁶⁷ At around the same age, 18 months, children also undergo a radical increase in their acquisition of new words, one estimate being that they acquire

⁶³ (Chomsky, 2000, p. 7). Chomsky makes the further provocative claim that 'in certain fundamental respects we do not really learn language; rather, grammar grows in the mind', (Chomsky, 1980, p. 134).

⁶⁴ Learning a language is thus *not* like learning how to play a game like chess. I do wish that people would stop using that analogy. E.g. (Dummett, 1959, p. 2); (Wright, 1976, p153; 1987, p. 206).

⁶⁵ (Brown & Bellugi, 1964).

⁶⁶ (Bloom P. , 1990); (Brown, 1973).

⁶⁷ (Gerken & McIntosh, 1993).

around nine new words a day.⁶⁸ By the age of 2 ½ years, children start to progress from two word utterances to fully formed sentences, including many of the syntactic categories previously omitted in speech production.⁶⁹ By the age of 3-4 years, the overwhelming majority of their sentences are fully grammatical.⁷⁰

In contrast, aspects of the child's 'Theory of Mind' (ToM) are comparatively slow to develop. I take 'ToM' in a fairly minimal sense, simply embodying our ability to understand, predict, and explain other people's mental states.⁷¹ Infants are typically able to report their own desires, and attribute desires to others, by age 2 years.⁷² Comprehending and producing propositional attitude locutions, however, only occurs at around 3 years of age.⁷³ The litmus test for a fully fledged ToM, namely performance in 'false belief tasks', typically is only achieved by around 4 years of age.⁷⁴ This leaves a significant period, typically between the ages of 18 months and 4 years, in which the child develops an almost fully developed grasp of syntax, and a broad understanding of word meanings, but during which her ability to comprehend the mental states of others is limited. If linguistic understanding consists in intentional, purposive participation in a practice, and iterated attitudes about shared beliefs amongst members of one's language community, then we

⁶⁸ (Carey, 1978).

⁶⁹ For a summary see the introduction in (Bloom P. , 1993).

⁷⁰ (Stromswold, 1999, p. 358).

⁷¹ I do not presuppose an 'internalist' notion of ToM, on which the theory is mentally represented, perhaps in modular form, perhaps innately endowed. My argument is compatible with taking ToM simply as a range of mind-reading *abilities*. My interpretation should thus be acceptable to Dummett. For further discussion, see (Ravenscroft, 2004).

⁷² (Gopnik & Slaughter, 1991).

⁷³ (de Villiers & de Villiers, 1999).

⁷⁴ (Baron-Cohen, Leslie, & Frith, 1985). Performance in a false-belief task involves a child, A, correctly attributing to another child, B, a false belief, typically as to the location of a toy or sweet that has been moved without B's knowledge.

should deny that such infants have anything like such understanding. But this seems an entirely unjustified conclusion, given the developmental evidence.

In summary, Dummett's abilities view of understanding has implausible commitments with respect to the intentional involvement and mental capacity of language learners. Language acquisition seems to be an effortless, passive process, and does not require intentional, purposive participation in a communal practice. The early development of infants' linguistic understanding seems unencumbered by their lack of ToM.⁷⁵ In the next section, I shall offer some reasons to favour the separation of our ability to use language, from our understanding of language.

2.4 The Separation of Competence and Performance

There exist persuasive apriori reasons to distinguish performance, abilities-based aspects of language use, from cognitive, representational aspects of our understanding. A first argument, introduced by Chomsky, is that speakers might improve their ability to *use* language, while intuitively showing no signs of improved understanding. Chomsky considers the case of Jones, who 'takes a public speaking course and improves his ability to speak and understand without any change in his knowledge of English, as we would describe the situation in normal usage'.⁷⁶ Chomsky also reports cases in which an accident, or stroke, has left a patient speechless, but who has later recovered to show no signs of damage to their

⁷⁵ In fact, studies reported in (Garfield, Peterson, & Perry, 2001) suggest that language development *enables* ToM development. While advanced early performance in ToM tasks proved to be no indication of advanced future linguistic performance, early linguistic ability proved a strong indicator of subsequent ToM ability. Studies on deaf children born to hearing families not fluent in sign language suggest that their delayed linguistic development strongly affects subsequent ToM performance. The control group, deaf children born to families fluent in sign language, showed no comparable deficit.

⁷⁶ (Chomsky, 1986, p. 10)

knowledge of language.⁷⁷ The argument in effect establishes that linguistic understanding does not co-vary with a speaker's ability to produce language: the latter ability can improve, or diminish, whilst understanding of language remains at a constant level. It seems most natural to ascribe a constant property of knowledge of English, which persists in the speaker's mind/brain throughout her temporary loss of these performance capabilities.⁷⁸ As Chomsky notes, her 'recovery provides evidence for the fact.'

In both cases, however, Chomsky seems to have taken quite a narrow reading of 'ability'. Both cases pertain specifically to a subject's ability to speak, or orally produce language. It might well be objected that 'ability to use' language is meant in the broader sense discussed above, signifying the intentional, rational, and purposive participation in a communal linguistic practice. This might allow the abilities theorist to accommodate the Jones and stroke patient cases. In both cases, they might say, the speakers retain an unchanged set of beliefs about communal usage, intentions to use language in conformity with that usage, and a belief that others intend to do likewise, and so on. In short, while their *diction* or *speech production* is altered, their 'ability to use' language, in this broader, intentional sense, remains constant.

A better argument against the abilities view will therefore dissociate our understanding of the syntax and semantics of English, both from our ability to produce language, *and* from any form of intentional, purposive states with respect to language use. Contemporary clinical research on patients in a vegetative state

⁷⁷ (Chomsky, 1988, p. 10)

⁷⁸ (Chomsky, 1980, pp. 51-2).

provides a rich source of evidence here. On traditional criteria, a patient is deemed to be in a vegetative state (VS) if there is 'no evidence of awareness of self or environment, no evidence of sustained, reproducible, purposeful or voluntary behavioural response to visual, auditory... stimuli and critically no evidence of language comprehension or expression'.⁷⁹ Interestingly, however, a study by Martin Coleman and colleagues found patients who satisfy all diagnostic criteria for VS, including lacking any signs of any willed responses to stimuli, except that they seem to retain aspects of language comprehension. Comprehension was measured in three stages: a) response to auditory, as opposed to silent stimuli; b) response to intelligible, as opposed to unintelligible speech; and c) response to both phonetically ('rain'/'reign'), and semantically ('bark'/'bank') ambiguous, as opposed to unambiguous sentences. Measurements were taken of differential brain activity, using fMRI techniques, when patients were presented with the contrasting forms of stimulus, and the results compared to responses in normal subjects. Of seven VS patients tested, three were found to respond differentially both to a) auditory/silent, and b) speech/non-speech stimuli. In addition, two patients showed significant signs of differential processing corresponding to c) ambiguous/unambiguous sentence stimuli. The interpretation of these results undoubtedly requires care. The marked responses to phonetically and semantically ambiguous sentences, however, strongly suggest that auditory inputs are being assigned phonetic and semantic interpretations. If true, this would mark a significant level of linguistic processing. The Chomskyan can accommodate this fact comfortably: on the I-language view, competence partly *consists* in the representation of rules governing phonetic and

⁷⁹ (Coleman, et al., 2007, p. 2495).

semantic interpretation. The absence in VS patients of even rudimentary signs of the potential for purposeful or voluntary behaviour, however, means that the abilities theorist is extremely short on explanatory options.

2.5 Language Acquisition and Poverty of Stimulus

So far, my agenda has been fairly negative. We have yet to see any significant, independent motivation for the Chomskyan view. Amongst all else, perhaps the single most important advance offered by the Chomskyan view of linguistic understanding is in explaining the *acquisition* of language. There are at least two discernible explanatory problems relating to language acquisition. First, the task facing the child learner in acquiring a language is of remarkable magnitude and complexity. That the child is able to complete this task at all, and moreover with an invariably high degree of success, demands explanation. Second, the child acquires an unboundedly creative and productive linguistic capacity. As put by Wilhelm von Humboldt, language involves the 'infinite use of finite means'. How the child acquires this capacity, given only exposure to minimal, and finite samples of language, also requires explanation. Both of these problems relate to so-called 'Poverty of the Stimulus' considerations. In this section, I shall argue that Chomskyan theorists are uniquely placed to explain how learners overcome these problems, and converge upon the correct syntactic and semantic rules for their language.

The core Poverty of the Stimulus (POS) consideration is that 'knowledge acquired in language acquisition far outstrips the information that is available in the

environment’.⁸⁰ That information is called ‘primary linguistic data’ (PLD), which include samples of language to which the child is exposed, the positive or negative responses to the child’s own language production, and such like. Corresponding to the first explanatory problem outlined above, the following four points outline the POS problem of ‘theory selection’ in language acquisition. First, the child learner has an indefinite number of alternative syntactic, semantic, phonological, principles that are consistent with its PLD. Second, the child’s PLD typically contain insufficient information to rule out most of these alternatives. Third, despite this, the child consistently and reliably arrives at the set of syntactic, semantic, phonological, principles for its language. Fourth, this convergence is inherently surprising, as the principles converged upon are typically not the simplest theoretical alternatives.

Taking the first problem of language acquisition, let us consider an example of a theory selection problem with respect to a particular *syntactic* principle.

Consider the following sentences:

1. Ecuador is in South America.
2. Is Ecuador in South America?
3. Julie’s telephone is working again now that the storm is over.
4. Is Julie’s telephone working again now that the storm is over?

⁸⁰ (Laurence & Margolis, 2001, p. 221). By using the term ‘knowledge’, I do not wish to prejudge whether this signifies *mental representations*, or *practical abilities*. POS considerations arise with equal force for the acquisition of abilities.

The simplest principle governing the transformation in these sentences is clearly to move the first 'is' in 1 and 3 to the front of the sentence, to form the questions in 2 and 4. But now consider the following pair of sentences:

5. That woman who is walking her dog is Tom's neighbour.
6. * Is that woman who walking her dog is Tom's neighbour?

In this case, application of the simple principle for 'is' on 5 yields an ungrammatical sentence in 6. The correct transformation would in fact be as follows:

7. Is that woman who is walking her dog Tom's neighbour?

The operative syntactic rule here is to move the auxiliary verb from the main clause to the front of the sentence. Sentence 6 is ungrammatical because the auxiliary that is moved is embedded in the subject noun phrase.⁸¹

The example is instructive, as children are rarely, if ever, exposed to utterances such as 7, from which they could rule out the simple rule for moving 'is'. It would be even rarer for a child to *produce* a sentence such as 6 or 7, from which they could receive negative or positive feedback, respectively. In short, there typically would be little or no information in the child's PLD that would inform them of the incorrectness of the simple rule, or of the correctness of the more complex rule. Nonetheless, children invariably acquire an understanding of the more complex rule, and display virtually error free performance in applying that

⁸¹ (*Ibid*, pp. 222-23).

rule.⁸² Even more remarkable, given the child's lack of exposure to explicit syntactic information, is that the rule acquired is *structure dependent*: the rule does not advert to surface transformations, such as *moving the first 'is'*, but makes essential reference to complex syntactic categories, such as AUXILIARY VERB and MAIN CLAUSE. This strongly suggests, however, that the infant *represents* such categories, at some level, and deploys them in the formation, or amendment of candidate rules. But such categories are *clearly* not marked in ordinary speech, or presented to the child through her PLD. The crucial issue, then, is *how* the child could possibly converge on such rules, given her hugely impoverished linguistic evidence.⁸³

The Chomskyan, cognitive theory of linguistic understanding excels in this explanatory endeavour. The crucial structural commitment is to the infants' language acquisition device (LAD), a mentally represented function from given linguistic inputs to a constrained range of output syntactic rules.⁸⁴ The LAD *constrains* the set of possible syntactic rules available to interpret these sentences: there are only a limited range of alternative principles output by the LAD. Once the LAD has output the syntactic rule, that rule is represented in the child's I-language, and serves to interpret novel sentences with the same structure.

In stark contrast, the abilities theorist lacks any sensible means of explaining the child's faultless achievement in acquiring such rules. As we saw above, Dummett believes that linguistic theory, whether syntactic or semantic, 'plays no

⁸² Karin Stromswold (Stromswold, 1999, p. 362) reports that in more than 66,000 utterances transcribed from the speech of 14 children, she found no clear errors in their use of auxiliary and lexical verbs.

⁸³ For a summary of the child's similarly sophisticated phonemic and phonetic knowledge, see (Plunkett, 1997).

⁸⁴ The particular details of the syntactic rules, or the nature of the theory in which they are stated, are of secondary importance. As noted by Stephen Laurence, (Laurence, 2003, p. 75), 'the Chomskian account of the nature of linguistics is not tied to any particular linguistic theory, and certainly does not require the truth of current linguistic theory in any detail.'

genuine role in the account of what constitutes a speaker's mastery of his language'. Furthermore, in specifying such a theory, he was 'not concerned to describe any inner psychological mechanisms which may account for' our linguistic abilities. If Dummett's abilities view of understanding *can* explain the relevant data with respect to our judgements of sentences 6 and 7, it remains to be demonstrated. As Chomsky rightly notes, 'someone who believes that this kind of understanding is "simply a practical ability," a "skill" of some sort, owes it to us to provide some account of the nature of this ability or skill; otherwise the proposal is empty'.⁸⁵

Let us turn now to the second POS problem, relating to the infant's acquisition of a potentially unbounded creative and productive linguistic capacity. All parties to the debate, including Dummett and Chomsky, agree that syntactic rules – whether mentally represented or not – must possess the property of *recursiveness*. Put simply, recursion allows us to generate increasingly complex, and indefinitely long, grammatical sentences given just a finite set of primitive terms, and rules for their well-formed combination. For instance, the rule ' $VP \rightarrow V S$ '⁸⁶ allows sentences (S) to occur embedded in verb phrases (VP), which in turn can feature in more complex sentences, such as 'Mary thinks [Harry stirred the stew]', and 'I suspect [Mary thinks [Harry stirred the stew]]', and so on.⁸⁷ Whereas Dummett believes that recursive syntactic theory has no basis in human psychology, however, the Chomskyan holds that these rules are mentally represented, and thus

⁸⁵ (Chomsky, 1986, p. 259)

⁸⁶ Such 'phrase structure' rules, typically of the form $X \rightarrow YZ$, signify that a sentential constituent X, or possibly a whole sentence, is broken down structurally into two sub-constituents Y and Z.

⁸⁷ (Smith N. V., 2004, p. 55)

causally implicated in the speaker's linguistic processing.⁸⁸ Dummett cannot avail himself of any causal-explanatory role for such recursive rules in our actual psychology. As such, whatever recursive properties are possessed by the syntactic theory, they will simply fail to offer any explanatory purchase on the issue as to how infants *in fact* acquire an understanding of those properties.⁸⁹

2.6 Problems with Chomsky

The Chomskyan view of linguistic understanding, despite its attractions, is not without faults. In this section, I shall outline some major challenges facing the cognitive theorist, and sketch some possible lines of response.

2.6.1 The Impropriety of 'Knowledge'

It will no doubt have been noticed that I have used the term 'knowledge' to describe the cognitive relation *R* borne by speakers to the rules and principles represented in their I-languages. This may well seem an inappropriate usage, however, as this putative cognitive relation evidently lacks many of the features ordinarily deemed distinctive of knowledge.

It is commonly assumed that, if *X knows p*, then 1) *p* is accessible to *X*'s consciousness, would be assented to by *X* if prompted, etc.; 2) *X* possesses the concepts deployed in *p*; 3) *p* is inferentially integrated with *X*'s other beliefs, knowledge etc.; and 4) *p* is true. In relation to the first feature, however, speakers are

⁸⁸ One possible worry is whether our mental representations of such rules would themselves be subject to recursive combination, perhaps in a 'Language of Thought' (LOT). If so, would we represent further recursive rules for our LOT? And would we require yet more recursive rules for *these* representations? An infinite regress looms large. See (Laurence & Margolis, 1997) for a defence of the LOT hypothesis against regress.

⁸⁹ I have benefited here from discussion of Dummett and Wright in (Antony, 1997).

utterly unaware of many of the rules and principles encoded in their I-language: these rules are *inaccessible to consciousness*.⁹⁰ Moreover, were these rules and principles to be described to the average speaker, they would probably not understand them, nor assent to them.⁹¹ On the second point, the concepts deployed in putative I-language representations arguably outrun speakers' explicit epistemic resources. Few other than professional linguists possess the concept AUXILIARY VERB, for instance; but from the examples considered in Section 2.6, we concluded that speakers must grasp this category at some level.⁹² On the third feature, I-language representations arguably are not inferentially integrated with the subject's other beliefs and knowledge. Let us grant, for instance, that my I-language represents the syntactic rule $VP \rightarrow V S$. Now suppose I avowedly believe, for some reason, that if $VP \rightarrow V S$,⁹³ then I shall cease to read linguistics. (Perhaps I have read a linguistics textbook, and concluded that, if *that* is the truth about the structure of verb phrases, then linguistics is simply too complicated for me.) Now if my I-language representation of the syntactic rule were inferentially integrated with that belief, I should be able to deduce that I shall cease to read linguistics. However, that representation is simply not available in such inferential tasks.⁹⁴ Finally, in relation to the fourth point, the relation *R* borne to I-language representations is arguably non-factive. For instance, I might through some erroneous process come to represent

⁹⁰ For discussion, see (Davies, 1989). There may be an intuitive difference here between the accessibility of syntactic and semantic principles. I shall not discuss that point here.

⁹¹ (Stich, 1971, pp. 485-6).

⁹² Crispin Wright (Wright, 1986, pp. 281-3) runs a similar 'argument from technicality' against the claim that speakers could represent a recursive *truth* theory.

⁹³ I am treating the phrase structure rule as if it were propositional, and thus apt to form the object of a belief.

⁹⁴ As noted in (Davies, 1989, p. 134), it is not quite right to say that our representations of such rules are inferentially isolated: they will combine in *linguistic* inferences, the products of which might well yield ordinary beliefs as consequences. There clearly is a sense, however, in which the charge of doxastic isolation is correct.

the semantic rule that ‘dog’ means FORK, say. But on most views, it would be strictly speaking *false* that ‘dog’ means FORK.

The cumulative weight of evidence strongly suggests that ‘knowledge’, at least as characterised by features 1-4, is not the correct term to describe the cognitive relation *R* borne to representations in one’s I-language. Chomsky has been fairly unsympathetic to this challenge, however. He pretty much sidesteps the issue by introducing the *term d’art* ‘cognisance’, thereby disengaging from more traditional philosophical concerns regarding the concept ‘knowledge’.⁹⁵

Underlying Chomsky’s dismissiveness, however, is a perfectly respectable line of thought. Chomsky is perfectly theoretically justified in positing *some* sort of representational states, encoding the syntactic and semantic rules for our language, in virtue of the resultant explanatory advantage with respect to language acquisition and processing. Even granting that the term ‘knowledge’ would cause unnecessary confusion, (or annoyance), it remains the best explanation of our linguistic capacities that language-specific representational states are causally implicated in our linguistic cognition. It is this *explanatory* focus, as opposed to the merely descriptive task of systematising our linguistic practice, which distinguishes the Chomskyan approach. This methodology is standard in the empirical sciences: we would not think twice if physicists told us that certain puzzling gravitational phenomena could only be explained by the existence of a hitherto unknown, and presently unidentifiable, particle with properties *X, Y, Z*.

Philosophers sympathetic to this thought, and to the explanatory program in the cognitive sciences, have attempted to explicate the requisite type of

⁹⁵ (Chomsky, 1980, pp. 92-102).

representation under monikers such as ‘tacit knowledge’, or ‘subdoxastic states’.⁹⁶

There has also been resistance to such notions from many quarters.⁹⁷ Many complex issues arise out of this debate. What is the causal role of such states? Under what conditions are we justified in attributing such states? What notion of content is involved in such states? I do not have time here, unfortunately, to discuss these themes. I am happy, however, simply to express my faith that *some* such notion of representation, once defined, will prove indispensable to any theorist working in the cognitive domain.⁹⁸ Explanatory spoils to the victor.

2.6.2 The Objectivity of Language

The Chomskyan approach to linguistic theory is unabashedly *internalist*, in both a methodological and an ontological sense. Methodologically speaking, in constructing linguistic theories, Chomsky believes we should take the psychological states of individual speakers as our object of study. Linguistic theory, then, is a theory about our I-languages, or linguistic competence. Ontologically speaking, a subject *S*’s linguistic competence with respect to a language *L* is taken to consist in a set of mentally represented rules and principles, which Chomsky ultimately takes to be a neurophysiological structure in *S*’s brain. In this sense, any neurophysiological duplicate of *S* would also be a duplicate of *S* with respect to competence in *L*.

⁹⁶ (Evans, 1985), (Davies, 1989), (Stich, 1978).

⁹⁷ Crispin Wright (Wright, 1986) expresses extreme doubts about the notion of tacit knowledge. Note that Michael Dummett, e.g. (Dummett, 1993, p. 69) uses the expression ‘implicit knowledge’ to refer to knowledge *how*. Dummett, like Wright, however, would not subscribe to a genuinely representational notion of implicit or tacit knowledge.

⁹⁸ As Zenon Pylyshyn notes, (Pylyshyn, 1999, p. 10), ‘to explain the simplest fact about the behaviour of most “higher” organisms, we must say how some aspect of the world is represented – and this applies even where the behaviour does not appear to involve reasoning or rational decision making’.

These internalist commitments have met with much philosophical resistance, however. The nature of these disagreements is of the utmost philosophical significance; I shall here offer just a superficial summary of their import. For a start, many theorists think that Chomsky's methodological internalism illegitimately circumscribes the object of linguistic study. Languages, such theorists contend, are *social* objects, and ought to be studied as such. A representative of this view, Michael Dummett, says,

A language, in the everyday sense, is something essentially social, a practice in which many people engage; and it is this notion, rather than that of an idiolect, which ought to be taken as primary.⁹⁹

Dummett is not alone; Bill Lycan agrees that 'what a linguist studies is the structure of a *public* language, the language of some speech community... Linguistics is the study of what the members of a speech community have in common'.¹⁰⁰

What is *driving* these claims? I suspect that it boils down to worries relating to the triumvirate of *meaning, reference, and communication*. Suppose we adopt Chomsky's method in constructing a theory of meaning, and set out to study the semantic features or properties that are represented in speakers' mental lexical entries for, say, 'dog'. There is simply no guarantee that the information represented in *my* I-language about the meaning of 'dog', will be the same as the information that *you* represent about the meaning of 'dog'. Furthermore, granted the assumption that a term's meaning (somehow) determines its *reference*, there is likewise presumably no guarantee that you and I refer to the same property by 'dog'. On the Chomskyan

⁹⁹ (Dummett, 1993, pp. 30-31).

¹⁰⁰ (Lycan, 1986, p. 247).

theory, then, you and I could conceivably mean, and refer to, different things by ‘dog’. But if that were so, then *how is it that we successfully communicate?*¹⁰¹ In summary, the worry is that, if meaning is to undergird successful communication, and shared reference, then meaning must be a property of public, shareable linguistic entities, *not* I-languages. These concerns are widespread: as Chomsky notes, ‘virtually all contemporary reflections about language and mind take the expressions that have sound and meaning to be in a “public language” that is external to the mind-brain’.¹⁰²

Chomsky might well complain, however, that our purported example of the internalist methodology is unrealistic. While the focus of linguistic theory undoubtedly is on the rules and principles mentally represented by individuals, the theory will invariably be highly idealised. As Chomsky notes, ‘in reality what [two individuals] have in their heads is about as interesting for naturalistic inquiry as the course of a feather on a windy day.’¹⁰³ The outcome of empirical inquiry, in contrast, will be an *abstraction* from the variant, idiosyncratic contents of individuals’ I-languages. While you and I undoubtedly possess a different understanding of the word ‘dog’, or ‘jazz’, nonetheless at some level of abstraction and with a certain amount of idealisation, we are ‘speaking/using the same language’.¹⁰⁴ Worries about communication are also misplaced. In practice, communication proceeds by assuming that the interlocutor is identical to oneself, and then determining exceptions, or required modifications, on the fly, by ‘any artifice available’, whether

¹⁰¹ These sorts of worries go back at least to Frege, and his ‘common store of thoughts’.

¹⁰² (Chomsky, 2003, p. 287).

¹⁰³ (Chomsky, 2000, p. 27)

¹⁰⁴ Chomsky suggests that ‘it may be that a kind of public (or interpersonal) character to thought and meaning results from uniformity of initial endowment, which permits only I-languages that are alike in significant respects’, (*Ibid*, p. 33).

pragmatic, semantic, or otherwise.¹⁰⁵ We need not assume that speakers possess exactly the same individual competence in order to communicate.¹⁰⁶

The point might be pressed, however, regarding the determination of *reference* by our I-language meaning representations. Haven't we learned from anti-individualist considerations that reference is often determined by factors beyond our idiolectic understanding? For instance, is it not commonly agreed amongst philosophers that the property we refer to by using 'arthritis', or 'water', is determined in part by social or environmental factors? This consideration threatens not just Chomsky's methodological internalism, but also his ontological internalism. Take two individuals, Gabriel and Evie, who occupy different language communities C1 and C2. In familiar Twin Earth style, we can suppose that Gabriel and Evie are identical with respect to their idiolectic understanding of 'dog', yet intuitively refer to different properties in using the term – *dog* in C1 and *dog** in C2, say. We might presume that competence with respect to 'dog' in C1 would require that we somehow secure reference to *dog*, and likewise in C2 for *dog**. But if the reference of 'dog' in C1 and C2 can vary independently of idiolectic understanding, securing the 'right' reference in each community must involve more than simply possessing such an idiolectic understanding.¹⁰⁷ But then how can competence *consist* solely in that idiolectic understanding? Gabriel's competence with respect to the

¹⁰⁵ (*Ibid*, p. 30). For related discussion, see (Davidson, 1984).

¹⁰⁶ To claim that similarity in idiolects is sufficient for successful communication is *not* (not, not) to claim that a notion of content, or meaning similarity will serve in lieu of content, or meaning identity *for a theory of public language meaning*. See (Fodor, 1998, pp. 30-34); (Fodor & Lepore, 1999). Rather, it is just to make the fairly self-evident point that individuals do in fact communicate, despite having individual differences in their idiolects.

¹⁰⁷ See e.g. (Putnam, 1975).

language of C1 must involve some *other* factors, such as that she is causal-historically embedded in the language community C1, for example.

Being convinced of the cogency of such Twin Earth considerations, I believe that we must find some way of accommodating anti-individualistic aspects of reference. That being said, I am also extremely persuaded by the Chomskyan account of linguistic understanding, and believe that such individualistic thought processes play an inalienable role in reference determination. Combining these two viewpoints requires extreme care, and more space than I possess here. We can clearly see, however, that some possible interpretations of the reference determining relation will rule out the possibility of such a union from first principles. Following Peter Ludlow, we might distinguish a reference relation R^1 , which ‘takes the reference of a term to simply be some direct (perhaps causal) relation between that [public language] term and the world’. If R^1 were adopted, that certainly would imply that reference determination should not advert in any way to I-languages. The mind of the language user is simply subtracted from the reference determining relation. The prospects for such a union would be improved, however, by adopting a more inclusive notion of reference, R^2 , which involves ‘at a minimum a four-place relation that involves the speaker, the expression used, context, and aspects of the world’.¹⁰⁸ R^2 would allow speakers’ I-language entries for the term ‘dog’ to feature in reference determination. Chomsky is in fact perfectly receptive to such a notion of reference. He acknowledges *some* role for non-individualistic factors in securing reference,

¹⁰⁸ (Ludlow, 2003, p. 142)

In the language of a given individual, many words are semantically indeterminate in a special sense: The person will defer to “experts” to sharpen or fix their reference... In the lexicon of this person’s language, the entries for “yaw1” and “ketch” will be specified to the extent of his or her knowledge, with an indication that details are to be filled in by others, an idea that can be made precise in various ways but without going beyond the study of the system of knowledge of language of a particular individual.¹⁰⁹

In principle, then, Chomsky has no objection to the idea that the reference relation essentially adverts to social or environmental features. These facts, however, can be revealed through study of the individual’s I-language. We could thus conceive of I-languages as part of a foundational, subvenient tier in a reference-determining structure that incorporates not only subjects’ thoughts, but also their context and worldly environment.

In summary, insofar as we are interested in the empirical study of the causal-explanatory properties underlying our linguistic competence, Chomsky’s internalist methodology seems the most plausible route to follow. This claim is strongly supported by the resultant explanatory power with respect to various aspects of language acquisition, processing, deficits, and so on. But Chomsky has nothing against theories of reference that incorporate non-individualistic features. Chomsky is simply sceptical that such theories would constitute a viable empirical *discipline*. Given that R² relates not just speakers, but also ‘aspects of the world’, and context, Chomsky fears that the study of this relation would explode into the ‘study of everything’.¹¹⁰ To paraphrase Tyler Burge, if linguists aspire to a *science* of I-

¹⁰⁹ (Chomsky, 1986, p. 18)

¹¹⁰ (Chomsky, 2000, p. 69).

languages, perhaps the best we can hope for is a *study* of reference determination.¹¹¹ We are just engaged in a *different enterprise* when we study our linguistic competence, as against providing a theory of reference. Perhaps a theory of reference can therefore never be completed, or at least never empirically confirmed. Perhaps the reference relation is simply too complex, chaotic, and multifaceted for us to make anything but trivial sense of it.¹¹² Who knows.

Before moving on, let us relate this discussion back to vagueness. As we saw in section 1.2, theories of vagueness have tended to explain the allure of the sorites by appealing to *reference related* features of vague expressions. Supervaluationism adverted to the arbitrariness, or ‘shiftiness’, of the extension precisifications of such expressions. Epistemicism appealed to the unknowability of the reference determining relation. It should now be even clearer, however, why these explanations fail. The reason is that theories of reference, although taking *input* from the study of I-languages, do not advert primarily to causal-explanatory features of our linguistic cognition.

We can consider a more direct argument to the effect that any explanation of vagueness must advert to individualistic aspects of linguistic understanding. It has been common to talk about terms *in a natural language* as being vague or not-vague. The allure of the major premise of the sorites, however, arguably is relative to a speaker’s idiolect. Suppose I erroneously believe that, within my language community, ‘thin’ is agreed to apply only to people with a body mass index (BMI) of

¹¹¹ (Burge, 1989, p. 176). Burge, *arch* anti-individualist, concurs with Chomsky that linguistic methodology should take idiolects as primary objects of study.

¹¹² Philosophical claims such as that ‘meaning determines reference’ are, in general, entirely trivial. Nobody *really* knows how reference gets determined; but the unerring faith is that it does – somehow. See e.g. (Williamson, 1994, pp. 205-12).

less than or equal to 25.¹¹³ My idiolectic understanding of ‘thin’ is that it is a sharply bounded term. In all other respects, I am a master of the term; in fact, I apply it in almost exactly the same way as others in my community. But given my erroneous belief, I would dissent from a major premise of the form, ‘for all n , if a person with BMI n is thin, then a person with BMI $n+0.01$ is thin’. Specifically, I would disbelieve the instance, ‘if a person with BMI 25 is thin, then a person with BMI 25.01 is thin’. Quite generally, a sorites major premise for a term T ceases to have allure for me, if my idiolectic understanding of T is that it is sharply bounded. We can even grant that, in my mouth, ‘thin’ is still vague: while I believe that my community has decided on a sharp boundary for ‘thin’, I remain unsure of this fact, and thus continue to defer to the public language meaning of ‘thin’, which happens to be vague, when speaking to others. ‘Thin’ is thus vague in natural language, and even vague in my mouth, *as I use it*, on account of my deference; but the major premise for ‘thin’ has no allure for me, because of my idiolectic conception of ‘thin’.

Likewise, I might find alluring a major premise constructed for a sharply bounded term T , where my idiolectic understanding of T is that it is vague. Suppose I come to believe that ‘prime number’ is defined as ‘has exactly two distinct natural number divisors *or* is too large to be comprehended by a human mind’. Perhaps I become confused when told that supercomputers are finding prime numbers longer than any human could write out in her lifetime. I mistake this for the claim that it is *sufficient* for being a prime number that no human could comprehend that number. I therefore update my idiolectic understanding of ‘prime number’, to include this

¹¹³ BMI = weight in kg/[height in m]².

sufficiency condition, along with that of having exactly two distinct natural number divisors. However, I now find myself subject to the following sorites series:

1. All numbers with 100 billion digits are prime.
2. For all n , if all numbers with n digits are prime, then all numbers with $n-1$ digits are prime.
3. Therefore, all numbers with 1 digit are prime.

I believe premise 1, (falsely, as it happens), as I think that a 100 billion digit number is almost certainly too long for any human to comprehend. I find premise 2 alluring, because I think that no single digit can mark the point at which a number becomes comprehensible to a human mind. But I disbelieve the conclusion 3, because I know that there are numbers with 1 digit, such as 1, 4, and 6, that are *not* prime. Again, we can assume that due to my lack of certainty and ultimate deference to experts, in my mouth ‘prime number’ actually means what mathematicians say that it means, and is thus a precise term as I use it. But the soritical major premise for ‘prime number’ still has idiolectic allure for me. In sum, if we are adequately to explain the allure of the major premise of the sorites, we must advert to facts about speaker’s idiolects, and their particular lexical entries for ‘thin’, ‘prime number’, as opposed to public language meanings. The allure of a tolerance principle for a term T can detach from the vagueness of T in natural language, given certain deviant idiolectic conceptions of T ’s meaning. This possibility of separation strongly suggests that an explanation of that allure must be individualistic in character.

2.7 Eklund's Inferential Role Semantics

Shortly, I shall proceed to relate the foregoing discussion of cognitive theories of linguistic understanding to Inconsistency Theories about vagueness in developing Cognitive Inconsistentism. Having discussed the problems with Dummett's abilities theory of linguistic understanding, however, we should first briefly address Matti Eklund's views on semantic competence.

As you will recall from section 1.3.2, Eklund proposes an inferential role semantics, on which linguistic understanding consists in accepting, or being disposed to accept as true or valid a certain set of meaning constitutive principles (MCPs). Accepting a principle, however, is an aspect of *performance*, not competence. Moreover, being *disposed* to accept a principle is no better candidate for a state of competence: as we noted in relation to patients in a vegetative state, aspects of linguistic competence remain when subjects are not disposed to do *anything*.

By way of diagnosis, Eklund seemingly has misconstrued the theoretical role of our acceptance of such principles. Eklund takes such acceptance to be *constitutive* of our linguistic understanding: if one failed to accept TP, or even be disposed to accept it, one would be deemed to lack semantic competence with respect to vague terms. Dispositions to accept such principles, however, are just another source of *evidence* for our theory of competence. Acceptance is not *constitutive* of competence; acceptance is *data* for a theory of competence. Eklund's theoretical shortcomings are surely symptomatic of this confusion. Consider an example from the theory of syntactic competence. Nobody would say that our disposition to judge 'the dog ran

quickly', but not 'ran the dog quickly', as grammatical is *constitutive* of our understanding of syntax. Such judgements simply provide evidence for the syntactic rules represented in our competence, (I-language). As with all evidence, the relationship between these judgements and linguistic theory is not a *logical* or *constitutive* one. Rather, we seek a theory that *best explains* such data. It is perfectly conceivable, for instance, that the datum of our accepting 'the dog ran quickly', but not 'ran the dog quickly', *could* have been best explained by a theory of pragmatics, (imagine the lost Gricean Maxim of Subject Salience: *always make the subject of your utterance salient by uttering it first*), rather than syntax. The fact that it is not reflects the success of the syntactic explanation of these linguistic intuitions.

3. Vagueness, Similarity, and Semantic Competence

3.1 Introduction

In this Chapter, I shall adapt the cognitive theory of understanding introduced above to the project of explaining vagueness. I have four main aims in this regard. First, I shall review some classic arguments for the role of tolerance in our linguistic understanding. In seeing how these arguments fail, we shall see that any adequate explanation of vagueness must respect two key features: the *centre-focus* of vague concepts, and the role of *similarity* in sorites susceptibility. My second aim will be to explore the nature of centre-focus through discussion of prototype/exemplar theories of concepts. Third, I shall illuminate the role of similarity in language acquisition and our conceptual classificatory processes. Finally, I shall argue that normal subjects mentally represent a *similarity constraint* with respect to conceptual classification, as part of their semantic competence. This constraint causally explains our linguistic judgements with respect to the major premise of the sorites.

3.2 The Inadequacy of Classic Arguments for Tolerance

Recall from section 1.3.1 Michael Dummett's argument that *observational* terms, such as 'red', must be governed by tolerance principles, such as TP'.¹¹⁴ Dummett argued that, if a term *F* is to be applied purely on the basis our perceiving some property *P*, which is subject to indiscriminable degrees of change ϵ , the conditions for the justified application of *F* could not discriminate as finely as ϵ . That is to say, for any two objects *a* and *b*, differing at most by ϵ with respect to *P*, either both are justifiably

¹¹⁴ TP' was part of our anti-nihilist strategy, stating tolerance in terms of conditions for 'justifiable predication', as opposed to truth.

called *F*, or *neither* is justifiably called *F*. In that earlier section, I suggested that the focus of this argument was too narrow. I now want to demonstrate why.

In two important respects, Dummett's argument underestimates the scope and variety of vagueness in natural language. First, there exist many *non-observational* vague terms, not applied just on the basis of looking. A term such as 'obligatory', as in 'morally obligatory', might be applied to an action, say, only after some considerable thought. Suppose you say that you intend to dedicate 1% of your time to your family. I might judge that act to be obligatory. Moreover, we would normally, if defeasibly, assent to the principle that, if committing $n\%$ of one's time to one's family is obligatory, then committing $[n+0.01]\%$ of one's time to one's family is obligatory. The term 'obligatory' is sorites susceptible, and thus vague, but it is what we might call an 'intellectual' term, applied on the basis of often significant ratiocination, and not just by looking. A complete explanation of vagueness will have to appeal not just to our limited powers of perception, but to deeper features of our language and cognition.

Second, there exist many terms that are vague with respect to *discriminable* changes ϵ , in some property *P*, on the basis of which we apply the term.¹¹⁵ The paradigmatically vague term 'old', while often applied by examining a person's appearance, can be applied on the basis of learning someone's *age*. Moreover, 'old' is certainly vague with respect to changes of one day, a week, or even a month, in a person's age. But it is absurd to claim that such degrees of change are imperceptible,

¹¹⁵ Interestingly, C. L. Hardin, (Hardin, 1988), argues that *any* two light wavelengths are in principle distinguishable, given sufficient pairwise judgements. If Hardin is correct, this would cause even more trouble for Dummett's focus on colour indiscriminability.

or that reported ages differing by up to a month are indiscriminable, to normal subjects. *Look*: '90 years 6 months', '90 years 5 months', etc.

Admittedly, Dummett's argument from observational terms is not the only source of motivation for the role of tolerance in our understanding. However, other 'classic' arguments for tolerance admit of similar problems. Consider the following argument from Crispin Wright, relating to general features of our *use* of vague terms:

We should have no use for a precisely demarcated analogue [of a vague term] in contexts in which the word is typically used. It would, for example, be ridiculous to force the question of obedience to the command, "pour out a heap of sand here", to turn on a count of the grains. Our conception of the conditions which justify calling something a heap is such that the appropriateness of the description will be unaffected by any change which cannot be detected by *casual observation*.¹¹⁶

Again, Wright's focus on changes that 'cannot be detected by casual observation' is too narrow. The term 'heap' is not vague simply with respect to changes of one grain of sand. There clearly exist degrees of change in the size of the heaps, which would be *perfectly* detectable by casual observation, but would nonetheless provide no basis for altering one's application of the term 'heap'. More fundamentally, as we saw with 'intellectual' terms such as 'obligatory', Wright's focus on observation *per se* is also too narrow.

Wright's second argument, concerning the *memorability* of putative conditions for correct usage, fares marginally better. Reflecting upon our learning of colour terms, Wright notes,

¹¹⁶ (Wright, 1976, pp. 157-8). Recall from Chapter 1 that Wright was actually *opposed* to Dummettian Incoherentism.

The [ostensive] character of our basic colour training presupposes the *total memorability* of the distinctions expressed by our basic colour predicates; only if single, unmemorable changes of shade never affect the justice of a particular basic description can the senses of these predicates be explained entirely by methods reliant upon our capacity to remember how things look.¹¹⁷

Call the state of an object with respect to some property P its P -state. Total memorability is the requirement that, for any objects x and y , which differ only by ϵ with respect to P , if that difference is to alter the applicability of a term F as between x and y , then the P -state of x must be memorably distinct from the P -state of y .¹¹⁸ In the case of colours, given that the colour-states of objects differing by $\pm\epsilon$ are ordinarily *not* memorably distinct, differences of that degree cannot affect the applicability of our colour predicates.

Although Wright's memorability argument is presented with respect to terms learned by ostension, and thus observation, it need not be so narrowly construed. There are many non-observational terms, 'obligatory', 'abstract', 'theory', and so on, which admit of non-memorable degrees of difference. Perhaps there is some intuitive difference in the abstractness of the concepts *justice* and *knowledge* – the former being more abstract than the latter, to my mind at least. Assuming that there is a difference between the abstractness-state of *justice* and the abstractness-state of *knowledge*, however, it surely is not a *memorable* one.¹¹⁹ That is,

The argument comes under pressure, however, from terms for which instances admitting of $\pm\epsilon$ difference in their applicatory basis *are* individually

¹¹⁷ (*Ibid*, p. 158).

¹¹⁸ Objects x and y might differ greatly with respect to all of their other properties; but it is just their difference with respect to P that is important to the case.

¹¹⁹ Again, it is the difference in these abstractness-states - *de re*, as it were – that is not memorable.

memorable. There is no reason why a child learning the word ‘old’, for instance, should not be able to remember the distinction between being 55 years and 55 years 1 day in age, presuming that she has the requisite concepts of number, year, day, perhaps time, and so on. There is nothing in *this* distinction, as perhaps there is in the case of small degrees of change in colour shade, or abstractness, which is fundamentally unmemorable.¹²⁰ But if the distinction between being 55 years old and 55 years 1 day old is memorable, then there must be some other reason why *that* distinction could not affect the applicability of our predicate ‘old’.

Several ‘classic’ arguments for tolerance have now been found wanting. Their focus on observational terms, perceptual capacity, indiscriminability, and memorability, is simply too narrow to explain the pervasiveness and intractability of vagueness in ordinary language. A different strategy is required. I shall introduce this strategy in two parts, the first relating to the notion of *similarity*, the second relating to the notion of *centre-focus*.

As we have seen, vagueness *qua* sorites susceptibility does not manifest only with respect to *indiscriminable* differences in the basis for applying a term. Our focus ought rather to be on the notion of *similarity*. The relation of similarity *subsumes* that of perceptual indiscriminability: two objects *a* and *b* are judged indiscriminable with respect to *P* just in case *a* and *b* are above some threshold for similarity in respect of *P*. Employing a notion of similarity, we can thus capture Dummett and Wright’s thought with respect to observational terms. The relation of similarity is *broadier* than that of indiscriminability, however: two objects *a* and *b* might be similar with respect

¹²⁰ There clearly do exist degrees of difference in age, say in the nanoseconds, which *would* be non-memorable; but this is beside the point.

to *P*, whilst being perfectly discriminable with respect to *P*. We can thus capture cases such as ‘old’, where two people of ages 90 years 6 months, and 90 years 5 months, are extremely similar with respect to age, but discriminably different.

Another lesson from the foregoing discussion was that vagueness is not at all limited to observational terms. Importantly, the notion of similarity is also separable from perceptual modalities. I might say that the concepts *justice* and *knowledge* are similar in respect of their abstractness, despite possessing no perceptual basis for this judgement. Equally, I might claim that dedicating 1% of one’s time to one’s family is extremely similar to dedicating 1.1% of one’s time in respect of their being obligatory. Employing a notion of similarity, we can now accommodate vagueness with respect to such ‘intellectual’, as opposed to merely observational, terms.

By focusing our inquiry on the notion of similarity, as opposed to indiscriminability, we can avoid the narrowness which plagues Dummett’s and Wright’s discussions of vagueness. Later in this Chapter, I shall argue that the notion of similarity plays a central role in our cognitive semantic competence. This will pave the way to explaining vagueness *qua* sorites susceptibility. First, however, recall the second element of our new strategy: that of *centre-focus*. It is remarkable that *all* of the classical arguments for tolerance employ what I shall call a *boundary-focused* strategy. The strategy, in sum, is to perform a *reductio ad absurdum* on the supposition that our understanding of vague terms could consist in the knowledge of sharp boundary conditions. To recall, it was argued that sharp boundaries would be a) indiscriminable by ordinary perception, b) impractical, and c) unmemorable. All of these arguments, however, neglect to register one crucial fact: vague terms

typically are neither used nor acquired with reference to peripheral cases, for which the existence of a boundary would be pertinent. As Mark Sainsbury has noted, infants generally do not learn to draw boundaries, or to partition objects cleanly into distinct sets. In contrast, they learn to *classify* the world in clusters around certain points, or ‘magnetic poles’ of conceptual attraction.¹²¹ Any explanation of the role of tolerance in our understanding of vague expressions must acknowledge this fact. Such understanding is not tolerant *because* vague expressions lack sharp boundaries, or knowable sharp boundaries, or *any* kind of boundaries. Boundary cases are simply *irrelevant* to our acquisition and use of most of language.

These two key features, similarity and centre-focus, will serve as the foundation for my explanation of vagueness *qua* sorites susceptibility. In the next section, I shall introduce prototype and exemplar theories of concepts, and suggest that such concepts underpin our semantic competence, or understanding of meaning. In section 3.4, I shall explain the central role of similarity with respect to such conceptual structures. My main argument in that section will be that our conceptual classification is subject to a *similarity constraint*, which causally explains our acceptance of the major premise of the sorites paradox.

3.3 Respecting Centre-Focus: Exemplars and Prototypes

Mark Sainsbury’s suggestions about centre-focus naturally lead us to consider prototype and exemplar theories of concepts. Sainsbury remarks,

¹²¹ (Sainsbury, 1990, p. 258).

It seems that very often a boundaryless concept is acquired on the basis of paradigms. We acquire the concept from the inside, working outwards from the central cases, and locating the central cases of contrary concepts, rather than from the outside, identifying boundaries and moving inwards.¹²²

Such theories of concepts have dominated in cognitive science for over thirty years.

To date, however, there has been little detailed investigation of the connection between these theories, and vagueness.¹²³ I suspect that this is because most theories of vagueness have been set at the level of referential, or model theoretic semantic theory, and not that of individualistic psychological explanation. As argued in Section 2.6.2, however, any fully adequate explanation of vagueness must advert to idiolects, and individualistic meaning representations. As such, the psychological study of concepts seems *precisely* relevant to the task at hand.

Before we pursue this thought, however, two important issues require clarification. The first relates to what we mean by ‘concept’, and the second concerns the supposed relationship between concept possession and semantic competence. As for the first issue, simplifying somewhat, what one means by ‘concept’ might seem to depend on whether one is doing philosophy or cognitive science. Whereas philosophers typically are concerned with concepts *qua* abstract, shareable, public entities, cognitive scientists are concerned with concepts *qua* structures in the mind/brain, we might say. This difference roughly matches that discussed in Section 2.6.2, with respect to Chomsky’s internalist leanings, as against Dummett’s focus on the social aspects of language. As with that debate, however, the dichotomy is

¹²² (*Ibid*, p. 262).

¹²³ In fact, Sainsbury expresses pessimism that prototype theories of concepts could have any bearing on the study of vagueness. He says, somewhat opaquely, that such accounts will always ‘miss the normative features’ of our use of vague terms, (*Ibid*, p. 263). However, the allure of sorites paradoxes is a *psychological* phenomenon, not a normative one.

probably overstated. Jerry Fodor, for instance, suggests that concepts must satisfy a type/token distinction. Concepts *tokens* are mental structures, featuring in individualistic causal explanations and such like.¹²⁴ Concept *types* are public and shareable, possessed by numerous distinct subjects.¹²⁵ No cognitive scientist would deny that concepts such as DOG, CHAIR, TABLE, are possessed by numerous subjects.¹²⁶ In practice, as Chomsky notes, concept tokens – what people have in their head – will be largely uninteresting for the purposes of naturalistic inquiry. All theorising and experimentation will treat subjects as if they are a ‘normal case’ of concept possession, tacitly presuming that their concepts are of the same type. Each data point, while strictly speaking pertaining to an individual concept token, will be taken as data for a theory of concept types. With any luck, the theoretical aspirations, if not the methodologies, of philosophers and cognitive scientists will thus coincide.

What other aspects of concepts are important for our purposes? Again following Jerry Fodor, we can say that concepts are *categories*, in the sense that objects in the world fall under them. The term ‘category’ is widely used by linguists and psychologists. Here is a representative claim:

Categories are groups of distinct abstract or concrete items that the cognitive system treats as equivalent for some purpose. Maintaining and using equivalence classes involves mental representations that encode key aspects about [*sic.*] category members.¹²⁷

¹²⁴ Calling such mental structures ‘concept *tokens*’ is not meant to undermine the deployment of such structures in multiple distinct thoughts. Perhaps we should operate with a tripartite distinction between concept types, mental tokens of those concepts, and *deployments* of those tokens.

¹²⁵ (Fodor, 1998, pp. 23-29).

¹²⁶ I am implicitly ruling out strongly holistic theories of concepts, on which concept types are individuated with respect to all inferential connections borne to other concepts. Given the low probability that any two subjects possess exactly the same inferential connections between their concepts, there is an equally low probability that any two subjects would possess tokens of the same concept types.

¹²⁷ (Markman & Ross, 2003, p. 593).

Concepts, then, equate to the ‘mental representations that encode key aspects [of] category members’.¹²⁸ I shall thus assume that possessing a concept *F* consists in possessing a mental representation encoding information about the members of *F*. Shortly, I shall offer prototype/exemplar theories as a means of explicating the nature of such mental representations.

There will be many diagnostics for the possession of a concept/category, chief amongst them the ability to *classify* objects as members of the same category. We should be careful to distinguish, however, between satisfying behavioural criteria for classifying an object using the term *F*, and acquiring the *category*, or *concept F*. Classification abilities are not logically sufficient for concept possession. As noted by Chin-Parker and Ross, those in the science community often fail to recognise this point.¹²⁹ This again is simply to respect the Chomskyan distinction between *performance* and *competence*: performance abilities are but one (inconclusive) source of evidence for competence.

Turning now to the second issue, in what ways does the Chomskyan study of semantic competence coincide with cognitive theories of concepts?¹³⁰ It is standard in both philosophy and cognitive science to assume a very close association between the meaning of an expression *F*, and the concept *F*. In philosophy, for instance, one

¹²⁸ Confusingly, the term ‘category’ is often used to refer to *object* categories, or *sortals*, such as CHAIR, BOTTLE, DOG. I shall take ‘category’ to include those concepts expressed by predicates, ‘tall’, ‘bald’, ‘red’, but also verbs, ‘run’, ‘talk’, ‘think’, and more.

¹²⁹ For instance, (Chin-Parker & Ross, 2004, p. 217) note that it is common in the psychology literature to assume that a successful classification of *x* as a *stapler*, for instance, is sufficient to conclude that the child possesses the category STAPLER. But this is at best *evidence* for such possession.

¹³⁰ This question is complicated somewhat by the fact that Chomsky, on the whole, has focused more on syntax than semantics. Furthermore, much of what philosophers would call ‘semantics’, he calls ‘syntax’. By ‘Chomskyan’, I therefore signify those theorists who have adopted Chomsky’s cognitive views on linguistic competence, but have focused on mentally represented conceptual structures, rather than syntax.

often claims that the concept CHAIR just is what is expressed by the term ‘chair’; the concept CHAIR is what ‘chair’ *means*. In cognitive science, a similar relationship is envisaged between one’s understanding the meaning of a term such as ‘chair’, and possessing the concept CHAIR. Here is an example from Paul Bloom:

An adequate theory of how children learn the meanings of words such as “dogs” and “giving” requires some account of what it is to possess the corresponding concepts of DOGS and GIVING.¹³¹

On this reading, semantic competence – understanding the meaning of an expression – seems to *entail* conceptual competence – possessing the concept for that expression.¹³² *Modulo* certain complicating factors, which will become clear as we proceed, I shall assume this entailment in what follows.

There is considerable empirical support for this proposed relationship. To mention but one instance, a study conducted by Alison Gopnik and Andrew Meltzoff found a significant correlation between the acquisition of sophisticated conceptual categorisation and linguistic abilities, specifically the so-called ‘naming explosion’ that occurs around 18 months.¹³³ As noted in Section 2.3, the naming explosion marks a remarkably consistent stage in linguistic development, during which children undergo a significant increase in vocabulary acquisition, at anything from around 45 to 63 words per week,¹³⁴ with a preponderance of nouns.¹³⁵ As for conceptual categorisation, while children as young as 9-10 months show some

¹³¹ (Bloom P. , 1993, p. 11).

¹³² ‘Entails’, because the relationship between competence with respect to ‘chair’ and possessing the concept CHAIR is one way: a French speaker plausibly could possess that concept, but not understand the term ‘chair’.

¹³³ (Gopnik & Meltzoff, 1987).

¹³⁴ (Bloom L. , 1973); (Corrigan, 1978).

¹³⁵ (Dromi, 1987); (Goldfield & Reznick, 1990).

rudimentary discrimination between distinct categories of object, at around 18 months children develop the more sophisticated ability to sort objects into piles corresponding to their categories. No child included in Gopnik and Meltzoff's experiments underwent the naming explosion *prior* to reaching this sophisticated stage of categorisation. Of twelve children tested, five experienced the shift in categorisation abilities and the naming explosion at the very same time. The evidence pointed towards a 'specific developmental link' between these cognitive and linguistic capacities, which in turn provides support for our assimilation of the acquisition of concepts, or mental categories, and lexical meanings.¹³⁶

The situation is complicated, however, by the fact that the acquisition of concepts/categories seems not only to subserve our acquisition of language, and hence semantic competence, but also a range of other cognitive faculties. Jerry Fodor goes so far as to say that 'most of cognitive psychology, including the psychology of memory, perception, and reasoning, is about how we apply concepts'.¹³⁷ This need not undermine the equation of conceptual and semantic competence, however. Our semantic competence – *qua* concept possession – would simply be engaged in other areas of cognition, such as in perceptual classification, prediction and inference, for instance. I see no particular reason to doubt that this might be the case in humans.

Let us now move on to discuss prototype and exemplar theories of concepts. Broadly construed, these theories claim that concepts represent a significant amount of information about their *central* or *focal* instances. This claim contrasts with a

¹³⁶ Other factors have been cited as influencing the onset of the naming explosion: the development of object permanence, (Corrigan, 1978); object concepts, (Nelson & Lucariello, 1985), and acquiring an understanding of the association between symbols and entities in the world, or a 'naming insight', (McShane, 1979).

¹³⁷ (Fodor, 1998, p. 24).

classical theory of concepts, for instance, on which subjects represent something like definitional, necessary and sufficient criteria for concept membership. I shall assume a degree of familiarity with the history of this classical-to-prototype progression in theories of concepts.¹³⁸ What sort of information is taken to be represented about these central, focal instances of concepts? On the *exemplar* theory, concepts represent specific information, or ‘memory traces’, about *particular examples* of a concept.¹³⁹ As new examples are encountered, they are added to memory, such that one’s representation will, over time, come to contain traces of many examples of the concept. As a way of modelling exemplar theory, a perceived stimulus can be represented by an ordered n -tuple, each value corresponding to a psychological dimension, feature or property, of that stimulus.¹⁴⁰ The delimitation, or selection of features to be represented, is plausibly constrained by attentional tendencies. A simple object such as an apple possesses innumerable properties, not all of which are taken into consideration in category formation: being on earth, consisting of n particles, and so on. An analysis as to how attention functions in such cases is beyond my remit, however.¹⁴¹

On *prototype* theory, in contrast, concepts do not represent information about, or traces of, particular exemplary instances, but rather some abstracted ‘summary’ of

¹³⁸ (Wittgenstein, 2001, §66) is usually credited with the first insight on prototype structure, and family resemblances, although his project was not that of illuminating *cognitive* representations. Eleanor Rosch, (Rosch, 1975a; 1975b), developed this insight within the confines of psychological theory.

¹³⁹ ‘Particular examples’ must be construed widely enough to accommodate uninstantiated concepts, such as WITCH or UNICORN, and also ‘exemplar descriptions’, as in cases where we do not acquire a concept by acquaintance with an instance, but rather through description.

¹⁴⁰ This model would presuppose that subjects possess a prior conception of each of these features. This would naturally suggest a hierarchical model of concept acquisition, with basic perceptual properties acquired first, and those properties deployed in increasingly complicated and abstract representations. For prominent versions of the exemplar model see (Nosofsky, 1986), (Kruschke, 1992), (Sheperd, 1987), (Lee & Navarro, 2002).

¹⁴¹ For discussion of the role of attention in category formation, see (Sheperd, Hovland, & Jenkins, 1961) and (Maddox, 2002). (Murphy & Medin, 1985), (Biederman, 1987), and (Yuille & Ullman, 1990) explore the role of perceptual systems in delimiting the range of salient properties in categorisation.

the features of those instances.¹⁴² This might be in the form of a central tendency, or mean average of each of the psychological dimensions associated with particular category exemplars.¹⁴³ It could equally be a representation of the modal, or most frequently encountered, features of exemplars. Another option would be an idealised representation, which is ‘maximally distinctive from other categories’.¹⁴⁴ As in exemplar theory, prototype summary representations are updated with new information as additional instances of the concept are encountered.

This cursory description of exemplar and prototype theories is intended merely as an overview. I shall not digress on the debate between exemplar and prototype theorists, as to which provides the best model of our conceptual structure.¹⁴⁵ The crucial aspect of these theories, for our purposes, is their *centre-focus*. There has been much experimental research into the pervasiveness of such centre-focus, known in the lingo as a *typicality effect*.¹⁴⁶ Typicality effects arise in three distinct ways. First, as Rosch demonstrated with respect to colour categories, certain so-called ‘focal colours’, or ‘reference stimuli’, are

better remembered and earlier learned than nonfocal colours, and that colour categories structured so that focal colours are the physical centre of the category are more easily learned than categories structured in other ways.¹⁴⁷

¹⁴² In the words of John Kruschke, (Kruschke, 2005), the representation contains a ‘global summary’, rather than ‘piecemeal information’, about category instances.

¹⁴³ This presumably presupposes at least a *transitional* exemplar representation, from which prototype representations are generated.

¹⁴⁴ (Kruschke, 2005). For examples see (Lynch, Coley, & Medin, 2000), (Palmeri & Nosofsky, 2001).

¹⁴⁵ Against exemplars, we have e.g. (Blair & Homa, 2001), (Minda & Smith, 2001). Against prototypes, we have e.g. (Ashby & Maddox, 1993), (Busemeyer & Myung, 1988). (Storms, DeBoeck, & Ruts, 2001) suggest theoretical pluralism. Perhaps one is not faced with a hard choice between exemplar and prototype theories.

¹⁴⁶ A related distinction in theories of learning is between *rule-based* and *inductive* learning. Whereas rule-based learning typically involves acquiring a representation of the *boundary* conditions of a concept, inductive learning normally involves the representation of a relatively *central case* of the concept, from which the subject tacitly generalises. See (Kruschke, 2005) for a summary.

¹⁴⁷ (Rosch, 1975a, p. 533).

This type of typicality effect may be limited to perceptually, or perhaps physiologically, basic categories such as colours. A second effect, however, arises with respect to many other ordinary categories such as FURNITURE, FRUIT, SPORT, and so on. The effect is that certain types of object or activity are consistently given a higher ‘rating’ as to whether they are ‘good examples’ of that category. Such graded ratings for concept membership are ubiquitous.¹⁴⁸ Moreover, there exists a high degree of commonality across subjects with respect to their typicality judgements: prototypes tend not to be overly idiosyncratic.¹⁴⁹ The third type of typicality effect is that verification times for prototype-involving sentences, such as ‘a robin is a bird’, are generally significantly lower than for non-prototypical cases, such as ‘an ostrich is a bird’.¹⁵⁰ These three typicality effects strongly suggest ubiquitous prototype/exemplar structure in our mental categories.¹⁵¹

Shortly, I shall move on to discuss a major problem with prototype/exemplar theories: *compositionality*. First, however, we should consider the relationship between prototype/exemplar theories and semantic theory. The discovery of typicality effects with respect to our cognitive concepts has frequently led prototype/exemplar theorists to claim that ‘concepts have graded structure’, or ‘concepts have degrees of membership’. This implies that *reference* determination, in which our individualistic semantic competence plays a significant role, (see Section 2.6.2), will yield graded degrees of membership, say, in the *extensions* of concepts.

¹⁴⁸ For a summary, see (Medin & Smith, 1984); (Mervis & Rosch, 1981).

¹⁴⁹ (Malt & Smith, 1982); (Rosch & Mervis, 1975).

¹⁵⁰ See (Rosch E. , 1975b) and (Smith, Shoben, & Rips, 1974).

¹⁵¹ Similar effects have been studied under the title of ‘Categorical Perception’: see e.g. (Bornstein & Monroe, 1980).

Such partisans have readily grafted a degree theoretic semantics, or fuzzy logic, onto prototype/exemplar theory.¹⁵² However, this seems to me a *huge* mistake, for two reasons.

As a first reason, I very much doubt whether we should allow psychological theories to constrain matters of *logic*.¹⁵³ One's choice of logic, it seems, should be determined at least in part by matters of simplicity, elegance, scope of theoretical application, and utility in reasoning.¹⁵⁴ Perhaps fuzzy logic will prove to be our best logic on these criteria. I doubt it.¹⁵⁵ But if our logic is not fuzzy, then there is no reason to think that reference determination will yield graded degrees of membership in the extensions of concepts. For instance, suppose that we possessed overwhelmingly independent reason to adopt classical logic. In that case, in semantic theory, we would take our reference determining relation to respect the law of excluded middle, amongst other things.¹⁵⁶ This would undoubtedly yield a more complicated relationship between individual graded responses to concept membership, and reference determination. Something like Eklund's principles of respect and approximation, (section 1.3.2), however, might serve an important purpose in this endeavour.

The second reason is that typicality effects occur with respect to seemingly non-graded concepts, such as ODD NUMBER, GRANDMOTHER, and TIGER. As

¹⁵² See e.g. (Fuhrmann, 1988).

¹⁵³ Psychologism in linguistics: good. Psychologism in logic: bad. Frege was half right.

¹⁵⁴ Relevantly, Tim Williamson, (Williamson, 1997, p. 218), suggests that one should hold 'one's logic fixed, to discipline one's philosophical thinking'.

¹⁵⁵ As Williamson notes, (Williamson, 1994, pp. 135-38), fuzzy logic and degree theoretic notions of truth lead to disastrous consequences for the truth-functionality of the logical connectives.

¹⁵⁶ Williamson thus adopts a *default principle* for reference determination: 'If there are not enough facts about use [and thought] and the environment to determine [a case] to belong [in] the [extension] of [a term], then that very shortfall is enough to determine it not to belong', (Williamson, 1999, p. 509).

such, typicality effects with respect to a concept *F* are insufficient to establish that there are ‘degrees of membership’ in the extension of *F*. The concept ODD NUMBER, for instance, permits of a perfectly precise definition, namely, *integer not divisible by two without remainder*. But interestingly, the number 3 is generally classed as a ‘better’ odd number than 19, and sentence verification times are quicker for ‘7 is an odd number’ than for ‘23 is an odd number’. Subjects are judged more typical of the concept GRANDMOTHER, if they possess wrinkles, grey hair, knit socks, and so on. Yet GRANDMOTHER also admits of a definition, namely *mother of a parent*, which precisely determines membership in the concept. Instances judged typical of the concept TIGER possess striped fur, live in jungles, and occasionally lunch on hapless villagers. Mainstream philosophical and linguistic opinion, however, has it that TIGER is a natural kind concept, referring to some scientifically isolable genetic trait. Whatever these typicality effects show, it is surely *not* that such concepts admit of degrees of membership in their extensions.

We might wonder, however, what *is* shown by such typicality effects? Some theorists have argued that, if typicality effects can occur with respect to sharply defined concepts, then such effects simply provide *no evidence whatever* as to the structure of our category representations.¹⁵⁷ Sharply defined concepts *clearly* lack graded membership, they claim, so typicality effects cannot pertain to *their* structure. Ergo, typicality effects cannot pertain to *any* cognitive conceptual structure. Armstrong, Gleitman, and Gleitman suggest that typicality effects pertain only to a ‘mentally stored *identification function* used to make quick sorts of things, scenes, and

¹⁵⁷ This finding is cited by (Sainsbury, 1990, p. 263) as a reason against pursuing prototype theory.

events in the world.’¹⁵⁸ In the case of GRANDMOTHER, for instance, the common identifying features *having wrinkles, grey hair, knitting socks*, are not part of conceptual structure *per se*, but merely associated with the concept’s identification function.

Armstrong et al. distinguish the identification function from what they call a ‘*categorical description*’, or ‘*conceptual core*’, which encodes conditions precisely determining concept membership. In the case of GRANDMOTHER, then, the conceptual core would consist in the (perhaps not fully articulated) representation, *mother of a parent*.

The dichotomy proposed by Armstrong et al. is deeply unappealing, however. As should be obvious, almost all concepts lack a categorical description.¹⁵⁹ Concepts such as GRANDMOTHER and ODD NUMBER are the exception, rather than the rule. Concepts lacking such categorical descriptions, such as RED, BALD, RICH, NEAR, DEMOCRACY, JUSTICE, BEAUTY, are far more common in ordinary thought and language. On the account of Armstrong et al, however, such concepts would lack a conceptual core. A more plausible account is that such concepts *consist* in the mental representation of the features associated with the identification function.¹⁶⁰ This is just what is claimed by prototype/exemplar theory.

Armstrong and colleagues object to this that many non-definitional concepts are consistently applied to instances completely lacking any ordinary identifying features for that concept:

¹⁵⁸ (Armstrong, Gleitman, & Gleitman, 1983, p. 292). The distinction between identification function/procedure and conceptual core stems from (Osherson & Smith, 1981).

¹⁵⁹ (McNamara & Sternberg, 1983) note that people often *believe* that concepts have defining features; but arguments in e.g. (Smith & Medin, 1981) suggest that such definitions are rarely forthcoming.

¹⁶⁰ (Armstrong, Gleitman, & Gleitman, 1983, p. 296) consider this reply.

Shouldn't subjects throw in the conceptual towel when asked whether a bird is still a bird even when plucked (or dewinged, or debeaked, or whatever) or a tiger still a tiger without its stripes? But on the contrary, subjects seem to be quite sanguine about having these identification features... removed, and even for concepts that allegedly consist of nothing else.¹⁶¹

Whatever rhetorical force is gained by these examples, however, surely derives from the fact that BIRD and TIGER are natural kind concepts, which although non-definitional, plausibly refer to scientifically isolable genetic traits. The examples therefore really do no more than reinforce familiar Twin Earth considerations. As I hope will be familiar, our intuitions about the reference of WATER seem to track the *kind* water, i.e. H₂O, as opposed to common identifying features of water, such as being clear, potable, found in streams, and so on. Something similar could be said for BIRD and TIGER. But we simply do not encounter similar intuitions with respect to concepts like BALD, RICH, and RED. In such cases, removing, or inverting, all common identifying features normally *would* change our intuitions about reference.¹⁶² If we take away all of a rich person's money, we normally deem them poor; if we increase a bald person's assignment of hair, then he typically ceases to be classed as bald; if we lighten a shade of red, we ordinarily no longer consider it red.¹⁶³ In such cases, typicality effects are best taken as evidence for the representation of prototypical features in our cognitive conceptual structures.

¹⁶¹ (Armstrong, Gleitman, & Gleitman, 1983, p. 269).

¹⁶² Although as I shall discuss below, such features are neither necessary nor sufficient for reference determination. There is no entailment from identifying features to the reference of a concept: such features will go into the subvenient tier of our reference determining facts along with context, aspects of the world, and so on.

¹⁶³ In the terminology of (Rey, 1983, p. 241), these concepts are *epistemologically taxonomic*, as opposed to *metaphysically taxonomic*, roughly in the sense that there are no scientific or naturally occurring identity conditions for the concept. This is not to say that concepts such as BALD lack a mind-independent basis for their application, as in the physical state of a person's hair and scalp: it is just to say that that basis forms no natural kind.

In any case, there is a less destructive, and more satisfying, means of interpreting typicality effects with respect to sharply bounded concepts. A more satisfying explanation, I believe, is that the *acquisition* of most concepts involves the initial formation of a prototype representation, consisting in a set of identifying features. In the case of GRANDMOTHER, for instance, infants would very rarely acquire its definition in the first instance. The concept would likely be acquired through interaction with, say, a *grandmother*, who may well have grey hair and wrinkles. Something similar is presumably true even of mathematical concepts, such as ODD NUMBER. I imagine it would be futile to teach a child the concept ODD NUMBER via the definition, *integer not divisible by two without remainder*. In fact, I doubt that most children would *ever* be taught this definition. The situation is rather that children encounter various ‘easy’ examples of odd numbers – 1, 3, 5 – and with age, and sophistication, gradually refine their concept, eventually converging on a (perhaps unarticulated) representation of a definition.¹⁶⁴ In such cases, our conceptual representation may develop a definitional structure, but retain its developmentally basic prototypical associations. This would explain the persistence of typicality effects with respect to such sharply bounded concepts. Such effects pertain to pervasive and developmentally basic aspects of our categorisation and conceptual structure. The determination of reference, however, is not achieved by

¹⁶⁴ The case might be complicated with respect to natural kind concepts. Research by Susan Carey, (Carey, 1985), indicates that infants develop a very early sense that types of animal, for instance, possess some unseen, unknown unifying biological basis. This is taken as evidence for the child’s knowledge of ‘folk biology’. See also (Keil, 1987; Keil, 1989). Nonetheless, as Robert Goldstone notes, acquiring the biological category DOG, for instance, may be rather like working towards a scientific theory of dogs: ‘information [about such features] may be sufficient to group dogs together in a common category. Once this category has been created, further abstract commonalities can be discovered. Scientists investigated the genetic similarity of dogs *because* of their more superficial similarities’, (Goldstone, 1994, p. 146). Likewise, perhaps the child’s concept DOG requires some initial prototypical grouping, in order to prime the child’s folk biological sense that a common kind underlies that grouping.

individualistic cognitive capacities alone: mathematical, causal-historical, and biological aspects of the world can constrain category extensions.

A different objection might be pressed. If the features associated with our prototype representation for RICH are to explain typicality effects, they surely must include such things as *wearing smart clothes, living in a mansion, travelling by helicopter*, and so on. Individuals possessing such features would normally be judged ‘better’ examples of the concept RICH than scruffy, ascetic, frugal billionaires. But such features surely cannot determine the reference of RICH: a scruffy billionaire is no less of a rich person for being sartorially challenged. The objection is well taken. The prototype/exemplar theorist, however, is not in the business of providing necessary and sufficient conditions on the reference of RICH. Even taking away all of a rich person’s money *might* not make them poor: there no doubt exist many cash-poor, asset-rich folk in the world. We could even imagine such cash-poor people remaining rich when stripped of their assets, perhaps through some elaborate system of government subsidies. Whatever role prototypical identifying features play in reference determination, it will not be one of necessitating classification. Reference determination will be a much messier matter, maximising the correctness of speakers’ intuitive classifications where possible, whilst respecting competing standards or constraints on reference. All the same, if our best psychological explanation of speaker’s conceptual judgements appeals to prototypical identifying features, then that consideration speaks highly in favour of ascribing prototype conceptual structures to them.

3.3.1 Prototypes and Compositionality

No discussion of prototype/exemplar theory would be complete without a mention of the problem of compositionality. Jerry Fodor describes the situation as follows:

Whatever conceptual content is, compositionality requires that complex concepts inherit their contents from those of their constituents... Accordingly, whatever is *not* inherited from its constituents by a complex concept is *ipso facto* not the content of that concept.¹⁶⁵

Compositionality – the contribution of constituent contents to the content of a complex concept – is, I assume, a non-negotiable requirement for any candidate theory of concepts. But now consider a complex concept such as PET FISH. A prototypical pet is probably a dog or cat, and a prototypical fish is something like a trout. But a prototypical *pet fish* is a goldfish, which seems *prima facie* not to be derivable from the combination of typical-pet and typical-fish properties. Ergo, if Fodor's claim is correct, typical-pet and typical-fish properties are not contents of the concepts PET and FISH respectively.¹⁶⁶ Similar problem cases abound:

CARNIVOROUS PLANT, GARDEN FURNITURE, CHILD ACTOR, and so on.

The compositionality problem is unfortunately too large to tackle here. There are various defensive strategies available, however; the prototype/exemplar theorist need not abandon hope. One strategy is to note that the PET FISH problem, as posed, seems to presuppose an exemplar-type view of conceptual representations, focusing on particular animals – cats, dogs, trout etc. – deemed typical of the category. Edward Smith and Daniel Osherson have explored the possibility of

¹⁶⁵ (Fodor, 1998, p. 100).

¹⁶⁶ Perhaps the first to mention this problem were (Osherson & Smith, 1982).

accommodating compositionality within the prototype model of abstracted, or averaged feature sets.¹⁶⁷ Hans Kamp and Barbara Partee consider evaluating typicality for complex concepts relative to a comparison class: a goldfish is perhaps not a good example of a pet *tout court*, but relative to the class of fish, a goldfish is as good a pet as one could hope to find.¹⁶⁸ A third strategy employed by Philip Robbins is to restrict the scope of the compositionality requirement.¹⁶⁹

Perhaps the most promising strategy, however, is to grant that a) concepts are not compositional *qua* prototypes, and b) concepts must be compositional *qua* something else, but maintain that c) this does not entail that conceptual structure is not prototypical. The tension between b) and c) is resolved by advocating a ‘dual model’ view of conceptual structure, on which prototype features play a central role in classification, but concepts only contribute a minimal, homophonic content in combination.¹⁷⁰ We might thus say, quite minimally, that the complex PET FISH is assigned a homophonic interpretation *pet fish*, in much the same way as homophonic theories of meaning would claim simply that ‘pet fish’ means *pet fish*. It takes a particularly austere or modest philosopher to claim that such homophonic interpretations could serve fully to explicate our understanding of ‘pet fish’, however. Although compositionality is rightly stressed as a non-negotiable requirement for a theory of concepts, it is far from the *only* important aspect of semantic competence. Prototypes or exemplars must remain a central feature of any theory of concepts powerful enough to explain the consistent typicality effects

¹⁶⁷ (Smith & Osherson, 1984).

¹⁶⁸ (Kamp & Partee, 1995). Jerry Fodor, (Fodor, 1998, pp. 102-3), takes issue with this view, and that of (Smith & Osherson, 1984).

¹⁶⁹ (Robbins, 2002).

¹⁷⁰ For a similar suggestion, see Lance Rips’s ‘Representations-of’ categories, which consist of ‘word-like symbols that can combine according to syntactic rules’, (Rips, 1995, p83, fn.4).

observed in language acquisition, classification tasks, and language processing. One possible line of future inquiry would be the idea that, once the compositional form of PET FISH has been determined, something close to the strategy of Smith and Osherson, or Kamp and Partee, is applicable to such complexes.¹⁷¹ For example, perhaps the prototype structure of PET FISH is determined by evaluating the typicality of PET relative to the comparison class of fish. This problem will be an essential focus for any future research.

3.4 The Similarity Constraint and Vagueness

In the previous section, I discussed prototype/exemplar theories of concepts, which respect one key feature of vague concepts, namely *centre-focus*. The other key feature was that vagueness *qua* sorites susceptibility manifests itself with respect to degrees of *similarity* ε , in some property P pertinent to the application of a term F . The remainder of the thesis will be focused on the notion of similarity. I shall argue that the role of similarity in sorites paradoxes corresponds to a fundamental aspect of human cognition and conceptual classification. In summary, I shall argue that humans represent a *similarity constraint* in their cognitive architecture, which is causally and explanatorily implicated in language acquisition and conceptual classificatory processes.¹⁷² Taking C as a ‘classification’ operator, let us state the similarity constraint as follows:

¹⁷¹ This avoids the problem, noted by (Fodor, 1998, p. 104), that identifying the correct comparison class for PET FISH, in the manner suggested by (Kamp & Partee, 1995), requires a prior grasp of its compositional structure.

¹⁷² Delia Graff describes a *contextualist* similarity constraint, which states that ‘whatever standard is in use for a vague expression, anything that is *saliently* similar, in the relevant respect, to something that meets the standard itself meets the standard; anything saliently similar to something that fails to meet the standard itself fails to

$$(SIM) \quad \forall x, \forall y, \text{ if } |P(x) - P(y)| \leq \epsilon, \text{ then } C[Fx \leftrightarrow Fy]$$

SIM states that, for any objects x and y , if x and y are within some bound of similarity with respect to P ,¹⁷³ then classify such that x is F iff y is F . SIM is a *soft* constraint, in a sense to be explicated below. The ‘classification’ operator must be understood *psychologically*: SIM is claimed to be mentally represented in our cognitive architecture, and operative on our classificatory processes. Adopting the metaphor of homunculi for a moment, C issues cognitive *directives* or *imperatives* as to how to deploy one’s concepts. SIM thus ‘directs’ our conceptual processes to classify in such a way as to ensure that x is F iff y is F , if x and y are similar in respect of P . The calculation of similarity as stated, $|P(x) - P(y)|$, is a simplification: there will be a complicated underlying similarity function, also represented in our cognitive architecture, determining the extent to which x and y are similar. I shall not go into the details of the candidates for this function.¹⁷⁴

Claiming that the similarity constraint issues cognitive ‘directives’ or ‘imperatives’ should not be taken as meaning that SIM is a procedural rule, applied strictly in conceptual classification. In contrast, the notion of constraint here is inspired by Optimality Theory (OT) in linguistic theory.¹⁷⁵ OT holds that linguistic

meet the standard’, (Graff, 2000, p. 14). She rejects the idea that the constraint has any special semantic status, or that it is cognitively represented by speakers.

¹⁷³ The ‘bound’ of similarity operating in the constraint will, I assume, be determined empirically.

¹⁷⁴ For a summary of various formal models for similarity judgements, see (Kruschke, 2005).

¹⁷⁵ OT is now a popular approach to phonology, morphology and syntax. The extension of OT to semantic theory has not been fully explored. Theories billed as ‘OT semantics’ in any case often seem to be in the business of specifying *pragmatic* constraints on competing semantic interpretations of sentences – e.g. (Hendriks & de Hoop, 2001). The extension of OT to cognitive conceptual classification appears to be a new line of inquiry.

interpretation, whether syntactic, phonological, or semantic, is not subject to strict rules, but rather to a number of constraints, which can come into conflict.

Constraints are strictly *ranked*, such that for any two constraints C_1 and C_2 , either C_1 outranks C_2 , or vice versa. The OT model of interpretation has two crucial elements: a ‘Generator’ produces a list of all candidate interpretations, and an ‘Evaluator’ is a function which determines which of these candidates is *optimal*, or *harmonic*, that is, which best satisfies or minimally violates the full set of constraints.¹⁷⁶ SIM is taken to be a *soft* constraint, in the sense that, if competing factors outrank or outweigh SIM, then we will not classify such that x is F iff y is F . In the casual phrase of Petra Hendricks and Helen de Hoop, constraints are of the form, ‘*if there is..., use it, unless...'*. Applying this rough heuristic to SIM, we can interpret the constraint as being that, *if there is* a high degree of similarity with respect to P between x and y , then (*use that basis to*) classify such that x is F iff y is F . The ‘*unless...*’ clause is simply a placeholder for other competing constraints, which have not been specified in SIM. SIM is thus *violable*, in theory, and in practice.

What other constraints operate on our conceptual classificatory processes?

Taking Pr^F as an operator ‘prototype of F ’, consider the following *prototypicality* constraint:

$$(PRO) \quad \forall x, \forall y, \text{ if } |P(x) - P(y)| \leq \epsilon \text{ and } Pr^F x, \text{ then } C[Fy]$$

¹⁷⁶ (Prince & Smolensky, 1993). The moniker ‘harmonic grammar’ has come to be used for OT theories of competence. OT has its roots in connectionist, or parallel distributed processing, theories of cognitive architecture. I do not want to discuss these historical associations; I am interested simply in the theoretical paradigm.

Allowing PRO to quantify over mental representations, as well as objects, the constraint would state that *if there is* a high degree of similarity with respect to P between x and y , and x is a prototype of F , then classify such that y is F . It is extremely plausible, given our discussion of prototype/exemplar theories of concepts, that PRO would play an important role in classification.¹⁷⁷ We shall return to this putative constraint in the next section 3.4.1.

My proposal is that SIM causally explains our linguistic judgements, or intuitions, with respect to the major premise of the sorites. When presented with this premise, we evaluate its plausibility relative to our conceptual classificatory mechanisms. Our mental representation of SIM, under normal circumstances, would cause the judgement that any two pairwise similar objects should be classified such that either both are F , or both are *not-F*. The causal explanatory notion of a conceptual *constraint*, as opposed to a semantic rule, highlights some fundamental differences between my approach, and that of Michael Dummett. Dummett, as we saw in section 1.3.1, situates tolerance principles at the level of *semantic rules*, dictating the conditions governing the correct use of language. Crispin Wright characterises Dummett's view as follows:

For any particular expression in the language, a set of such rules is completely determinant of when the expression is used correctly... Clearly, however, the feasibility of such a picture requires that the rules associated with an expression... be consistent. For if they issue in conflicting verdicts upon the correctness of a particular application of the expression, it cannot be explained just by appeal to the rule why we agree that the application is e.g. correct.¹⁷⁸

¹⁷⁷ It might even be possible to derive PRO from SIM, with the added assumption that $Pr^F x \rightarrow C[Fy]$. I am currently unaware, however, whether this derivation corresponds to any actual aspect of cognitive development.

¹⁷⁸ (Wright, 1976, p. 171).

The presumption of such theories is that semantic rules are *strict, non-violable, and non-conflicting*: any use of language that failed to comply with a tolerance principle would be deemed indicative of a failure of understanding. Merely *selective* implementation of these rules, or outright *failure* to implement them, would be deemed a failure of understanding. Of course, if tolerance principles were semantic rules, then *all* uses of language would in a sense be failures on our part, as tolerance principles are not satisfiable by any consistent, non-trivial pattern of use. Within the cognitive framework I have elucidated, however, the possibility of conflicting, possibly inconsistent, constraints on conceptual classification can be accommodated. My view has the potential to explain the seemingly tolerant aspects of our concepts and meanings, without succumbing to radical incoherence.

Moving on, my argument for the similarity constraint will proceed in three stages. First, I shall motivate the role of similarity in language acquisition and semantic competence by appealing to Poverty of the Stimulus considerations. Second, I shall argue that the similarity constraint plays a distinctive role in explaining our productive capacities with concepts and language. Third, I shall produce evidence that the similarity constraint SIM can be differentiated from our prototypicality constraint PRO. Although often overruled by outranking constraints such as PRO, when prototype representations are unavailable in naming tasks, SIM plays a prominent classificatory role.

3.4.1 Similarity and Poverty of the Stimulus

In this section, I shall argue that the notion of similarity plays a central role in the related areas of language acquisition, semantic competence, and conceptual classification. My argument will be based around Poverty of the Stimulus (POS) considerations, as first introduced in section 2.5. Recall from that section the two major explanatory problems with respect to language acquisition. The first problem, that of ‘theory selection’, concerned how infants are able, with remarkable speed and consistency, to acquire an understanding of the syntactic, semantic, phonetic, rules and principles for their language. The second problem concerned the child’s acquisition of an unboundedly creative and productive linguistic capacity. The key POS consideration, if you recall, was that the ‘primary linguistic data’ (PLD) available to the child learner are insufficient to adjudicate between the range of alternative rule hypotheses for their language. Moreover, such PLD were insufficient ever to account for the child’s creative capacity. In section 2.5, our examples focussed on the acquisition of *syntactic* rules and principles. The overwhelming conclusion was that theory selection in language acquisition is constrained by a prior body of domain-specific ‘knowledge’, or ‘cognisance’, as to the possible syntactic rules for the infant’s language. The question of productivity, or creativity, in turn was explained by the fact that these rules were *recursive* in nature.

These problems of language acquisition and the related issues of poverty of the stimulus are not limited to syntactic rules. The problems arise with equal force with respect to the acquisition of word meaning and semantic competence. The problem of semantic theory selection was most famously noted by Quine, albeit in a

somewhat different dialectic endeavour. Consider the position of an infant acquiring the meaning of the word 'dog', say. Simplifying somewhat, suppose that the child's PLD for the meaning of this term consisted largely in ostensions of dogs concurrent with utterances of the word 'dog'. Given such minimal PLD, however, Quine asked what would enable the child to select between indefinitely many meaning hypotheses for 'dog'? The problem, I assume, is familiar: it is compatible with such PLD that the interpretation of 'dog' could be *undetached dog-part*, or *the universal dogness*, *shape of dog*, or ... whatever.¹⁷⁹ Quine considered this apparent indeterminacy to be irresolvable: what could *possibly* dictate that 'dog' meant one thing for us, rather than another? The notion of meaning for Quine possessed no stable theoretical basis. Semantic irrationalism ensued.

Quine's scepticism seems somewhat premature, however. The overwhelmingly obvious fact of the matter is that infants *do* consistently converge on more or less the same understanding of the meaning of 'dog'. We therefore need an explanation as to *how* this semantic theory selection occurs. The same goes for the second problem of language acquisition: how does the child learner, given absolutely minimal PLD, acquire a *productive* and *creative* understanding of the meaning of 'dog'? That is, how does the child acquire an understanding of 'dog' that explains its potential application in indefinitely many novel and varied cases? I shall argue that the notion of similarity, and my proposed similarity constraint SIM, lie at the very heart of this explanatory endeavour.

¹⁷⁹ (Quine, 1960). Nelson Goodman, (Goodman, 1966), noted the existence of another range of alternate meaning hypotheses, such as *dog until 2010*, or *werewolf thereafter*.

The two explanatory problems of meaning acquisition are in fact even more puzzling than Quine makes out. The infant's PLD not only underdetermine her semantic theory selection: they are also in radically short supply. Most children acquire their first referential words between 9 and 15 months.¹⁸⁰ Studies on children at 13 months demonstrate that it can take as little as six to nine presentations of a novel term for the child to acquire its meaning.¹⁸¹ By 24 months, children have reached a stage known as 'fast mapping',¹⁸² on which word meanings are learned after hearing a word only once or twice. This achievement is all the more remarkable, given the child's ability thereafter to apply such terms successfully to novel and varied instances. The child usually receives no explicit guidance as to 'how to go on' in applying the term. By their very nature, PLD are finite, relative to specific instances, and rarely contain any information as to the extent or breadth of the category being acquired.

Theorists of a Chomskyan persuasion are optimally placed to provide an explanation of these abilities. The Chomskyan strategy here is analogous to that considered in Section 2.5, which hypothesised the existence of constraints on the range of syntactic rules used to interpret inputted utterances. The 'constraints/principles' approach to *concept* acquisition, now an extremely popular position in cognitive science,¹⁸³ posits six lexical principles by means of which a child narrows the range of possible meaning interpretations,¹⁸⁴ and converges upon the

¹⁸⁰ (Benedict, 1979).

¹⁸¹ (Woodward, Markman, & Fitzsimmons, 1994).

¹⁸² (Carey, 1978), (Carey & Bartlett, 1978), (Heibeck & Markman, 1987).

¹⁸³ Kim Plunkett, personal communication.

¹⁸⁴ Douglas Behrend says that 'the concept of a constraint on development is implicit in the term *range of reaction*, which is commonly used in developmental behaviour genetics. .. The range of reaction is a familiar

most likely interpretation for a particular word.¹⁸⁵ These principles appear at different stages in the child's development, but three in particular seem to be present very early in language acquisition. These 'first tier' principles include the principle of Reference, which dictates that words are mapped onto representations of objects, actions, events, or attributes in the environment.¹⁸⁶ The principle of Object Scope dictates that words refer to objects over actions or events, and name whole objects, as opposed to object parts or attributes.¹⁸⁷ Finally there is the principle of Extendibility, according to which 'words may be used to label similar referents, even if the child has not yet heard someone else label that referent with that word.'¹⁸⁸ It is not my intention here to explicate the principles of Reference or Object Scope. I hope that it is intuitively clear how they might be employed to combat Quine's worries about meaning indeterminacy. In any case, the fact that there exists a successful scientific program investigating precisely Quine's problem should allay any strong form of scepticism about the acquisition of meaning.

The principle of Extendibility is more directly relevant to our present concerns. The principle of Extendibility is stated with direct reference to the notion of similarity. Why should we think that the notion of similarity is so central to language acquisition? There is strong evidence that the 'career of similarity', to borrow a phrase from Dedre Gentner and Mary Jo Ratterman, follows the path of

theoretical construct which states that one's genetic makeup (genotype) determines the set of potential expressions of that makeup (phenotype)', (Behrend, 1990, p. 317).

¹⁸⁵ (Golinkoff, Mervis, & Hirsh-Pasek, 1994), (Mervis & Bertrand, 1994). There exists a sizeable contingent in the linguistics community that eschews the cognitive constraint/principles approach in favour of more social/pragmatic factors of joint attention, pointing, salience, etc. For a summary, see (Hollich, et al., 2000, pp. 1-16). I cannot arbitrate on this issue here, although I suspect that a hybrid theory of the sort outlined by (Hirsh-Pasek, Golinkoff, Hennon, & McGuire, 2004) provides the most reasonable position.

¹⁸⁶ (Macnamara, 1982).

¹⁸⁷ Empirical support is found in (Markman & Wachtel, 1988) and (Woodward, 1992). The term 'object' should be taken to mean 'spatiotemporally bounded entity', or some such.

¹⁸⁸ (Mervis & Bertrand, 1994, p. 1647).

language acquisition remarkably closely. At the earliest stage of development, infants typically only judge two objects to be similar if they are matching in respect of almost all featural and contextual aspects.¹⁸⁹ This has been called the *holistic*, or *global* stage of similarity.¹⁹⁰ During the global similarity stage of categorisation, infants display several closely related tendencies in their use of language.

Experiments conducted by Kathy Hirsh-Pasek and colleagues, indicate that 10-12 month old children exhibit a ‘proper-name bias’, treating most terms as if they were names for particular objects, as opposed to categories.¹⁹¹ A study by Margaret Harris and colleagues found that around 50% of the first forty words acquired by a sample of English-speaking infants were ‘context-bound’, used to refer to particular objects only in the specific contextual circumstances in which the expression was first introduced.¹⁹² For instance, one infant studied would produce ‘*brrrm brrrm*’ only when pushing along a particular toy car.¹⁹³ These naming tendencies suggest that initial representations associated with words are grouped on the basis of global similarity. The child’s similarity judgements are indicative of their nascent conceptual organisation.

Kathy Hirsh-Pasek and colleagues performed experiments confirming the proper name bias in young infants. Infants were first taught a label for an object with which they were playing, (‘danu’). They were then presented with the initial danu

¹⁸⁹ At the pre-verbal stage, evidence of similarity judgements is typically gathered from a) visual preference, (Bomba & Siqueland, 1983), (Quinn, 1987); b) object examination, (Oakes & Tellinghuisen, 1994); and c) sequential touching tasks, (Nelson K. , 1973), (Starkey, 1981). For an overview of these methodologies, see (Mareschal & Quinn, 2001).

¹⁹⁰ Global similarity judgements may be related to a lack of perceptual refinement in young infants. Evidence suggests that very young children often perceive features that are separable by adults, such as brightness and shape, as integral. See (Garner & Felfody, 1970) and (Shepp, 1978).

¹⁹¹ (Hirsh-Pasek, Golinkoff, Hennon, & McGuire, 2004, p. 190).

¹⁹² (Harris, Barrett, Jones, & Brookes, 1988). See also (Barrett, 1986).

¹⁹³ (Gillis, 1986; Gillis, 1987).

along with a completely dissimilar object, to confirm that they would apply the name 'danu' to the correct object. The next stage was to present the initial danu along with a perceptually very similar object, differing only slightly in colour. Infants at 10 to 12 months old would only apply the term 'danu' to the specific original object, exhibiting a lack of extension of the term to the similar object.¹⁹⁴

At around 14 months, however, under some experimental conditions infants would extend the term 'danu' to the perceptually similar object. This coincides with the beginning of a well documented shift in the child's similarity judgements, from global similarity to what Gentner and Ratterman call *mere-appearance*, or *feature-based* similarity. Between 14 and 24 months of age, infants' similarity judgements start taking into account individual features of objects, such as size, shape, colour, or form, independently of their contextual association. Similarity in respect of such features seems to be the predominant basis for conceptual groupings at this stage, and likewise for the extension of novel terms. Gentner and Ratterman suggest that 'young children may be operating under the assumption that the extensions of object names are based on physical similarity'.¹⁹⁵ Pre-school infants, age 2 onwards, are particularly biased towards extending terms to objects of a similar *shape*.¹⁹⁶

Nelson Goodman has objected to the use of similarity in explaining language acquisition that immature appearance-based similarity fails to illuminate our

¹⁹⁴ Factors such as the perceptual salience of the original danu, and the syntactic frame within which the word 'danu' is introduced were controlled in these experiments. For instance, experimenters did not introduce 'danu' in the frame '*the* danu', which might imply uniqueness.

¹⁹⁵ (Gentner & Ratterman, 1991, p. 256).

¹⁹⁶ On the 'shape bias', see (Landau, Smith, & Jones, 1988), (Clark, 1973); (Anglin, 1977); (Bowerman, 1976); (Tomikawa & Dodd, 1980). (Soja, Carey, & Spelke, 1991) have shown that shape-based extension is reduced with nonsolid substances. (Becker & Ward, 1991) show similar results with animate objects. They emphasise the role of tacit biological/physical theory in supporting generalisations. These results show, as if it were in any doubt, that the child's acquisition of meaning is not one-dimensional, and may be subject to a host of differing constraints. This does not undermine the central role of perceptual similarity as outlined here.

understanding of more abstract language.¹⁹⁷ Goodman's objection misses its mark, however, as infants beyond about 2 years of age develop what is known as a *mature* principle of Extendibility, which takes into account non-perceptual information.¹⁹⁸ For example, it has been widely documented that similarity in respect of *relational* properties can affect naming with comparable consistency to appearance-based properties.¹⁹⁹ This includes basic relations such as *being identical to*, spatial properties such as *relative location*, and *relative size*, relational verbs such as *falling from*, or *giving to*, more abstract relations such as *permitting*, as well as functional properties, such as *being used to cut things*.²⁰⁰ As with Dummett's narrow focus on observational terms, Goodman's equally constricted conception of similarity fails to do justice to the full range of similarity judgements involved in lexical acquisition.

Another persistent objection to the role of similarity in naming is that the notion is just too *context-sensitive* and *theory-driven* to provide the basis for our relatively stable understanding of language. To consider but one example, white hair and grey hair are typically deemed more similar than grey hair and black hair, whereas grey clouds and black clouds are deemed more similar than grey clouds and white clouds.²⁰¹ A plausible explanation of this asymmetry is that grey hair and white hair are linked by a tacit theory of *aging*, whereas grey clouds and black clouds

¹⁹⁷ (Goodman, 1972).

¹⁹⁸ (Malt, Sloman, Gennari, Shi, & Wang, 1999, pp. 232; 237-8). (Goldstone, 1994, p. 144) suggests that the experiments criticised by Goodman are biased towards eliciting judgements of purely perceptual similarity.

¹⁹⁹ (Gentner & Ratterman, 1991, pp. 226-7). On relational similarity, see (Smith L. B., 1989); (Gentner, 1983; Gentner, 1989). The 'relational-shift hypothesis' of (Gentner, 1988) conjectures that judgements with respect to object-based, featural commonalities precede those made with respect to relational commonalities.

²⁰⁰ (Malt & Johnson, 1992) discuss functional similarity groupings for artefact concepts.

²⁰¹ (Medin & Shoben, 1988). See also (Roth & Shoben, 1983), (Hardiman, Dufresne, & Mestre, 1989), (Sjoberg, 1972), (Barsalou, 1983) and (Medin, Goldstone, & Gentner, 1993). (Suzuki, Ohnishi, & Shigemasa, 1992) found that judgements of similarity were dependent on the subject's goals.

are linked by a tacit theory of *precipitation*.²⁰² There is no doubt that similarity judgements are complex, multifaceted, and often variable phenomena. Attempts have been made to model such context sensitivity within prototype/exemplar theories of concepts. Itamar Gati and Amos Tversky, for instance, suggest that features might be assigned different *weights*, depending on context.²⁰³ The function from contexts to weightings, however, remains elusive – as with much else about the specifics of context dependence. This remains an important area for future research.

In summary, I doubt anybody would deny that naming and conceptual classifications are extremely complicated processes. Infants bootstrap from a wide range of social,²⁰⁴ semantic, syntactic,²⁰⁵ intonational information, cognitive biases and tendencies. The principle of Extendibility clearly will not provide an *exhaustive* account of the acquisition of semantic competence; but given the evidence, its role cannot be denied.

Let us relate this discussion of similarity, and the principle of Extendibility, back to the problems of language acquisition discussed above. Once into the fast-mapping stage, infants seem to acquire an understanding of the meaning of a novel term given as few as one or two exposures. Moreover, that understanding licenses creative and productive use of such terms to novel and varied instances. The task facing the child learner seemingly would be insurmountable, were it not for a range

²⁰² So-called ‘*Theory Theory*’ holds that categorisations are *theory* relative, in the sense of being sensitive to the representation of certain tacit, or folk principles. See e.g. (Carey, 1985) and (Gelman, 1988). Theory Theorists do not tend to deny the role of prototypes in such representations, however, but hold that more is needed to explain certain categorising phenomena.

²⁰³ (Gati & Tversky, 1984).

²⁰⁴ Hirsh-Pasek and colleagues note that the mature principle of Extendibility marks an increased use of *social* information pertaining to naming conventions, sometimes in the complete absence of perceptual information, in extending a novel term. See also (Clark, 1993) and (Golinkoff, Mervis, & Hirsh-Pasek, 1994).

²⁰⁵ On ‘syntactic bootstrapping’ in acquiring the meaning of a word, see e.g. (Gleitman, 1990).

of prior constraints on the range of possible meaning interpretations. The notion of similarity has a crucial role to play in this respect. The child's immediately productive linguistic capacity with respect to category terms such as 'dog', 'red', 'chair', 'bald', and so on, requires that they cognise that 'dog' applies to anything *similar* to what is known to be, or represented as being a prototypical or exemplary dog. The principle of Extendibility, then, plays an important explanatory role in respect of our most basic linguistic capabilities.

The principle of Extendibility does not simply constrain classification with respect to similarity *to a prototype*, however. As noted by Barbara Malt and colleagues, naming is often 'influenced by something other than perceived similarity to central examples: It is influenced by *similarity to a near neighbour that may be at some distance from central examples, which in turn may have been influenced by another near neighbour*'.²⁰⁶ That is, for *any* two objects x and y similar in respect of P , no matter whether one is a prototype representation, we often classify such that if x is F , then y is F , and vice versa. Our constraint SIM thus enters the fray. Reflecting again upon the child's acquisition of a productive linguistic capacity should lead us to see the psychological necessity of such a constraint. Suppose that a child encounters three dogs in sequence, a , b , and c . Dogs a and b are pairwise similar, as are b and c . Dogs a and c are comparatively dissimilar. Now suppose that dog a forms the basis for the child's nascent prototype representation DOG. When presented with b , the child might deploy the prototypicality constraint PRO to classify b as a dog. Beyond this point, the child's extension of 'dog' cannot rely merely upon similarity to a prototype. The extension of DOG to c requires that the child deploy a principle on

²⁰⁶ (Malt, Sloman, Gennari, Shi, & Wang, 1999, p. 234), italics mine.

which c is classified the same way as b , which was classified as a dog. The extension of that concept requires that the child deploy our similarity constraint, SIM.

Evidence suggests that subjects frequently classify objects as F on the basis of similarity to other objects classed as F , and not necessarily on proximity to a prototype for F . This phenomenon, known in cognitive psychology as ‘category chaining’, is an essential aspect of creative language use, and the Extendibility of terms. Barbara Malt and colleagues discuss the term ‘bottle’, and cognates in other languages. ‘Bottle’ is used to classify many containers judged quite dissimilar from prototypical bottles, and they argue that this depends on a ‘chain of intermediate objects’ linking the containers to the prototypes.²⁰⁷ Intuitive reflection upon a host of other terms suggests similar chaining effects.

It might be objected that SIM is not required as an independent constraint. Simply operating with PRO will suffice. Suppose that once b is classified as a dog, the subject’s nascent prototype representation DOG is updated with any features that b possesses, but a lacks.²⁰⁸ It is perfectly conceivable that such an updated prototype, call it a' , would possess greater similarity to c , such that c could be classified as a dog via direct similarity to the prototype representation. We can represent this formally, using the feature matching model of Itamar Gati and Amos Tversky. On this model, an object x is represented as a feature set X . The similarity of two objects x and y , $S(x, y)$, is calculated as a function of three arguments: $F(X \cap Y, X - Y, Y - X)$, corresponding to the features shared by x and y , those possessed by x but

²⁰⁷ The example admittedly is somewhat speculative and could conceivably depend on a range of social, conventional features, as opposed to cognitive classificatory principles.

²⁰⁸ Recall from section 3.3 that exemplar/prototype representations are held sometimes to be updated with new information as novel instances are encountered. I shall not discuss the specific details behind such updating.

not y , and those possessed by y but not x respectively. Now let us represent our three dogs a , b , and c by the following features sets: $A=\{P_1, P_2, P_3\}$; $B=\{P_1, P_2, P_3, P_4\}$; $C=\{P_2, P_3, P_4\}$. Now consider these three similarity functions:

$$S(a, b) = F(A \cap B=3, A-B=0, B-A=1)$$

$$S(b, c) = F(B \cap C=3, B-C=1, C-B=0)$$

$$S(a, c) = F(A \cap C=2, A-C=1, C-A=1)$$

Even without defining the function F on these arguments, we can see intuitively that a and b will be judged fairly similar, as will b and c . Each pair share three properties in common, and differ by at most one property. Dogs a and c , however, are intuitively more dissimilar: they have a lower number of shared features, and a greater difference in features. Now suppose that a is updated to reflect the novel feature P_4 encountered with b , yielding the new prototype representation $A' = \{P_1, P_2, P_3, P_4\}$. The similarity judgement between that prototype and c is calculated as follows:

$$S(a', c) = F(A' \cap C=3, A'-C=1, C-A'=0)$$

The degree of similarity between a' and c is now intuitively the same as that between a and b . 'Dog' might therefore be extended to c via a deployment of PRO, our prototypicality constraint, as opposed to SIM. Perhaps this sort of model captures what is happening in other category chains. Does this undermine the potential role

for SIM in language acquisition and our extension of terms? Doesn't SIM simply collapse into PRO?

There are two responses to this challenge. In the next section, I shall provide evidence of dissociation between SIM and PRO, which supports the claim that SIM is an independent classificatory constraint. First, however, we should note that the 'prototype updating' model might threaten to undermine the *distinctiveness* of our conceptual representations. If opposing or contrary prototypes, for instance, were updated with all information encountered in novel instances falling 'between' the two classificatory poles, the representations would increase their overall similarity to one another. But assuming then that in many cases our prototype representation *a* would maintain its original feature profile *A*, and not be updated with feature *P₄* from *B*, *c* will thus be classified as a dog via SIM as opposed to PRO, if at all.

3.4.2 Dissociations between PRO and SIM

If SIM and PRO really are distinct constraints on our conceptual classification, we might expect there to be dissociations between them. That is, we might expect to find situations in which SIM is applied, but not PRO. Remarkably, clinical cases of such dissociations have been encountered. Debi Roberson and colleagues report the plight of a patient LEW, who has severely limited classification and naming capabilities.²⁰⁹ LEW is largely unable to produce colour names in response to particular colour instances. Crucially, LEW is also unable to make typicality judgements with respect to colour categories. Given a range of red colours, LEW consistently selected colours outside of the normal focal area for red as the 'best' examples of that colour. LEW's

²⁰⁹ (Roberson, Davidoff, & Braisby, 1999).

similarity judgements, however, are within normal range. The most revealing task for LEW involved free-sorting colours. LEW was presented with a large sample of 58 colour chips, all of which were very closely spaced in the colour continuum. Whereas normal patients would free-sort the chips around prototypical focal points of the various colour categories, LEW was unable to group colours in this way. Instead, LEW grouped the colours by assessing pairwise similarity.²¹⁰ This resulted in a small number of groupings, each containing a range of different colours. One group, for instance, contained pink, yellow and light purple chips. LEW's colour groupings, although quite unlike those of normal subjects, were not random, however. The defining characteristic of LEW's groupings was that each member chip abutted, or bordered, another member in colour space. Roberson and colleagues summarise this finding as follows:

It seems as though LEW, focusing as he does on pairwise comparisons between colour chips, is adopting precisely [a] sort of Sorites reasoning... Somehow, normal subjects are managing to avoid drawing these seemingly inevitable conclusions.²¹¹

LEW's predicament is obviously a complicated one, and any candidate explanations must be approached with caution. Given LEW's patent failure on typicality judgements, and his seemingly associated deficit in colour naming, however, a very likely explanation would seem to be as follows. As Roberson and colleagues suggest, in the absence of any prototype information, or perhaps *access* to such information,²¹²

²¹⁰ LEW's classificatory tendencies were not limited to colour. The experiments were repeated with similar results in free-sorting pictures of facial expressions.

²¹¹ (*Ibid*, pp. 11-12)

²¹² Roberson et al. consider and reject the hypothesis that LEW might have a working, or short term memory deficit, which restricts him from utilising stored prototype/exemplar information in the classification task.

the only classificatory principles left for LEW are those involving pairwise similarity. Normal patients are able to employ pairwise similarity, but only *to a point*, in forming colour groupings around prototypical focal colours. For LEW, however, in the absence of such *centre-focussed* information, there are no competing constraints to limit his pairwise similarity classifications. Within the optimality framework described above, such cases can be easily accommodated. The case of LEW provides strong evidence for the mental representation of SIM in normal subjects. If this account is accepted, it will provide a compelling account of the intuitions of ordinary speakers with respect to the major premise of the sorites.

3.4.3 Explaining Borderline Cases

The optimality framework of competing constraints can, I believe, also explain the phenomenon of borderline cases. Recall that a borderline case *b*, with respect to a concept *F*, lies somewhere between clear cases of *F* and clear cases of *not-F*, although it is unclear and undecidable, for a normal, competent language user, whether *b* is in, or out, of the extension of *F*, no matter how far *S* extends her evidence, within her capacity, with respect to *b*. My suggestion is that *b* constitutes a borderline case when there is *no optimal categorisation of b either as F or as not F*. Allow me to expand.

As noted by Mark Sainsbury, vague 'concepts tend to come in systems of contraries: opposed pairs like child/adult, hot/cold, weak/strong, true/false, [*sic.*] and the more complex systems exemplified by our colour terms.'²¹³ Such lexical contrast, or semantic opposition is a pervasive feature of our natural language meanings and concepts. Much of language acquisition therefore must require the

²¹³ (Sainsbury, 1990, p. 258). See also (Dummett, 1975, p. 106).

infant to organise her nascent conceptual representations in a way that captures these oppositions. Consider two concepts F and G , which are represented as lexically contrasting, such as ‘child’ and ‘adult’. Further consider some object x that is ‘equidistant’ between the prototypes for F and G . In such cases, there will be conflicting classificatory directives for x , issuing from the similarity chains, via SIM, existing between x and F , and x and G . No outranking constraint exists that could determine whether x is more similar to F than to G , or vice versa. The subject is left in a state of conceptual indecision: while there exists no outranking constraint to *override* the subject classifying x as F , the same can be said for classifying x as G . Our conceptual Evaluator function, if you will, fails to yield an optimal classificatory directive. This perfectly captures the psychology of borderline cases: one feels a pull in either direction, neither of which is conclusive or decisive, neither of which is clearly wrong or overruled. One is in a psychological state of *quandary* as to how to classify the case.²¹⁴ Note how different this is to the putatively borderline case of *child** considered in Section 1.1. Presented with a 17 year old, we do not have any *competing* or *conflicting* classificatory principles as to how to classify this case with respect to *child**. Rather, we simply lack *any* principle governing this case.

3.5 The Need for a Cognitive Similarity Constraint

Why think that the similarity constraint SIM is *mentally represented* in our cognitive architecture? Could we not account for the child’s extension of novel terms via a simpler *behavioural* explanation, one that altogether eschews the attribution to infants of constraints on conceptual classification? A case in point would be Quine, who

²¹⁴ (Wright, 2001).

viewed infants' creative capacity to extend novel terms merely as a sophisticated form of *imitation*. Quine does take 'perceptual similarity' to play a role in such imitation; but he employs a thoroughly behavioural notion of similarity. Quine says that 'an episode *a* is proved to be perceptually more similar to *b* than to *c* when the subject has been conditioned to respond in some fashion to *b* and not to so respond to *c*, and then is found to respond in that fashion to *a*'. The subject's response to *a* need not be immediate: rather, it is 'a question of the subject's disposition to submit to conditioning in one way and another; hence of his disposition to *acquire or change* his habits of response' to *a*.²¹⁵ Quine exploits this behavioural notion of similarity in explaining language acquisition:

[C]onsider the learning of the word 'red'. Suppose the child happens to utter the word in the course of the random babbling that is standard procedure in small children, and suppose a red ball happens to be conspicuously present at the time. The parent rewards the child, perhaps only by somehow manifesting approval... On a later occasion a red shawl is conspicuously present. Its colour makes for a degree of perceptual similarity between the pleasant earlier episode and the present, thus enlivening the [memory] trace of that episode. The child contorts his speech muscles so as to add what more he can to the similarity: he again says 'red', and we may hope that the similarity is yet further enhanced by a recurrence of the reward.²¹⁶

The explanation of the child's extension of 'red' seems to be as follows. First, the child's association between the production of a novel term *T* and an object *b* is positively reinforced by some act upon the child, such as offering approval, which the child finds pleasurable. The child is thereby conditioned to respond to *b* by saying *T*. Second, some object *a*, the shawl, is presented, which is perceptually

²¹⁵ (Quine, 1974, p. 18), italics added.

²¹⁶ (*Ibid*, p. 29).

similar to *b*, the ball. Under Quine's analysis of perceptual similarity, this means that the subject will be disposed to acquire or change his response habits with respect to *a*, to bring them into line with those to *b*. Third, this perceptual similarity supposedly triggers the child's memory (a 'trace') of the positive episode associated with *b*. Fourth, in order to replicate the recalled pleasurable outcome associated with *b*, the child will endeavour to recreate in full the conditions that led to the positive appraisal in that instance. This will induce the child to reproduce the term *T*, whose use was positively reinforced in that initial episode, in the presence of *a*. The term *T* is thus extended from *b* to *a*. Quine concludes that 'the child would be said in this instance to have imitated the parent, not only in saying ['red'] but in saying it in circumstances similar to those in which the parent said it.'²¹⁷

Quine's account leaves *much* to be desired. For one, his behavioural explication of similarity simply begs the question as to which aspects of cognition *cause* and *explain* the child's behavioural disposition to respond the same way to *a* as she does to *b*. The role of *imitation* in Quine's account is also troubling. As with all other Chomskyan ascriptions of mentally represented constraints, rules or principles, our motivation for claiming that SIM was a *cognitive* similarity constraint stemmed from Poverty of the Stimulus considerations. The untutored creativity of young infants far exceeds the samples of language to which they are exposed, and which they could imitate. This fact demands a substantive cognitive explanation, in which the notion of similarity is likely to feature prominently. Quine's purported explanation, in contrast, provides only a superficial behavioural account of these puzzling phenomena.

²¹⁷ (*Ibid*, pp. 30-31).

It should be noted that the role of imitation in language acquisition remains a hotly debated topic, and a highly contentious region of cognitive linguistic research. The famous debate between Piaget and Chomsky in the 1970s revolved in no small part around the role of imitation, with Piaget advocating nonverbal imitation as a ‘developmental foundation for the emergence of language’, and Chomsky advocating a separation of the two domains.²¹⁸ The true story, as in so many cases, probably involves elements of both views. I cannot settle the matter here. Quine’s proposed explanation of infants’ extension of novel terms, however, seems thoroughly unlikely to match the scope and power of our cognitive account.

²¹⁸ (Meltzoff & Gopnik, 1989, p. 23).

4. Conclusions

In this thesis, I have defended a novel explanation of vagueness: Cognitive Inconsistentism. This theory holds that vagueness *qua* sorites susceptibility is best explained by an important component of our semantic competence: the similarity constraint. The constraint causally explains our intuitions with respect to the major premise of the sorites. The constraint is independently motivated by its explanation of important aspects of language acquisition, in particular our productive and creative capacity with respect to word meanings.

The similarity constraint was presented within the framework of a cognitive, broadly Chomskyan, theory of linguistic understanding. In contrast to a Dummettian, abilities view of linguistic understanding, the Chomskyan theory possesses significant explanatory power. The Chomskyan theory also undermines Eklund's inferential role theory of understanding, which took aspects of linguistic performance, or dispositions to perform, as constitutive of semantic competence. Dummett's and Eklund's theories of understanding lacked the potential adequately to explain the psychological question with respect to vagueness. Cognitive Inconsistentism, on the other hand, provides an empirically viable, philosophically persuasive, and explanatorily satisfying account of the psychology of vague terms and concepts.

Bibliography

- Aitchison, J. (1989). *The Articulate Mammal: An Introduction to Psycholinguistics* (3rd ed.). London: Unwin Hyman.
- Anglin, J. M. (1977). *Word, object, and conceptual development*. New York: Norton.
- Antony, L. M. (1997). "Meaning and semantic knowledge". *Proceedings of the Aristotelian Society, Supplementary Volume*, 71, 177-207.
- Armstrong, S. L., Gleitman, L. R., & Gleitman, H. (1983). "What some concepts might not be". *Cognition*, 13, pp263-308.
- Ashby, F. G., & Maddox, W. T. (1993). "Relations between prototype, exemplar and decision bound models of categorization". *Journal of Mathematical Psychology*, 37, pp372-400.
- Baron-Cohen, S., Leslie, A., & Frith, U. (1985). "Does the autistic child have a Theory of Mind?". *Cognition*, 21, pp37-46.
- Barrett, M. (1986). "Early semantic representations and early word-usage". In M. Barrett (Ed.), *The development of word meaning*. New York: Springer-Verlag.
- Barsalou, L. W. (1983). "Ad hoc categories". *Memory & Cognition*, 11, pp211-27.
- Becker, A. H., & Ward, T. B. (1991). "Children's use of shape in extending novel labels to animate objects: Identity versus postural change". *Cognitive Development*, 6, pp3-16.
- Behrend, D. A. (1990). "Constraints and development: A reply to Nelson". *Cognitive Development*, 5, pp313-30.
- Benedict, H. (1979). "Early lexical development: Comprehension and production". *Journal of Child Language*, 6, pp183-200.

- Biederman, I. (1987). "Recognition-by-components: A theory of human image understanding". *Psychological Review*, 94, pp115-147.
- Black, M. (1939). "Vagueness: an exercise in logical analysis". *Philosophy of Science*, 4, pp427-55.
- Blair, M., & Homa, D. (2001). "Expanding the search for a linear separability constraint on category learning". *Memory and Cognition*, 29 (8), pp1153-64.
- Bloom, L. (1973). *One word at a time: The use of single-word utterances before syntax*. The Hague: Mouton.
- Bloom, P. (1990). "Syntactic distinctions in child language". *Journal of Child Language*, 17, pp343-55.
- Bloom, P. (Ed.). (1993). *Language Acquisition: Core Readings*. Cambridge: Harvester Wheatsheaf.
- Bomba, P., & Siqueland, E. R. (1983). "The nature and structure of infant form categories". *Journal of Experimental Child Psychology*, 35, pp294-328.
- Bornstein, M. H., & Monroe, M. D. (1980). "Chromatic information processing: rate depends on stimulus location in the category and psychological complexity". *Psychological Research*, 42, pp213-25.
- Bowerman, M. (1976). "Semantic factors in the acquisition of rules for word use and sentence construction". In D. M. Morehead, & A. E. Morehead (Eds.), *Normal and deficient child language* (pp. 99-179). Baltimore: University Park Press.
- Brown, R. (1973). *A first language: The early stages*. Cambridge, MA: Harvard University Press.

Brown, R., & Bellugi, U. (1964). "Three processes in the child's acquisition of syntax".

In E. H. Lenneberg (Ed.), *New directions in the study of language*. Cambridge: MIT Press.

Burge, T. (1989). "Wherein is language social?". In A. George (Ed.), *Reflections on Chomsky* (pp. 175-91). Oxford: Blackwell Publishers.

Busemeyer, J. R., & Myung, I. J. (1988). "A new method for investigating prototype learning". *Journal of Experimental Psychology: Learning, Memory and Cognition*, 14, pp3-11.

Carey, S. (1978). "The child as word learner". In M. Halle, J. Bresnan, & A. Miller (Eds.), *Linguistic Theory and Psychological Reality* (pp. 264-93). Cambridge: MIT Press.

Carey, S. (1985). *Conceptual changes in childhood*. Cambridge, MA: MIT Press.

Carey, S., & Bartlett, E. (1978). "Acquiring a Single New Word". *Papers and Reports on Child Language Development*, 15, pp17-29.

Chin-Parker, S., & Ross, B. H. (2004). "Diagnosticity and Prototypicality in category learning: A comparison of inference learning and classification learning". *Journal of Experimental Psychology: Learning, Memory and Cognition*, 30 (1), pp216-26.

Chomsky, N. (1980). *Rules and Representations*. Oxford: Basil Blackwell.

Chomsky, N. (1986). *Knowledge of Language*. New York: Praeger.

Chomsky, N. (1988). *Language and the problems of knowledge: The Managua Lectures*. Cambridge: MIT Press.

Chomsky, N. (1995). "Language and nature". *Mind*, 104 (413), pp1-61.

Chomsky, N. (2000). *New horizons in the study of language and mind*. Cambridge: Cambridge University Press.

- Chomsky, N. (2003). "Replies to Commentators". In L. M. Antony, & N. Hornstein (Eds.), *Chomsky and his critics*. Oxford: Blackwell Publishers.
- Clark, E. V. (1973). "What's in a word? On the child's acquisition of language in his first language". In T. E. Moore (Ed.), *Cognitive development and the acquisition of language* (pp. 65-110). New York: Academic Press.
- Clark, E. V. (1993). *The Lexicon in Acquisition*. Cambridge, MA: Cambridge University Press.
- Coleman, M. R., Rodd, J. M., Davis, M. H., Johnsrude, I. S., Menon, D. K., Pickard, J. D., et al. (2007). "Do vegetative patients retain aspects of language comprehension? Evidence from fMRI". *Brain*, 130 (10), pp2494-2507.
- Corrigan, R. (1978). "Language development as related to Stage 6 object permanence development". *Journal of Child Language*, 5, pp173-89.
- Cowie, F. (2008). "Innateness and Language". In E. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Winter 2003 Edition). URL = <http://plato.stanford.edu/entries/innateness-language/>.
- Davidson, D. (1984). "Communication and Convention". In *Inquiries into Truth and Interpretation* (2nd ed., pp. 265-80). Oxford: Oxford University Press.
- Davies, M. K. (1989). "Tacit knowledge and subdoxastic states". In A. George (Ed.), *Reflections on Chomsky* (pp. 131-52). Oxford: Basil Blackwell.
- de Villiers, J., & de Villiers, P. (1999). "Linguistic determinism and the understanding of false beliefs". In P. Mitchell, & K. Riggs (Eds.), *Children's reasoning and the mind*. New York: Psychology Press.

- Dromi, E. (1987). *Early Lexical Development*. Cambridge and New York: Cambridge University Press.
- Dummett, M. (1959). "Truth". *Rpt. in Truth and Other Enigmas*, pp1-24.
- Dummett, M. (1975). "Wang's Paradox", in Keefe & Smith (Eds.). pp99-118.
- Dummett, M. (1981). *Frege: Philosophy of Language* (2nd ed.). Gerald Duckworth & Co.
- Dummett, M. (1993). *The Seas of Language*. Oxford: Oxford University Press.
- Eklund, M. (2002). "Inconsistent Languages". *Philosophy and Phenomenological Research*, 64 (2), pp251-75.
- Eklund, M. (2005). "What vagueness consists in". *Philosophical Studies*, 125 (1), pp27-60.
- Evans, G. (1985). "Semantic theory and tacit knowledge". In S. Holtzman, & C. Leich (Eds.), *Wittgenstein: To follow a rule* (pp. 322-42). Routledge and Kegan Paul.
- Fine, K. (1975). "Vagueness, Truth and Logic", in Keefe & Smith (Eds.). pp119-150.
- Fodor, J. A. (1998). *Concepts: Where cognitive science went wrong*. Oxford: Oxford University Press.
- Fodor, J. A., & Lepore, E. (1999). "All at Sea in Semantic Space: Churchland on Meaning Similarity". *The Journal of Philosophy*, 96 (8), pp381-403.
- Frege, G. (1979). *Posthumous Writings*. (H. Hermes et al., Ed.) Oxford: Basil Blackwell.
- Fuhrmann, G. (1988). "'Prototypes' and 'Fuzziness' in the logic of concepts". *Synthese*, 75, pp317-47.
- Garfield, J., Peterson, C., & Perry, T. (2001). "Social cognition, language acquisition, and the development of the Theory of Mind". *Mind and Language*, 16 (5), pp494-541.

- Garner, W. R., & Felfody, G. L. (1970). "Integrality of stimulus dimensions in various types of information processing". *Cognitive Psychology*, 7, pp225-41.
- Gati, I., & Tversky, A. (1984). "Weighting common and distinctive features in perceptual and conceptual judgements". *Cognitive Psychology*, 16, pp341-70.
- Gelman, S. A. (1988). "The development of induction within natural kind and artifact categories". *Cognitive Psychology*, 20, pp65-95.
- Gentner, D. (1983). "Structure-mapping: A theoretical framework for analogy". *Cognitive Science*, 7, 155-70.
- Gentner, D. (1988). "Metaphor as structure mapping: The relational shift". *Child Development*, 59, pp47-59.
- Gentner, D. (1989). "The mechanisms of analogical learning". In S. Vosniadou, & A. Ortony (Eds.), *Similarity and analogical reasoning* (pp. 199-241). Cambridge: Cambridge University Press.
- Gentner, D., & Ratterman, M. J. (1991). "Language and the career of similarity". In S. A. Gelman, & J. P. Byrnes (Eds.), *Perspectives on Thought and Language: Interrelations in Development* (pp. 225-77). London: Cambridge University Press.
- Gerken, L., & McIntosh, B. J. (1993). "The interplay of function morphemes and prosody in early language". *Developmental Psychology*, 29, pp448-57.
- Gillis, S. (1986). "The child's nominal insight is actually a process: The plateau-stage and the vocabulary spurt in early lexical development". *Antwerp Papers in Linguistics*, 45.

- Gillis, S. (1987). "Words and categories at the onset of language acquisition: Product versus process". In A. de Houwer, & S. Gillis (Eds.), *Perspectives on child language* (pp. 37-53). Brussels: University de Bruxelles.
- Gleitman, L. (1990). "The Structural Sources of Verb Meanings". *Language Acquisition*, 1 (1), pp3-55.
- Goldfield, B. A., & Reznick, J. S. (1990). "Early lexical acquisition: rate, content, and the vocabulary spurt". *Journal of Child Language*, 17, pp171-183.
- Goldstone, R. L. (1994). "The role of similarity in categorization: providing a groundwork". *Cognition*, 52, pp125-57.
- Golinkoff, R. M., Mervis, C., & Hirsh-Pasek, K. (1994). "Early object labels: The case for a developmental lexical principles framework". *Journal of Child Language*, 21, pp125-55.
- Goodman, N. (1951). *The Structure of Appearance*. Cambridge, MA : Harvard University Press.
- Goodman, N. (1966). "The New Riddle of Induction". *The Journal of Philosophy*, 63, pp281-331.
- Goodman, N. (1972). "Seven strictures on similarity". In N. Goodman (Ed.), *Problems and Projects*. New York: Bobbs-Merrill.
- Gopnik, A., & Meltzoff, A. (1987). "The Development of Categorization in the Second Year and its Relation to Other Cognitive and Linguistic Developments". *Child Development*, 58 (6), pp1523-31.
- Gopnik, A., & Slaughter, V. (1991). "Young children's understanding of changes in their mental states". *Child Development*, 62, 98-110.

- Graff, D. (2000). "Shifting Sands: An interest-relative theory of vagueness". *Philosophical Topics*, 28 (1), pp45-81.
- Greenberg, M. (2005). "A New Map of Theories of Mental Content: Constitutive Accounts and Normative Theories". *Philosophical Issues*, 15 (1), pp299-320.
- Greenough, P. (2003). "Vagueness: A Minimal Theory". *Mind*, 122 (446), pp235-81.
- Hardiman, P. T., Dufresne, R., & Mestre, J. P. (1989). "The relation between problem categorisation and problem solving among experts and novices". *Memory & Cognition*, 17, pp627-38.
- Hardin, C. L. (1988). "Phenomenal Colours and Sorites". *Nous*, 22, pp213-34.
- Harris, M., Barrett, M., Jones, D., & Brookes, S. (1988). "Linguistic input and early word meaning". *Journal of Child Language*, 15, 77-94.
- Hauser, M., Chomsky, N., & Fitch, W. T. (2002). "The Faculty of Language: What Is It, Who Has It, and How Did It Evolve?". *Science*, 298, pp1569-79.
- Heibeck, T. H., & Markman, E. M. (1987). "Word Learning in Children: An examination of fast mapping". *Child Development*, 58, pp1021-34.
- Hendriks, P., & de Hoop, H. (2001). "Optimality Theoretic Semantics". *Linguistics and Philosophy*, 24 (1), pp1-32.
- Hirsh-Pasek, K., Golinkoff, R. M., Hennon, E. A., & McGuire, M. J. (2004). "Hybrid theories at the frontier of developmental psychology: The emergentist coalition model of word learning as a case in point". In D. G. Hall, & S. R. Waxman (Eds.), *Weaving a lexicon* (pp. 173-204). Cambridge, MA: MIT Press.
- Hollich, G. J., Hirsh-Pasek, K., Golinkoff, R. M., Brand, R. J., Brown, E., Chung, H. L., et al. (2000). "Breaking the language barrier: An emergentist coalition model for the

origins of word learning". *Monographs of the Society for Research in Child Development* , 65 (3), ppi-iv + 1-135.

Kamp, H., & Partee, B. (1995). "Prototype theory and compositionality". *Cognition*, 57 (2), pp129-191.

Keefe, R., & Smith, P. (Eds.). (1999). *Vagueness: A Reader*. Cambridge: MIT Press.

Keil, F. C. (1987). "Conceptual development and category structure". In U. Neisser (Ed.), *Concepts and Conceptual Development: The ecological and intellectual factors in categorization*. Cambridge: Cambridge University Press.

Keil, F. C. (1989). *Concepts, kinds, and cognitive development*. Cambridge: Cambridge University Press.

Kruschke, J. K. (1992). "ALCOVE: An exemplar-based connectionist model of category learning". *Psychological Review*, 99, pp22-44.

Kruschke, J. K. (2005). Category Learning. In K. Lamberts, & R. L. Goldstone (Eds.), *The Handbook of Cognition* (pp. 183-201). London: Sage.

Landau, B., Smith, L. B., & Jones, S. S. (1988). "The importance of shape in early lexical learning". *Cognitive Development*, 3, pp299-321.

Laurence, S. (2003). "Is linguistics a branch of psychology?". In A. Barber (Ed.), *Epistemology of Language* (pp. 69-106). Oxford: Oxford University Press.

Laurence, S., & Margolis, E. (1997). "Regress arguments against the language of thought". *Analysis* , 57 (1), pp60-66.

Laurence, S., & Margolis, E. (2001). "The Poverty of the Stimulus Argument". *British Journal for the Philosophy of Science*, 52, pp217-276.

- Lee, M. D., & Navarro, D. J. (2002). "Extending the ALCOVE model of category learning to featural stimulus domains". *Psychonomic Bulletin and Review*, 9, pp43-58.
- Ludlow, P. (2003). "Referential semantics for I-languages?". In L. M. Antony, & N. Hornstein (Eds.), *Chomsky and his critics* (pp. 140-61). Oxford: Blackwell Publishing.
- Lycan, W. (1986). "Tacitness and Virtual Beliefs". *Mind and Language*, 7 (3), pp240-63.
- Lynch, E. B., Coley, J. D., & Medin, D. L. (2000). "Tall is typical: Central tendency, ideal dimensions, and graded category structure among tree experts and novices". *Memory and Cognition*, 28, pp41-50.
- Macnamara, J. (1982). *Names for Things*. Cambridge, MA: MIT Press.
- Maddox, W. T. (2002). "Learning and attention in multidimensional identification and categorization: Separating low-level perceptual processes and high-level decisional processes". *Journal of Experimental Psychology: Learning, Memory and Cognition*, 28 (1), pp99-115.
- Malt, B. C., & Johnson, E. C. (1992). "Do artifact concepts have cores?". *Journal of Memory and Language*, 31, pp195-217.
- Malt, B. C., & Smith, E. E. (1982). "The role of familiarity in determining typicality". *Memory and Cognition*, 10, pp69-75.
- Malt, B. C., Sloman, S. A., Gennari, S., Shi, M., & Wang, Y. (1999). "Knowing versus naming: Similarity and the linguistic categorization of artifacts". *Journal of Memory and Language*, 40, pp230-62.
- Mareschal, D. C., & Quinn, P. C. (2001). "Categorization in Infancy". *Trends in Cognitive Science*, 5 (10), pp443-50.

- Markman, A. B., & Ross, B. H. (2003). "Category use and category learning". *Psychological Bulletin*, 129 (4), pp592-613.
- Markman, E. M., & Wachtel, G. F. (1988). "Children's use of mutual exclusivity to constrain the meanings of words". *Cognitive Psychology*, 20, pp121-57.
- McNamara, T. P., & Sternberg, R. J. (1983). "Mental models of word meaning". *Journal of verbal learning and verbal behavior*, 22 (4), pp449-74.
- McShane, J. (1979). "The development of naming". *Linguistics*, 17, pp879-905.
- Medin, D. L., & Shoben, E. J. (1988). "Context and structure in conceptual combination". *Cognitive Psychology*, 20, pp158-90.
- Medin, D. L., & Smith, E. E. (1984). Concepts and concept formation. (M. R. Rosenzweig, & L. W. Porter, Eds.) *Annual Review of Psychology*, 35, pp113-118.
- Medin, D. L., Goldstone, R. L., & Gentner, D. (1993). "Respects for similarity". *Psychological Review*, 100, pp254-78.
- Meltzoff, A. N., & Gopnik, A. (1989). "On linking nonverbal imitation, representation, and language learning in the first two years of life". In G. E. Speidel, & K. E. Nelson (Eds.), *The Many Faces of Imitation in Language Learning* (pp. 23-51). Berlin: Springer.
- Mervis, C. B., & Bertrand, J. (1994). "Acquisition of the Novel Name-Nameless Category (N3C) Principle". *Child Development*, 65 (6), pp1646-62.
- Mervis, C. B., & Rosch, E. (1981). "Categorization of natural objects". *Annual review of psychology*, 32, pp89-115.

- Minda, J. P., & Smith, J. D. (2001). "Prototypes in category learning: The effects of category size, category structure, and stimulus complexity". *Journal of Experimental Psychology: Learning, Memory and Cognition*, 27 (3), pp775-799.
- Murphy, G. L., & Medin, D. L. (1985). "The role of theories in conceptual coherence". *Psychological Review* , 92, pp289-316.
- Nelson, K. (1973). "Some evidence for the cognitive primacy of categorization and its functional basis". *Merrill-Palmer Quarterly*, 19, pp21-39.
- Nelson, K., & Lucariello, J. (1985). "The development of meaning in first words". In M. Barrett (Ed.), *Single-word speech* (pp. 59-86). New York: Wiley.
- Nosofsky, R. M. (1986). "Attention, similarity, and the identification-categorization relationship". *Journal of Experimental Psychology: General*, 115, pp39-57.
- Oakes, L. M., & Tellinghuisen, D. J. (1994). "Examining in infancy: does it reflect active processing?". *Developmental Psychology*, 30, pp748-56.
- Osherson, D. N., & Smith, E. E. (1981). "On the adequacy of prototype theory as a theory of concepts". *Cognition*, 9 (1), pp35-58.
- Osherson, D. N., & Smith, E. E. (1982). "Gradedness and conceptual combination". *Cognition* , 12 (3), pp299-318.
- Palmeri, T. J., & Nosofsky, R. M. (2001). "Central tendencies, extreme points, and prototype enhancement effects in ill-defined perceptual categorization". *Quarterly Journal of Experimental Psychology Section A: Human Experimental Psychology* , 54 (1), pp197-235.
- Peirce, C. S. (1902). "Vague". In J. M. Baldwin (Ed.), *Dictionary of Philosophy and Psychology* (p. 748). New York: Macmillan.

- Plunkett, K. (1997). "Theories of early language acquisition". *Trends in Cognitive Sciences*, 1 (4), pp146-53.
- Prince, A., & Smolensky, P. (1993). "Optimality Theory: Constraint Interaction in Generative Grammar". In J. J. McCarthy (Ed.), *Optimality Theory in Phonology: A Reader* (pp. 3-17). Oxford: Blackwell Publishers.
- Putnam, H. (1975). "The meaning of 'meaning'". In *Philosophical Papers, Vol. 2*. Cambridge: Cambridge University Press.
- Pylyshyn, Z. (1999). "What's in your mind?". In E. LePore, & Z. Pylyshyn (Eds.), *What is cognitive science?* (pp. 1-25). Oxford: Blackwell Publishers.
- Quine, W. V. (1960). *Word and Object*. Cambridge, MA: MIT Press.
- Quine, W. V. (1974). *The Roots of Reference: The Paul Carus Lectures*. La Salle, IL: Open Court Publishing.
- Quinn, P. C. (1987). "The categorical representation of visual pattern information by young infants". *Cognition*, 27, pp145-79.
- Ravenscroft, I. (2004). "Folk Psychology as a Theory". (E. N. Zalta, Ed.) *The Stanford Encyclopedia of Philosophy*, URL = <<http://plato.stanford.edu/entries/folkpsych-theory/>>.
- Rey, G. (1983). "Concepts and stereotypes". *Cognition*, 15 (1-3), pp237-62.
- Rips, L. J. (1995). "The current status of research on concept combination". *Mind and Language*, 10 (1/2), pp72-104.
- Robbins, P. (2002). "How to blunt the sword of compositionality". *Nous*, 36 (2), pp313-34.

- Roberson, D., Davidoff, J., & Braisby, N. (1999). "Similarity and categorisation: neuropsychological evidence for a dissociation in explicit categorisation tasks". *Cognition*, 71, pp1-42.
- Rosch, E. (1975a). "Cognitive Reference Points". *Cognitive Psychology*, 7, pp532-47.
- Rosch, E. (1975b). "Cognitive representations of semantic categories". *Journal of Experimental Psychology: General*, 104, pp192-233.
- Rosch, E., & Mervis, C. B. (1975). "Family resemblance: Studies in the internal structure of categories". *Cognitive Psychology*, 7, pp573-605.
- Roth, E. M., & Shoben, E. J. (1983). "The effect of context on the structure of categories". *Cognitive Psychology*, 15, pp346-78.
- Sainsbury, M. (1990). "Concepts without boundaries". In R. Keefe, & P. Smith (Eds.).
- Sainsbury, M. (1991). "Is there higher-order vagueness?". *The Philosophical Quarterly*, 41 (163), pp167-82.
- Sheperd, R. N. (1987). "Toward a universal law of generalization for psychological science". *Science*, 237, pp1317-23.
- Sheperd, R. N., Hovland, C. L., & Jenkins, H. M. (1961). "Learning and memorization of classifications". *Psychological Monographs*, 75 (13).
- Shepp, B. E. (1978). "From perceived similarity to dimensional structure: A new hypothesis about perceptual development". In E. Rosch, & B. B. Lloyd (Eds.), *Cognition and Categorization* (pp. 135-67). Hillsdale, NJ: Erlbaum.
- Sjoberg, L. (1972). "A cognitive theory of similarity". *Goteborg Psychological Reports*, 2 (10).

- Smith, E. E., & Medin, D. L. (1981). *Categories and concepts*. Cambridge, MA: Harvard University Press.
- Smith, E. E., & Osherson, D. N. (1984). "Conceptual combination with prototype concepts". *Cognitive Science*, 8 (4), pp337-61.
- Smith, E. E., Shoben, E. J., & Rips, L. J. (1974). "Structure and process in semantic memory: A featural model for semantic decisions". *Psychological Review*, 81, pp214-41.
- Smith, L. B. (1989). "From global similarities to kinds of similarities: The construction of dimensions in development". In S. Vosniadou, & A. Ortony (Eds.), *Similarity and analogical reasoning* (pp. 146-78). Cambridge: Cambridge University Press.
- Smith, N. J. (2005). "Vagueness as closeness". *Australasian Journal of Philosophy*, 83 (2), pp157-83.
- Smith, N. V. (2004). *Chomsky: Ideas and Ideals* (2nd ed.). Cambridge: Cambridge University Press.
- Soja, N., Carey, S., & Spelke, E. (1991). "Ontological categories guide young children's inductions of word meaning: Object terms and substance terms". *Cognition*, 38, pp179-211.
- Starkey, D. (1981). "The origins of concept formation: Object sorting and object preference in early infancy". *Child development*, 52, pp489-97.
- Stich, S. P. (1971). "What every speaker knows". *The Philosophical Review*, 80 (4), pp476-96.
- Stich, S. P. (1978). "Beliefs and Subdoxastic States". *Philosophy of Science*, 45 (4), pp499-518.

- Storms, G., DeBoeck, P., & Ruts, W. (2001). "Categorization of novel stimuli in well-known natural concepts: A case study". *Psychonomic Bulletin & Review*, 8 (2), pp377-84.
- Stromswold, K. (1999). "Cognitive and neural aspects of language acquisition". In E. LePore, & Z. Pylyshyn (Eds.), *What is Cognitive Science?* (pp. 356-400). Oxford: Blackwell Publishers.
- Suzuki, H., Ohnishi, H., & Shigemasu, K. (1992). "Goal-directed processes in similarity judgement". In *Proceedings of the Fourteenth Annual Conference of the Cognitive Science Society* (pp. 343-48). Hillsdale, NJ: Erlbaum.
- Tomikawa, S. A., & Dodd, D. H. (1980). "Early word meanings: perceptually or functionally based?". *Child Development*, 51, pp1103-1109.
- Tye, M. (1995). "Vagueness: Welcome to the quicksand". *Southern Journal of Philosophy*, 33 (Supplement), pp1-22.
- Unger, P. (1979). "There are no ordinary things". *Synthese*, 41 (2), pp117-54.
- Weatherson, B. (Forthcoming). "Vagueness as indeterminacy". *Cuts and Clouds: Essays in the nature and logic of vagueness*. (R. Dietz, & S. Moruzzi, Eds.) Oxford: Oxford University Press.
- Wheeler, S. (1979). "On that which is not". *Synthese*, 41 (2), pp155-94.
- Williamson, T. (1994). *Vagueness*. London: Routledge.
- Williamson, T. (1997). "Imagination, Stipulation and Vagueness". *Philosophical issues*, 8, pp215-28.
- Williamson, T. (1999). "Schiffer on the epistemic theory of vagueness". *Philosophical Perspectives*, 13, pp505-517.

- Wittgenstein, L. (2001). *Philosophical Investigations*. Oxford: Blackwell Publishers.
- Woodward, A. L. (1992). "The effect of labelling on children's attention to objects". In E. V. Clark (Ed.), *Proceedings of the 24th annual child language research forum*. Chicago: University of Chicago Press.
- Woodward, A. L., Markman, E. M., & Fitzsimmons, C. M. (1994). "Rapid word learning in 13- and 18-month olds". *Developmental Psychology*, 30 (4), pp553-66.
- Wright, C. (1976). "Language-mastery and the sorites paradox", in Keefe & Smith (Eds.). pp151-73.
- Wright, C. (1986). "Theories of meaning and speakers' knowledge". In S. G. Shankar (Ed.), *Philosophy in Britain today*. London and Sydney: Croom Helm.
- Wright, C. (1987). "Further reflections on the sorites paradox", in Keefe & Smith (Eds.). pp204-50.
- Wright, C. (2001). "On being in a quandary. Relativism vagueness logical revisionism". *Mind*, 110 (437), pp45-98.
- Yuille, A. L., & Ullman, S. (1990). "Computational theories of low-level vision". In D. Osherson, S. Kosslyn, & J. Hollerbach (Eds.), *Visual cognition and action (Vol.2): an invitation to cognitive science* (pp. 5 - 39). Cambridge, MA: MIT Press.