Semantic Structure and Language Development

by

Paul Bloom

B.A. Psychology (Honors), McGill University
(1985)

Submitted to the Department of Brain and Cognitive Sciences in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

at the

Massachusetts Institute of Technology

September, 1990

Signature of Author _____

Department of Brain and Cognitive Sciences
September 5, 1990

Certified by _____

Susan Carey
Professor, Brain and Cognitive Sciences
Thesis Supervisor

Accepted by _____

Emilio Bizzi
Chair, Department of Brain and Cognitive Sciences

Semantic Structure and Language Development

Paul Bloom

Submitted to the Department of Brain and Cognitive Sciences in September, 1990 in partial fulfillment of the requirements for the Degree of Doctor of Philosophy in Cognitive Science.

Abstract

This thesis concerns the relationship between grammar and non-linguistic cognition and its role in language development. It is argued that children's capacity to categorize novel words and phrases as belonging to the syntactic categories "count noun", "mass noun", and "NP" is the result of innate links between language and thought, links that also exist as part of the adult's knowledge of language. In the first chapter, this view -- the "Semantic Competence Hypothesis" -- is sketched out, and contrasted with other proposals, such as distributional learning, cognitive categories, prosodic bootstrapping, and semantic bootstrapping.

In the second chapter, analyses of the spontaneous speech of 1- and 2-year-olds are presented. These suggest that even very young children categorize words like *dog* and *water* as nouns, and words like *Fred* and *she* as lexical NPs. Further, they appear to treat adjective-nominal strings (e.g., *big cookie*) as NPs, while treating nominal-adjective strings (e.g., *cookie big*) as sentences. Alternative, non-syntactic, explanations of the children's competence are discussed and argued to be descriptively inadequate.

In the third chapter, the question of how children syntactically categorize words like *Fred* and *dog* is addressed. It is argued that there exists a semantic level intermediate between grammar and non-linguistic cognition, and that children can exploit both the mapping between grammar and semantics, and the mapping between semantics and cognition, in the process of language development. With regard to the semantic representation of nominals, the view I present is an extension of the theory of lexical semantics developed by Jackendoff, Pinker, and others -- with the added empirical claim that one of the semantic features (what Jackendoff calls THING) is always expressed as an NP. Some apparent anomalies, such as expletives, idiom chunks, and predicational NPs, are also discussed.

The fourth chapter concerns the distinction between count nouns (words like *dog*) and mass nouns (words like *water*). It is proposed that this grammatical distinction maps onto a semantic contrast, i.e., [+/- INDIVIDUAL]. This semantic contrast relates to the cognitive notion of "discrete bounded entity", which explains why names for objects tend to be count nouns and names for non-solid

substances tend to be mass nouns. But the semantic nature of the count/mass distinction also extends to other domains. In particular, the theory predicts that mass superordinates like *furniture* should have subtly different cognitive properties than count superordinates like *animal*. A second claim is that adults should be sensitive to the semantic properties of count/mass syntax even for nouns describing non-material entities, such as sounds and sensations. Both predictions were supported in studies of the intuitions of naive adult subjects.

In the fifth chapter, I critically examine several studies that purport to show that young children initially treat the count/mass distinction as semantically arbitrary. It turns out that these claims are based on either a mistaken characterization of semantic theory, or on problematic experimental results.

In the final chapter, I present a theory of how children acquire the language-specific aspects of the count-mass distinction -- which words are count nouns, which words are mass nouns, and what the selectional properties of specific determiners are. One prediction of this theory is that even very young children should be sensitive to the semantic correlates of count/mass syntax. To test this, children's early usage of mass nouns that name objects (e.g., *furniture*) was examined. It was found that children miscategorized these words as count nouns far more frequently than they miscategorized mass nouns that name non-solid substances (e.g., *water*). Finally, two experimental studies suggest that 3- and 4-year-old's inferences about an ambiguous entity -- even one that is non-material, such as a sequence of sounds -- is affected by whether the entity is described with a count noun or a mass noun. These studies suggest that young children are sensitive to the semantic properties of count/mass syntax.

Thesis Supervisor: Dr. Susan Carey
Title: Professor of Brain and Cognitive Sciences

Acknowledgements

I am especially grateful to three people. Susan Carey has been unstinting in her intellectual guidance throughout my graduate career. She possesses the ideal combination of creativity and clear-headedness, and discussions with her have shaped my views to a striking degree. My interactions with Steve Pinker have been immensely valuable, and co-authoring a paper with him was terrific fun. Both Susan and Steve have made extensive, extremely helpful, comments on the style and content of this thesis. More importantly, they have had a significant influence on my own intellectual style and content. And it was having John Macnamara as a teacher and a friend, and being influenced by his ideas and his enthusiasm, that led me to enter this field.

Jane Grimshaw has been my linguistics guru; she is a wonderful person to talk with and a fine drinking buddy. Discussions with Ken Wexler have forced me to reconsider my views on many matters, and have added considerable intellectual stimulation to my graduate career. I am grateful to both of them.

Jane Grimshaw, John Kim, Sandeep Prasada, and especially Nancy Soja have all provided very helpful comments on earlier drafts of this thesis. Many of the ideas within have been developed through discussions and correspondence with Melissa Bowerman, Lila Gleitman, Roberta Golinkoff, John Macnamara, Jerry Samet, Virginia Valian, Sandy Waxman, and Rob Wilson. Ray Jackendoff and Ronald Langacker have been kind enough to provide me with their unpublished work.

Ned Block, Paul Horwich, Paul Pietroski and especially Ed Stein have greatly affected my views on matters philosophical, while Leda Cosmides and John Tooby have lured me into worrying about the evolution of mind and language.

Pat Claffey has helped me find countless books and articles, and Jan Ellertsen has patiently guided me through various administrative and financial labyrinths.

My first studies on children's understanding of count/mass were done in collaboration with Renee Oatway, and her ideas influenced much of the discussion in Chapter 5. Patrick Donovan has provided skillful assistance with the experiments presented in Chapter 6. I am also grateful to the children and teachers of Agassiz Day Care, Beacon Hill Nursery School, Bright Horizons Day Care, and Children's Cooperative for their kind cooperation.

Though it's easy to forget, there's more to life than the relentless search for truth, and I owe a lot to all of my friends who kept me busy with arguments, drinking, squash, tennis, and awful movies. This includes a lot of people, but

especially Roberta Brawer, Jacob Feldman, Ezra Samet, Jerry Samet, Scott Smith, Nancy Soja, Ed Stein, Karin Stromswold, Mike Tarr, and Debbie Zaitchik.

I am very grateful to my family, both in Boston and Montreal, for their love and support.

Most of all, I thank Karen Wynn. She has provided me with extensive comments and advice through every stage of this thesis, from the conceptual structure of the arguments to the design of the experiments to the wording of sentences in the final draft. She has been a constant source of encouragement and love; she is a friend, a colleague, and much much more. Since this thesis is about meaning, it is only fitting that I dedicate it to Karen, because she has brought meaning into my life.

Table of Contents

Chapter 1

Semantic competence and language acquisition

> It would be absurd to suppose that this innate linguistic theory, which determines the general form and structure of language, should not be connected in the most intimate manner to the fundamental properties of meaning and language use. -- Noam Chomsky, *Language and Responsibility* (1977), p. 140.

1.1. Introduction

Humans have the capacity to understand the meanings of novel sentences. This simple fact sets the stage for a scientific theory of semantics, which studies the part of the mind that underlies what we can call, following Creswell (1978), "semantic competence". This thesis concerns one aspect of the relation between language and thought – the correspondence between syntactic structure and cognitive structure -- and its role in language development.

While the precise nature of the correspondence is a matter of lively debate, there is no doubt that systematic links between grammar and cognition exist (see Enc, 1988 for discussion). So *Mary kissed John* and *John kissed Mary* have the same words but mean different things, presumably because the words occupy different syntactic positions. Or consider anaphora: *John thinks Fred likes himself* means that John thinks Fred likes Fred, not that John thinks Fred likes John, and this is due to the structural relationships that hold between the NPs *himself*, *Fred*, and *John* (e.g., Chomsky, 1981). Or predication: *John visited his friends drunk* means that John was

drunk, but *John visited his drunk friends* means that John's friends were drunk (e.g., Williams, 1980). In all of these cases, there is a systematic relationship between some cognitive aspect of sentence understanding – the different roles that participants play in the description of an action, the different possible referents of a pronoun, the way in which properties should be attributed -- and the syntactic structure of the sentence.

These sorts of correspondences have been extensively studied for verbs, and there are intricate theories of both the mappings between verb meaning and grammatical structure (e.g., Dowty, 1989; Grimshaw, 1990; Jackendoff, 1983, 1987, 1990a; Levin and Rappaport, 1986; Talmy, 1975, 1983, 1985), and of the roles that these mappings play in language acquisition (e.g., Gleitman, 1990; Pinker, 1989; Pye, 1990). Two ideas from this domain have proven particularly fruitful and I will adopt (and defend) them in the chapters to follow. The first is that much of the necessary representation is unlearned. Thus both the cognitive primitives and the syntactic primitives are innate; they are not somehow induced from experience. Further, the mappings from thought to grammar are highly constrained, leaving only minor variation to be acquired by children. The second idea is that there is an intermediate level between syntactic structure and cognition -- a "semantic level" with its own properties and primitives. Positing this level allows for a better explanation of the systematicities that hold between language and thought than would otherwise be possible.

In sum, it is clear that knowledge of language includes semantic competence, and that some of semantic competence involves the representation of mappings between grammar and thought. More controversially, at least some of these mappings are part of our innate linguistic endowment, wired into the language faculty by natural selection, not learning.

Consider now what might appear to be an entirely different issue. Jerry Fodor, in a paper written in 1966, was the first to point out that no matter how much of syntactic knowledge is innate, children must learn which words and phrases

correspond to which categories. So assume that the syntactic principle called the "Case Filter" is an innate part of grammar. This principle states that all NPs must be in a certain syntactic configuration, and it rules out sentences such as **John to leave would be a shame,* where *John* is not case-marked. But in order for children to apply the Case Filter they have to identify NPs in the environment. Thus children learning English will hear sentences like *John threw the ball* and they somehow have to figure out which linguistic categories these words and phrases belong to. There are no obvious psychophysical cues that could help the child -- nouns don't sound any different from verbs and they don't have any universal marking or position associated with them. So how does the child work out that *John* is an NP and *threw the ball* is a VP, instead of vice-versa?

Note that this problem exists even if there isn't any innate language-specific knowledge. It's uncontroversial that syntactic categories are needed to explain adult knowledge, so any adequate theory of language acquisition must explain how children syntactically categorize novel words and phrases. This is what Pinker (1984) has called the bootstrapping problem: how do children ever "break into" the language system; how can they get from strings of sounds into an abstract system of rules and principles defined over categories such as "NP" and "VP"?[1]

In the chapters that follow, I will defend what can be called the Semantic Competence Hypothesis (SCH). This is the view that children use their innate semantic competence to solve the acquisition problem. The key point of this proposal that distinguishes it from the semantic bootstrapping theory of Grimshaw (1981) and Pinker (1984) (see below) is that it posits that the mappings are part of knowledge of language, and not special to language acquisition. Thus the SCH can be divided into four sub-claims, with only the last distinct from the bootstrapping view.

---

[1]I will assume throughout that prior to the acquisition of grammar children can parse the speech string into at least some words, though they may not fully grasp their meanings. This assumption is supported by the fact that children use words in isolation long before they start to combine them into syntactic structures (see, e.g., Benedict, 1979).

- There exist systematic mappings between cognition and syntax.
- Knowledge of these mappings is innate.
- Children exploit these mappings to infer the linguistic categorization of new words.
- These mappings are part of semantic competence, and are not special to a language acquisition device.

There is a motivation for expecting the SCH to be true. Totally aside from issues of language acquisition, we know that semantic competence exists, that there are links between grammar and cognition that are part of the language faculty itself. It thus seems plausible that children could exploit these links when first learning the syntactic categorization of new words. This would be the simplest solution: the capacity to solve the acquisition problem would just "fall out" of knowledge of language itself.

Of course, just because a solution is the most elegant does not mean that it is correct. If it turns out that the cognition-grammar links are either (i) not present in very young children, or (ii) not sufficient to map words onto categories, then the SCH is wrong and some other solution must be found. In the following chapters, two domains are explored: the noun/NP distinction and the count/mass distinction. For each, it is argued that the SCH can explain how children determine the linguistic status of new words and can adequately characterize the developmental process.

1.2. Alternatives

Here I will briefly review alternative theories of how children solve the categorization problem. The point of this summary is just to show that the acquisition problem has not already been solved and that all existing theories have serious difficulties (see Pinker, 1987 for a similar conclusion). More precise versions of these alternatives will be discussed throughout the thesis.

1.2.1. Distributional learning

To the extent that developmental psycholinguistics has a party line, this is it (see, for example, the papers collected in Levy, Schlesinger, and Braine, 1988). This is the view, discussed in detail by Maratsos and Chalkley (1980), that children initially categorize parts of speech according to their "distributional properties", such as what words they go before, what words they go after, and their absolute position within an utterance. As a result of this analysis, children come to cluster words and phrases into appropriate linguistic classes, such as "noun" and "verb", "count noun" and "mass noun". Only later -- if ever -- do they grasp meaning differences between these linguistic classes. This is an old idea: distributional learning was called a "discovery procedure" by linguists in the 1940's, and Fodor (1966) himself tentatively adopted it as a solution to the acquisition problem.

There are serious problems with distributional theories, however. For one thing, nobody has the foggiest idea about how such a correlational analysis could work, and there have been no explicit theories proposed. Given the immense number of possible alternatives that an unconstrained analysis would have to sift through in order to converge on the correct adult categorizations (over 8 billion for the count/mass distinction - see Gordon, 1982), there is no reason to believe that such a procedure even exists. (It was for similar reasons that the notion of "discovery procedure" was abandoned in linguistic theory; see Chomsky, 1977). This problem becomes more vexing when one considers that children younger than 2-and-a-half appear to possess syntactic rules and principles (Pinker, 1984; Valian, 1986; see

chapter 2), and that categorization errors are vanishingly rare (Brown, 1973; Maratsos, 1982).

Furthermore, there is considerable evidence that even very young children are sensitive to the relationship between grammar and cognition. For example, Katz et al. (1974) and Gelman and Taylor (1984) have shown that 1- and 2-year-olds are sensitive to the semantic contrast between nouns and NPs when inferring the meaning of new words. Brown (1957) has shown that 3- to 5-year-old children understand that count nouns tend to describe objects and mass nouns tend to describe non-solid substances, and Soja (1990) has found that even children younger than 2-and-a-half are sensitive to a word's count-mass syntax when acquiring its meaning. Maratsos (1976) found that 3-year-olds are sensitive to the semantic distinction between definite and indefinite NPs. All of these examples concern the nominal system, but there is perhaps even more evidence for verbs. Children's errors in verb categorization suggest that they are sensitive to remarkably subtle semantic properties (Bowerman, 1982), and this has been supported in several experimental studies (e.g., Gropen, 1988; Pinker et al., 1987; see Pinker, 1989 for a review). Moreover, even children younger than two are sensitive to grammatical structure when inferring the meanings of new verbs (Gleitman, 1990; Naigles, 1990). In light of all this, the distributional view seems rather implausible.

There is a range of motivations for distributional theories, perhaps the major one being the general theoretical assumption that language-learning is best viewed as a special case of the child's ability to solve formal puzzles (see, e.g., Levy, 1988a, 1988b). One motivation is more specific, however, and goes as follows: Children can rapidly acquire semantically arbitrary distinctions, such as the gender contrast in a language like French (Levy, 1988b). So children can learn at least some distinctions without the aid of semantics. And since they can use a distributional analysis some of the time, perhaps they use it all of the time.

But there is a slippery move here, involving a blurring of the distinction between

"non-semantic" and "distributional". Children succeed at acquiring contrasts like gender, which shows that they are not limited to categorizing linguistic forms on the basis of meaning. One should be careful, however, not to confuse the fact that children can acquire semantically arbitrary distinctions with the view that they must do so through a surface distributional analysis. It seems more likely, given the cross-linguistic similarities in phenomena like subject-verb agreement and gender marking, that children acquire these semantically arbitrary properties of language through highly constrained analyses of the syntactic tree. For instance, children may attend to agreement markers on adjectives because they are disposed to search for morphological regularities between adjectives and nouns -- not because they are analyzing all possible relations among words within a utterance. Moreover, they might only be able to acquire this sort of semantically arbitrary regularity after they have used cognition-grammar mappings to determine which words are nouns and which words are adjectives in the first place.

Finally, one implication of the distributional approach is that there are no biologically specified universals of grammar. Under this view, children do not start off with categories such as "noun", "verb", etc. Rather, they converge on these categories through a correlational analysis of adult speech. As a result, there can be no innate knowledge that is special to these categories, no knowledge of the sort "All NPs must satisfy the Case Filter" or "Noun inflections encode plurality" or "NPs are bounding nodes", etc. For the same reason, there can be no innate relationship between cognition and grammar. It is historical accident, then, that nouns (and not modals) are names for objects, that languages form questions by movement and not by reversing the order of words, that there is subject-verb agreement and not subject-preposition agreement, and so on. From the standpoint of the child, one correlation is no better or worse than other.

All of this appears to be mistaken. There is strong evidence that children innately possess abstract constraints on the nature of grammatical structure and on the relationship between cognition and grammar. This has been found through

detailed linguistic analyses of particular languages (e.g., Bresnan, 1982; Chomsky, 1986), by cross-linguistic studies (e.g., Bybee, 1985; Greenberg, 1966; Talmy, 1985), and analyses of language acquisition (e.g., Pinker, 1984, 1989; Slobin, 1985).

1.2.2. Prosodic information

A different possibility is that children use prosodic cues to break up the sentences that they hear into distinct phrases (see Gleitman et al., 1987 for a review). Phrase boundaries are often marked by changes in prosodic structure (e.g., Cooper and Paccia-Cooper, 1980), and there is recent evidence that infants are sensitive to the psychophysical correlates of phrase boundaries -- 7- to 10-month-olds prefer to listen to an excerpt from mother's speech where there is a pause between phrases than where there is a pause within phrases (Hirsh-Pasek et al., 1987).

One interesting aspect of this proposal is the notion that phrases are developmentally prior to words, which runs contrary to the position that children first learn the syntax of words and then use principles of grammar to infer phrasal structure. In the next chapter, I present evidence suggesting that this view is correct within the nominal system -- under certain circumstances, NPs are syntactically categorized prior to nouns. As for prosody itself, there are two difficulties.

One general critical point was made by Pinker (1984, 1987). For prosody to be part of the solution to the acquisition problem, the relevant psychophysical correlates of phrase structure must exist across all languages. The evidence is scarce, and to my knowledge there is no evidence at all for tonal languages or signed languages. Unlike the link from cognition to grammar, where the evidence for universal mappings is very robust, there is no strong reason to believe this claim for prosody.

In any case, the major difficulty with prosody is that it does not solve the acquisition problem. Suppose 1- and 2-year-olds can use prosody to infer phrase boundaries, and can thus infer that *the dog chased the cat* has two main phrases: *the dog* and *chased the cat*. But which categories do these phrases belong to? Prosody

itself tells the children nothing. Thus a phrasal parse can perhaps play a role in the solution, but it is not the solution itself.[2]

1.2.3. Conceptual categories

A different view is that children start off by categorizing words according to the perceptual/cognitive categories that they belong to, categories like "object word" and "action word". These are clearly not equivalent to adult linguistic categories and thus proponents of this sort of theory have to argue that somehow the child's immature categories get transformed into the proper syntactic ones. So "object word" somehow gets transformed into "noun" and "action word" somehow gets transformed into "verb". Schlesinger (1981: 230) sums up this view as follows:

> ... the child's earliest semantic categories are narrowly circumscribed and ... he extends them gradually into the broader categories of the adult grammar through semantic assimilation.

This theory is very different from the Semantic Competence Hypothesis. The SCH posits that the child starts off with a syntactic level of representation, which is linked to non-linguistic cognition. Schlesinger's view is that there are no innate grammatical categories, that the adult grammatical representation emerges as the result of the shifting and elaboration of narrow conceptual classes.

One could argue, following Fodor (1975), that the whole notion of "semantic

---

[2]Gleitman (1990: 37) cites Joshi and Levy (1982) as providing "evidence that much of labeling, or its equivalent, can be derived from 'skeletal' representations in which there are configurations but no overt labels." But in fact, Joshi and Levy only show that skeletal representations are in certain senses equivalent to context-free phrase structure rules: they do not show that one can correctly label the nodes of a skeletal representation in the absence of any other syntactic information. Indeed, no strictly formal approach could ever provide such an analysis, as it would depend on certain empirical assumptions about grammar (is the VP universal? what sort of movement is permitted? what sort of X-bar constraints exist?, etc.) Moreover, even if one already knows the language, it is likely that there are skeletal representations that can be described in multiple ways, e.g., as both an NP and an AP. One insight of modern linguistics is that different phrasal categories have very similar (if not identical) syntactic configurations (e.g., Abney, 1987). If so, this could make the problem faced by children far more difficult, at least if they could not use semantic information.

assimilation" -- of a category becoming more abstract as a result of experience -- is conceptually incoherent. However, there are well-documented cases where children's categories *do* change (e.g., the child's concept of "animal", see Carey, 1986), and similar cases with adults (e.g., the shift in the notion of "species" in the 19th century, see Mayr, 1982). Whatever unknown mechanism is responsible for cognitive change in these cases might also cause changes in the child's linguistic categories. Nevertheless, while within the realm of possibility, the absence of any clear theory of how the child's categories get more abstract makes it impossible to evaluate this view.

There are other difficulties. For one thing, this theory, along with the distributional theory, posits no innate syntactic categories and is at a loss to explain cross-linguistic universals. For another, even very young children use nouns that do not describe objects, verbs that do not describe actions, and so on. Similarly, as Maratsos (1982) has argued, children do not make the errors that one would expect if they were limited to narrow linguistic categories. For instance, nouns that seemingly name events (e.g., *nap*) are not miscategorized as verbs, even by very young children (see Macnamara, 1982 for discussion).

1.2.4. Semantic bootstrapping

The phrase "semantic bootstrapping" is used in two different ways. Most generally, it is the claim that children use semantics to infer grammatical structure. Thus they use the fact that dogs are objects to infer that *dog* is a noun, that running is an event to infer that *runs* is a verb, and so on. Under this reading, the SCH proposed here is a particular version of semantic bootstrapping. But there is another version of "semantic bootstrapping" that is rather different from the SCH, and thus is worth discussing.

Grimshaw (1981) and Pinker (1984), who developed this view, are both at pains to argue that the form-meaning correspondences are *not* part of knowledge of language. Rather, they are the result of special mechanisms of language

acquisition. Thus Grimshaw (1981: 171) posits as a principle of UG (Universal Grammar) that "form and function are independent".[3] Given this, whatever correspondences do exist are because "the LAD [Language Acquisition Device] ... gives priority to a grammar with a one-to-one correspondence between form and function." Finally, Grimshaw begins the section introducing the semantic bootstrapping hypothesis with the claim "It is universally agreed by linguists that the syntactic categories of a language are defined in structural and not semantic terms." Similarly, Pinker (1984: 39) introduces the theory by stating, "Although grammatical entities do not have semantic definitions in adult grammars, it is possible that such entities refer to identifiable semantic classes in parent-child discourse. ...", and he notes later that these correspondences are only sure to hold in "basic sentences", not in passives, contextually dependent sentences, etc. (but see Pinker, 1989 for a different proposal).

It is over these points that the SCH is in conflict with the bootstrapping hypothesis. The SCH is the view that the form-function mappings exploited by children are actually part of knowledge of language, not part of the language acquisition device. Defending this will be the burden of the following chapters, but it is worth pointing out that there are some reasons to doubt the bootstrapping theory.

First, the semantic bootstrapping theory makes a very strong prediction, that the only systematic relation between cognition and grammatical categories will occur with classes of word meaning that are somehow related to perception, e.g., "names for objects are nouns". But this is false; it is hardly a matter of chance that words like *day* and *conference* are nouns and not prepositions or modals. The same words appear as nouns across different languages and, more importantly, if an adult speaker was taught these words without syntactic support, he or she would

---

[3]This view is not unanimously held. Thus Chomsky (1957: 108) states "we do find many important correlations, quite naturally, between syntactic structure and meaning", and Jackendoff (1983, 1990a) provides an extensive discussion of such correlations within the framework of generative grammar.

correctly categorize them. In general, there are clear cognition-grammar relations that go well beyond the perceptual domain, and these show up cross-linguistically and in the intuitions of adults categorizing novel words.

Second, although the generalizations described by the theory are descriptively adequate, there is a large explanatory gap. By simply stipulating the perception-syntax mappings, the bootstrapping theorist cannot explain why these mappings occur and not others. So there is no theory of why the mapping is from object names to nouns instead of (say) from object names to modals. Would it be a possible language if the mappings led children to categorize *dog* as a modal and *hit* as a quantifier? It seems unlikely; for one thing, other aspects of the language faculty, such as structural constraints on co-reference, could not naturally apply. The point is, so long as the form-function mappings exploited in language acquisition are presumed to be distinct from the rest of linguistic theory, systematic relationships between the role of semantics in language acquisition and the role of semantics as part of the language faculty itself are left unexplained. This motivates capturing the (correct and useful) generalizations discussed by Grimshaw and Pinker in terms of a richer theory of linguistic semantics.

Finally, there is an argument by Gleitman (1990) which was directed against semantic bootstrapping, but also applies against the SCH. I will discuss it in detail in Chapter 3, but for the sake of completeness I will briefly mention it here. Gleitman notes the tremendous poverty-of-stimulus problem in the acquisition of word meaning, and discusses the many difficulties faced by the child when acquiring new words (see also Carey, 1982). She makes the suggestion that at least some word meanings are not acquired by observation at all, but through information inferred from the syntactic structures that the words appear in ("syntactic bootstrapping"; see also Landau and Gleitman, 1985; Naigles, 1990). And if children need to know the syntactic category of a word to infer its meaning,

then they certainly cannot be using the word's meaning to infer its syntax.[4]

1.3. A note on the autonomy of syntax

One motivation for semantic bootstrapping, made explicit by Grimshaw (1981), is to move form-function correlations out of the language faculty and into the language acquisition device. And the motivation for this is the autonomy of syntax: "It is a consequence of the autonomy of syntax that syntactic form and semantic type will not be in one-to-one correspondence in any principled way" (Grimshaw 1981: 167).

There is considerable motivation to accept the autonomy thesis (see the arguments in Grimshaw, 1979, for instance), but it is not clear that the existence of syntax-semantics mappings would violate it. Grimshaw advances a very strong version of autonomy -- where syntactic structure can have no relationship to semantics at all. But there are milder versions, one being the view that "rules of syntax ... are organized in terms of 'autonomous' principles of mental computation and do not reflect in any simple way the properties of the phonetic or semantic 'substance' or contingencies of language use." (Chomsky, 1980: 213). Under this view, the SCH is consistent with autonomy. Even if all NPs share a common semantic feature (which is probably too strong a claim; see Chapter 3), there could still be properties of syntax that apply independently of this definition, e.g., all NPs must satisfy the Case Filter.

More generally, just because a primitive at one level can be defined at a lower level, this does not necessarily mean that the levels can be collapsed. The situation would be different if *all* the primitives at one level mapped one-to-one with all the primitives at another level -- because then there would be no motivation for

---

[4]Gleitman herself actually has a more moderate position, where both forms of inference (semantics -> syntax, and syntax -> semantics) play some role. Nevertheless, this sort of radical "syntax-first" view is worth considering, as it makes explicit some important aspects of the acquisition problem.

positing distinct levels in the first place. But a partial mapping is not inconsistent with autonomy. The fact that we know the precise chemical composition of certain neurotransmitters surely does not imply that neuroscience is superfluous and should be reduced to chemistry.

In the case of language, there are phenomena that can only be adequately explained if one posits a syntactic level of mental representation, where linguistic generalizations apply independently of semantics, phonology, etc. In fact, as Newmeyer (1990) argues, there are excellent reasons to expect such an autonomous level of language to exist, given the functions for which language has evolved. The role of syntax is not just to convey thought; it is to bridge the gap between cognition and phonology, and thus requires special properties of its own. This is similar to why one needs intermediate levels of visual processing: something must bridge the gap between the retinal array and the higher-levels of brain functioning. Given the evolutionary motivation for autonomy, and its parallels in other mental faculties, it is a rather reasonable assumption about the nature of language.

1.4. Outline of thesis

In the next chapter, I will discuss some phenomena in child language which strongly suggest that 1- and 2-year-old children possess the syntactic categories "noun" and "NP". In Chapter 3, a theory of the semantic basis of nouns and NPs will be presented. It will be argued that there are innate links from non-linguistic cognition to a semantic level, and from the semantic level to grammar. Although the mapping between the grammatical category and the semantic level is not one-to-one (i.e., it is not the case that all and only NPs have a certain semantic feature), it is sufficiently informative so that children can exploit the relationship between semantics and grammar to determine which words and phrases are nouns and which are NPs.

In Chapter 4, I develop a theory of the count/mass distinction, and present some empirical studies of adult knowledge suggesting that the count/mass

categorization of different nouns can be partially explained in terms of their semantic properties. In Chapters 5, arguments that children do not treat this distinction as semantic are discussed and found to be based on either a mistaken characterization of semantic theory or on problematic experimental results. In Chapter 6, I present a theory of how children acquire the language-specific aspects of the count-mass distinction -- which words are count nouns, which words are mass nouns, and what the selectional properties of the determiners are. This theory is supported by analyses of the spontaneous speech of 2- to 5-year-olds and by experimental studies with 3- and 4-year-olds.

With the possible exception of the final chapter, all of the chapters can be read independently from one another. This is at the cost of some slight redundancy, and anyone who plans to go through this thesis from cover to cover should prepare for an occasional case of *deja vu*.

Chapter 2

Nouns and NPs in child language

2.1. Introduction

In recent times, Chomsky has resurrected the Rationalist position that knowledge of language is mostly unlearned, part of the specific nature of humans (e.g., Chomsky, 1986). Under this view, language learning (in whatever sense one could still view it as "learning") mostly reduces to the acquisition of minor variants on universal principles; this is sometimes conceptualized as the setting of parameters. One tenet of this theory is that the representational format is innate. For instance, children must learn whether their target language is head-first or head-final; this will determine, among other things, whether the language has verb-object order like English or object-verb order like Japanese. But the notions of "head" and "complement", "verb" and "object", are themselves innate, and not built up from prior non-linguistic knowledge or somehow abstracted from experience.

The alternative, which is the orthodox view within psychology, is that the linguistic knowledge of adults emerges through the interaction of non-linguistic learning mechanisms and the environment. For instance, Skinner argued that language is shaped primarily through principles of reinforcement, generalization, and discrimination (Skinner, 1957), while Piaget and his colleagues have suggested that the emergence of grammar is the by-product of cognitive development (Piaget, 1980). More recently, scholars have stressed, among other things, children's desire to communicate (Bates and MacWhinney, 1989), their ability to construct

hierarchical structures (O'Grady, 1987), and their capacity to recognize statistical generalizations (Levy, 1988b). Despite their other differences, these scholars are united in the belief that there is no innate capacity special to language, no language-organ or module that all normal humans are born with. Two of the theories discussed in the previous chapter -- distributional learning and cognitive categories -- emerge from this tradition, as they hold that the language-specific knowledge of adults is the product of some non-linguistic learning mechanism.

In light of this debate, the language of young children becomes interesting, because the two kinds of theories make strikingly different predictions about its nature. The nativists argue that from the very start of language acquisition, children assign words to syntactic categories and, having done so, order words through rules and principles that manipulate such abstract categories (e.g., Chomsky, 1986; Hyams, 1983; Pinker, 1984, Valian, 1986).[5] It follows from the empiricist view that children will initially start off with some representation different from adults, one that reflects a more natural untutored bias. So children may start off with rules that order classes like "agent" and "action" (conceptual categories), or by utilizing patterns of associated words learned by attending to adult speech (distributional learning). Only much later -- perhaps by age four or five -- will children come to possess the same linguistic categories that adults do (e.g., Bates and MacWhinney, 1987, 1990; Bowerman, 1973; Braine, 1976; Gathercole, 1986; Macnamara, 1982; Maratsos, 1983; Schlesinger, 1981, 1988).

This chapter will provide some support for the nativist view, by presenting evidence that 2-year-olds possess at least some abstract linguistic categories: nouns

---

[5]Importantly, this does not imply that children's first word combinations need follow these principles, since language-specific neural structures might mature sometime after the start of spontaneous speech (see Gleitman, 1981 for discussion). After all, other unlearned capacities (like bladder control, the ability to walk, stereo vision, etc.) mature after birth and there is no *a priori* reason to reject the view that the language capacity matures around the second year of life. Since very young children might start to combine words before this faculty matures, the existence of a "non-linguistic" phase of speech production is quite plausible. Some scholars have argued that different aspects of linguistic knowledge mature at different times, so that 4-year-olds might have some, but not all, aspects of Universal Grammar (e.g., Borer and Wexler, 1987).

and NPs. Although the same data suggest that adjectives and sentences are also present in the grammars of young children, the focus of the chapter is on the noun/NP distinction and these other categories will be discussed only in passing. In the next chapter, a semantic theory of how children determine which words are nouns and which are NPs will be presented.

2.2. The phenomena

What should we expect to find if young children's grammars contain nouns and NPs? One way to approach this is to ask how adult grammars distinguish the two, and then look for the same distinctions in the speech of children.

One phenomenon is that, in English, adjectives precede nouns to form NPs, but cannot precede NPs to form more NPs. Thus *Those are big dogs* is fine, but **Those are big my dogs* and **That is big John's dog* are not, presumably because *my dogs* and *John's dog* are already NPs and cannot further co-occur with prenominal adjectives. Under the assumption that pronouns and proper names are lexical NPs, we can also capture the fact that sentences like **There is big Fred* and **There is big he* are ungrammatical. Adjectives can only modify NPs when they appear in predicate position, as in *My dogs are big* and *Fred/He is big*. The relevant contrast between nouns and NPs is summarized in (1)-(4) below.

(1) big dogs
(2) *big the dog
  *big my dog
  *big Fred
  *big he
(3) Dogs are big
(4) The dog is big
  My dog is big
  Fred is big
  He is big

There are some exceptions to the generalization that adjectives cannot precede

pronouns and proper names. There are nicknames (*Proud Mary*), and certain exclamations (*Good Fifi!, Bad Freddie!, poor you!, silly me!*). These exclamations only occur in a pragmatically restricted range of contexts and are rare in adult-to-adult speech, although they are more frequent in adult speech to children. A second class of exceptions occurs when a proper name is used to pick out a class of entities that are contrasted over space (e.g., *Big John is over there and little John is over here*), or time (e.g., *You used to be little Joey, and now you're big Joey*). But once again, these sound rather odd and are only acceptable when the proper name is given a special interpretation (see Section 2.7 for discussion).

Putting these cases aside, we can characterize the adult competence as below. For the purposes here, "prenominal adjectives" describes those that co-occur with a nominal to form an NP (e.g., *cold water*), while "predicate adjective phrases (APs)" co-occur with nominals to form sentences (e.g., *the water is cold*).

- The English words *dog, water*, etc. are nouns; *she, Fred*, etc. are NPs.
- Predicate APs appear with NPs.
- Prenominal adjectives appear with nouns, not NPs.

These properties are a matter of empirical fact, and do not follow from the definitions of "prenominal adjective" and "predicate AP". It could have been, for instance, that English speakers could combine predicate adjectives with nouns to form sentences of the form [N - AP], or that they could combine prenominal adjectives with NPs to form NPs of the form [A - NP] -- it just happens otherwise.

If children act in accord with these properties of English, this suggests that they possess the categories "noun" and "NP". This ia not to imply that these particular generalizations are universal, but only that in order for children to order words in accord with these aspects of language (whatever their status), the children must know which words are nouns and which are NPs, and must use these words in different syntactic contexts. Thus showing that children obey these generalizations will provide some support for a nativist theory of language acquisition.

In the following three sections, the speech of several young children is analyzed. In the first analysis, children's use of prenominal adjectives and predicate adjectives is compared. It is found that while children frequently say things like *big dogs, dogs big,* and *he big,* they almost never say anything like *big Fred* or *big he.* The second analysis shows that such violations are rare even in children's very first word combinations. The third analysis presents some evidence that 2-year-olds only use adjective-nominal sequences (*big dogs*), not nominal-adjective sequences (*dogs big*), in positions which select for an NP in the adult grammar, such as the object position of a verb like *hit.* This suggests that children mentally represent sequences like *big dogs* as NPs and sequences like *dogs big* as sentences. Some non-syntactic theories of the children's knowledge are discussed and it is argued that these alternatives fail to account for certain properties of child language. Instead, the best explanation for the child's competence is in terms of the categories "noun" and "NP".

2.3. Study 1: Use of prenominal and predicate adjectives

2.3.1. Subjects

The children whose speech was analyzed were Adam, Eve, and Sarah (see Brown, 1973), and Peter (see Bloom and Lahey, 1978). These corpora are stored on computer text files as part of the Child Language Data Exchange System -- CHILDES (MacWhinney and Snow, 1985). The ages of the children ranged from 1;6 to 2;10 and their MLUs ranged from 1.04 to 2.39 (see Table 2.1).

Table 2.1 -- Summary of transcripts

Child	Age range	MLU range	Samples
Adam	2;3 to 2;6	2.00 to 2.25	8 2-hour samples
Eve	1;6 to 1;9	1.50 to 2.25	8 2-hour samples
Sarah	2;3 to 2;10	1.75 to 2.25	34 1-hour samples
Peter	1;9 to 2;1	1.04 to 2.39	8 3-hour samples

2.3.2. Sampling Method

To get a large and representative sample, I searched the transcripts for 43 adjectives found in the spontaneous speech of 2-year old children (taken from the studies of Nelson, 1976 and Valian, 1986) (see Table 2.2). All of these adjectives are acceptable in both prenominal and predicate positions; there were no adjectives like *afraid* (**the afraid man*) or *former* (**the president was former*).

Table 2.2 -- Adjectives used in transcript search

bad	happy	silly
bare	hot	sleepy
big	hungry	smart
black	little	soft
blue	long	stinky
broken	missing	tall
brown	new	ticklish
cold	nice	tight
colored	old	tiny
cool	pink	unhappy
dead	pretty	white
different	purple	yellow
dirty	red	yummy
good	sharp	
green	sick	

The utterances were extracted from the corpora by means of a computer search program. If an utterance was a verbatim repetition of what an adult just said, it was not included in the analyses. Some utterances were bracketed in the transcripts, which signals that the scorer was not sure if this was what the child was actually saying. In these cases, context was used to determine whether the interpretation was plausible and if it was, the utterance was included. For adjective-nominal

sequences, both those that occur in isolation (e.g., *little kitty, big dog*) and those that appear in a larger context (e.g., *little boy sleeping, see nice kitty*) are included. For nominal-adjective sequences, most utterances were two words long, although the few with a copula (e.g., *this is dirty*) or with expanded nominals (e.g., *Eve tapioca cold*) were included in the analysis.

All the analyses below concern utterance types, not tokens. A second set of analyses was performed on utterance tokens. For each token analysis, the results were more statistically significant than for the corresponding type analysis, so they will not be discussed further. Finally, use of the terms "adjective" and "nominal" is not meant to prejudge the issue; they should be viewed as shorthand for "words that are categorized in the adult grammar as adjectives/nominals".

2.3.3. Results

The purpose of the study is to see if children use pronouns and proper names following adjectives. The results are shown in Table 2.3, with some examples of what the children said shown in Table 2.4.

Table 2.3 -- Proportion of sequences where the nominal is a pronoun or a proper name (P-PN) or a common noun (CN)

Child	Adjective-Nominal		Nominal-Adjective	
	P-PN	CN	P-PN	CN
Adam	4% (2)	96% (44)	40% (6)	60% (9) *
Eve	0% (-)	100% (15)	57% (4)	43% (3) *
Sarah	1% (1)	99% (88)	87% (20)	13% (3) **
Peter	7% (2)	93% (25)	100% (6)	0% (0) **
TOTAL	3% (5)	97% (172)	71% (36)	29% (15)

* $p < 0.005$, ** $p < 0.00005$.

Table 2.4 -- Examples with pronouns and proper names (P-PN) and with common nouns (CN)

Adjective-Nominal		Nominal-Adjective	
P-PN	CN	P-PN	CN
Big strong Cromer	Dirty diaper	Bambi green	Pillow dirty
Broken dis?	Hot water	I sleepy	Girl bad
Big it!	Little cowboy	Mommy tired	Floor wet
It's my little Butch	He bad boy	Dat black	Tapioca hot
Oh bad Teppy	buy new shoes	He hungry	Car broken
right dere			

Over 70% of the children's nominal-adjective sequences had a pronoun or proper name as the nominal. In contrast, there were just five occurrences (3% of the total) of adjective-nominal sequences where the noun was not a common noun. These were the following: Adam: *Big Strong Cromer* (2;5), *Broken dis?* (2;6); Peter: *It's my little Butch* (1;10), *Big it!* (1;11); Sarah: *oh bad Teppy right dere* (2;7).

There are two reasons why the number of true violations may actually be smaller than five. First, immediately before Peter said *big it!*, both he and his father said *fix it*. So it is possible that the sequence is either a speech error or an error in transcription. Also, the utterance *bad Teppy* -- and, to a lesser extent, the other two sequences with adjectives followed by proper names -- is similar to a type of expression some adults use when talking to children: using a modifier before a person's name to express approval or disapproval towards that person. For instance, in the same session that Sarah said *Bad Teppy*, an adult said *Bad Mike* to a misbehaving dog. In light of this, it is not clear whether such a phrase is really a

violation, as opposed to an instance of an exceptional form that is modeled by adults. In fact, it can be taken as a striking demonstration of the child's competence -- apparently, when children hear adults use adjectives before NPs, they do not infer that adjective-NP sequences can be used freely across all contexts, but instead restrict them to the very same limited pragmatic contexts that adults do.

The difference between the use of pronouns and proper names within adjective-nominal sequences and their use in nominal-adjective sequences is significant for each of the children (Eve, $p = 0.0048$, Peter, $p = 0.00025$, Fisher Exact Tests; Adam, $\chi^2(1) = 12.62$, $p < 0.005$; Sarah, $\chi^2(1) = 82839$, $p < 0.00001$). In sum, children appear to have a systematic bias not to use pronouns and proper names with prenominal adjectives, but have no such bias with predicate adjectives.

2.4. Study 2: Use of prenominal adjectives in children's first word combinations

The first analysis compared use of prenominal adjectives with use of predicate adjectives. It is well known, however, that prenominal adjectives emerge before predicate adjectives in the linguistic development of English children (e.g., Bloom, 1970; Bowerman, 1973; but see Kim, 1987 for a different pattern in the acquisition of Korean). How are these first adjectives used? To study this, I analyzed the available data on English children's first word combinations.

2.4.1. Subjects

The speech of 10 children was analyzed: Kieran (from Macnamara, 1982), Kendall (from Bowerman, 1973), Eric, Gia, and Kathryn (from Bloom and Lahey, 1978), Jonathan (from Braine, 1976), Allison (from Bloom, 1973), June and April (from Higginson, 1985), and Naomi (from Sachs, 1983). The transcripts for Allison, June, April, and Naomi are stored on CHILDES (MacWhinney and Snow, 1985).

The children's ages are shown on the second column of Table 2.5. For each child, the available transcripts were analyzed one by one in ascending order by age until

at least 10 adjective-nominal sequences were found. Some of the children had fewer than 10 sequences altogether, but as long as there were at least 4 adjective-nominal sequences in the corpora, the child was included in the analyses.

2.4.2. Sampling Method

For the transcripts that were on CHILDES, I used a computer search program to extract the adjectives shown in Table 2.2. For the others, I went through them by hand and included as an "adjective" any prenominal modifier that is an adjective in the adult grammar. In all other regards the method was the same as for the first analysis.

2.4.3. Results

The results are shown in Table 2.5. Out of 125 utterances, there were 5 sequences (4%) where an adjective preceded a pronoun or proper name. These were the following: Kieran: *Old Mommy* (1;8), *Old Freddie* (1;8), Kendall: *red Kendall* (1;11), *blue Mommy* (1;11), and Eric: *little this* (1;11). The other seven children made no errors. Since all of these children used pronouns and proper names in other contexts, this suggests that young children understand the restriction against using pronouns and proper names following adjectives.[6]

---

[6]As in the first study, it is unclear how many of the exceptional utterances are really violations. Just as with Sarah's utterance of *bad Teppy*, it may be that Kieran was either using nicknames or was over-extending adult expressions like *bad Mike*. As for Kendall's utterances, they were used to describe pictures that had reddish and blueish casts (Bowerman, 1973: 291). Since young children frequently omit parts of sentences, it is possible that Kendall had meant to say *red picture of Kendall* and *blue picture of Mommy*. This is supported by the fact that in the same sample she had said *picture Kendall*, meaning "picture of Kendall" (see Bloom, 1990b for a discussion of omissions in child language).

Table 2.5 -- Adjective-nominal sequences from children's very first word combinations

Child	Age	Adjective-nominal sequences	
		Pronouns and PNs	Common Ns
Kieran	1;8	14% (2)	86% (12)
Kendall	1;11	40% (2)	60% (3)
Eric	1;10 - 2;1	8% (1)	92% (11)
Gia	1;10 - 1;11	0% (0)	100% (10)
Kathryn	1;9 - 1;10	0% (0)	100% (15)
Jonathan	1;11	0% (0)	100% (35)
Allison	1;7 - 2;6	0% (0)	100% (7)
June	1;6 - 1;9	0% (0)	100% (4)
April	1;10 - 2;1	0% (0)	100% (13)
Naomi	1;8 - 1;10	0% (0)	100% (10)
TOTAL:		4% (5)	96% (120)

2.5. Study 3: Adjective-nominal and nominal-adjective sequences in NP-positions

The studies reported above suggest that children distinguish pronouns and proper names from common nouns in their spontaneous speech -- only common nouns can follow adjectives. It was argued that this is because the children

categorize pronouns and proper names as lexical NPs and, in English, NPs cannot follow adjectives. It would be useful, however, to have evidence for the existence of the category "NP" in children's grammars that is distinct from evidence that they have the more narrow category "pronouns and proper names". This motivates the following study.

It is often argued that children don't encode phrases like *big dogs* differently from *dogs big* -- they are produced by similar or identical algorithms and once they are produced the children's grammar doesn't make any distinction between the two (e.g., Braine, 1976). The nativist alternative is that children have an adult-like competence; thus when children say *big dogs*, they are producing an NP, and when they say *dogs big*, they are producing a sentence, which adults would express as *the dogs are big* (note that young children frequently omit determiners and the copula from their spontaneous speech). The analysis below tests the prediction that children treat adjective-noun sequences -- and not noun-adjective sequences -- as NPs.

2.5.1. Analysis

There are certain positions where only NPs can go and sentences cannot. If children draw the NP/sentence distinction, then they should only use phrases like *big dogs* in these position, never phrases like *dogs big*. These positions include the object position of many verbs, the subjects of verbs, and the subjects and objects of predicative statements. This is shown in (5) - (8).

(5) Big dogs bite people
*Dogs big bite people
(6) I like big dogs
*I like dogs big
(7) Big dogs are mean
*Dogs big are mean
(8) Those are big dogs
*Those are dogs big

To test this, I analyzed all the utterances from Study 1 and Study 2 which contained an adjective and noun appearing in (what would be for adults) an NP-position, and recorded how many were adjective-noun sequences and how many were noun-adjective sequences. The prediction is that only adjective-noun sequences should appear.

2.5.2. Results

There were 54 utterances where a noun and an adjective appeared together in an NP-position. For all of these utterances, it was an adjective-noun sequence, not a noun-adjective sequence. In other words, children say things like *want big cookie* but they never say anything like *want cookie big*. Once again, this is exactly what one would expect if sequences like *big cookie* were NPs and sequences like *cookie big* were sentences in children's grammar.[7]

One might argue that children's competence can explained more simply. They could have a rule like the one below, where the bracketed constituents are optional, and therefore it is unnecessary to posit an NP constituent at all.

(adjective) - noun - verb - (adjective) - noun

However, this rule does not adequately explain the children's spontaneous speech, since not all of their utterances contain two nouns and a verb. Children use verbs without subjects (e.g., *want a cookie*) and without objects (e.g., *I fell*). This might suggest a representation more like the following:

(adjective) - (noun) - verb - (adjective) - (noun)

---

[7]Lyn Frazier has suggested to me that a more robust test would involve looking at verbs that take sentential complements; presumably these verbs should be followed by noun-adjective sequences (e.g., *I think dogs big*). Unfortunately, I found no utterances where such verbs were followed by nominals and adjectives, in any order. However, I would predict that as soon as such utterances appear in child language, they should be in noun-adjective order.

But this is also inadequate, because it permits sequences that children do not produce, such as adjective-verb-adjective and adjective-verb-noun. A better account would capture the fact that children process the pre-verb and post-verb adjective-noun pairs as distinct units, as shown below:

((adjective) - noun) - verb - ((adjective) - noun)

This brings us full circle, since the "units" containing adjectives and nouns are equivalent to NPs in the adult grammar, i.e., a verb has a subject NP and may have an object NP, and these NPs can expand to an adjective and a noun. The fact that the utterances of 1- and 2-year-olds have the same hierarchical structure as those of adults is further evidence that they have adult-like grammars -- and, more specifically, that their grammars contain the categories "noun" and "NP" (see also Valian, 1986 for similar analyses).

2.6. Alternative explanations of children's knowledge

Above, I proposed an account of the nature of the restriction against using prenominal adjectives with pronouns and proper names, which was based on the distinction between nouns and NPs. Are there alternative descriptions of the child's knowledge that do not involve the attribution of such categories?

2.6.1. Cognitive categories

One possibility is that children understand the restriction in terms of rules that order categories that play a role in perception or non-linguistic cognitive processes, but which are not special to language. The most explicit proposal that I have come across is that children have rules like the following, which are slightly simplified versions of those proposed by Macnamara (1982: 252). These rules would generate sequences like *big dogs* and *dogs big*, as well as *Fred big* and *I big*. However, neither of these rules would produce ungrammatical utterances like *big Fred* or *big he*.

modifier + object word

object word
proper name + modifier
pronoun

For two reasons, these rules cannot explain the adult competence. First, they lack even the rudiments of hierarchical structure. So they will produce *big dogs* only as a complete utterance, while the adult grammar allows this structure to appear as the subject of a verb, the object of a verb or preposition, a predicate nominal, a fronted topic, part of a conjunct, and so on. The results of Study 3 show that this sort of systematicity appears in the speech of even very young children, suggesting that the rules above are inadequate as a theory of the children's knowledge (see also Pinker, 1984; Valian, 1986).

Second, at least for adults, the categories themselves are inadequate. Categories such as "object word" and "modifier" (i.e., the class of words that denote properties) play no role in any adequate theory of natural language. Macnamara is aware of this, and is explicitly positing a "discontinuous theory", one where the child's immature categories are later transformed into the linguistic categories of adults. In Chapter 5, I will discuss a similar class of theories proposed for the count/mass distinction.

One could object that these categories are not entirely immature. In particular, "pronoun" plays a role in another aspect of linguistic theory, the aspect which is concerned with explaining patterns of acceptable co-reference relationships in the adult grammar. Thus *John thinks Bill likes him* can mean that John thinks that Bill likes John, but not John thinks that Bill likes Bill, while *John thinks Bill likes himself* shows the opposite pattern. Systematic grammatical principles govern the differential patterns of pronouns, reflexives, and referring expressions (see e.g., Chomsky, 1981). Thus there is a sense in which the category "pronoun" *is* part of the adult grammar. Nevertheless, proper names have no special status in the

theory of co-reference relationships, patterning the same as other "R-expressions" such as definite descriptions (e.g., *the man*). And "object word" and "modifier" are not aspects of any part of the adult grammar.

In any case, even ignoring the issue of hierarchical structure, the rules proposed by Macnamara are clearly inadequate when it comes to explaining the child's competence. In particular, the 1- and 2-year-olds analyzed above allow words following adjectives that are best described as nouns, not "object words". There are three classes of words that children use which motivate abandoning Macnamara's description of their linguistic competence.

2.6.1.1. Substance words

Least importantly, some of the words they used describe substances or substance-like entities -- words that in the adult grammar would be categorized as mass nouns. The relevant utterances are shown below.

hot water
hot cereal
nice cheese
big bread
yellow hair
black hair
good cake
little coffee

This set of counter-examples could easily be accounted for without abandoning the notion of cognitive-perceptual categories, by replacing "object word" with "object word/substance word".

2.6.1.2. Abstract words

Some of the modifier-word combinations have an adjective followed by a word that does not describe either an object *or* a substance. These are more interesting, because they suggest that the categorical status of words that can follow adjectives cannot be defined in terms of categories that relate directly to perception. The cases are shown below.

broken side
big fire
little fall
big noise
nice nap
green light
big hug
nice time
new show

2.6.1.3. The word "One"

Finally, there is at least one word that has the grammatical privileges of a noun yet is not the name of *anything* -- this is the word *one*. In the adult understanding of word order, it behaves the same as nouns like *dog* and *water*; it can follow adjectives, determiners, it can take a relative clause, and so on. But it is not an 'object word'. Rather, it is a pro-form, appearing in sentences like *John had a cold beer and Mary had a warm one*, where *one* is co-referential with the noun *beer*.[8] Thus from a conceptual point of view, *one* falls into the same class as *Fred* and *he*, not *dog* and *water*, but from the standpoint of language-specific grammatical structure, *one* falls into the same class as common nouns, not pronouns and proper names. If children understand word order through rules ordering conceptual categories then, unlike adults, they should not produce utterances where an adjective precedes the word *one*, e.g., they should not say things like *big one*.

The results -- taken from the analyses above -- are as follows. There were a total of 34 utterances where an adjective is followed by *one*. Some examples are shown below.

---

[8]More precisely, it has been argued that *one* is a pro-N', it can substitute for a category intermediate between nouns and NPs (Lakoff, 1970; Jackendoff, 1977; see Radford, 1988 for discussion). If so, then *one* may actually be a lexical N' (or "N-bar"), not a noun. It is unclear whether 1- and 2-year-olds categorize *one* as a noun or as an N'. However, since both categories behave identically with regard to prenominal modifiers -- both can follow adjectives -- the distinction is irrelevant for the purposes here. So long as children use *one* following determiners and adjectives, they are not clearly treating the word as belonging to the same class as proper names and pronouns like *it* and *she*.

big one
blue ones
yellow one
That little one
this a new one
this one's dirty
Buy new one
I want whole one

Each of the children studied in the first analysis produced at least four such sequences (Adam: 4, Eve: 4, Sarah: 14, Peter: 6), and three of the younger children also produced some (Kieran: 1, Eric: 3, April: 2). All of this suggests that the children's understanding of word order is not expressed through rules ordering conceptual categories. The children's use of *one* is of particular interest, as it suggests that their restriction against producing utterances like *big it* cannot be simply expressed as "pronouns and proper names cannot appear after adjectives". Instead, children know that *NPs* cannot appear after adjectives, and thus they have no general restriction against using pronouns in this position -- so long as these pronouns are not NPs.

2.6.2. Pragmatic rules

A different possibility is that children have some pragmatic restriction against using adjectives before proper names. One suggestion is as follows: At least for adults, one role of prenominal adjectives is to specify the reference of a noun, not to provide new information. Since proper names usually have only a single referent, there is rarely any reason to use them with adjectives. This can be summarized as follows:

- Pronouns and proper names have specific reference.
- Prenominal adjectives are used to specify reference.
- Thus, prenominal adjectives will not be used with pronouns and proper names.

One obvious problem is that this account does not work for pronouns, which are frequently ambiguous. So the pragmatic hypothesis cannot explain why children would not say things like *big he* to pick out a particular male from a group of males. But a more serious problem is that 2-year-olds clearly do not order words in accord with the principles above. If a child wants to express the belief that a particular dog is big, he or she is far more likely to say *big dog* than to say *dog big*. In the first analysis, the vast majority of two word utterances used as comments were adjective-nominal sequences, not nominal-adjective sequences (see Braine, 1976: 76-77 for a similar observation). Given this, one would expect that when children want to make a remark about someone, they should produce sequences like *big Fred* or *big he*. Interestingly, they do not.

A related proposal (suggested to me by Melissa Bowerman) is that when children say something like *big dog*, they are actually attempting to say (something like) *THAT'S a big dog*, but are not producing the initial demonstrative pronoun. And so, the argument goes, it would be redundant for a child to use a pronoun or proper name following the adjective, since the initial null pronoun already picks out a particular individual or group of individuals. To put it another way, the initial null pronoun may be the topic and the pragmatic role of the overt adjective-nominal sequence produced by the child is to provide new information about the topic. Thus saying *big it* would be equivalent to the child saying *It is big it*, which is clearly redundant and could be ruled out on pragmatic grounds.

This hypothesis also runs into problems, however. It predicts that in non-ostensive contexts (where presumably there is no need for a null pronoun), children would no longer be restricted as to what follows the adjective. As found in Study 3, however, children do use adjective-nominal sequences as the subjects and objects of verbs; moreover, the same children showed no hesitation in using a bare pronoun or proper name in subject or object position, e.g., *I want that*. Thus children will frequently say things like *want big cookie* and *want that*, but will never

say anything like *want big that*. Once again, this runs contrary to the pragmatic hypothesis.

To sum up, it seems as if neither the conceptual hypothesis or the pragmatic alternative can adequately describe the child's knowledge.

2.6.3. Implications

It is easy to underestimate the child's knowledge; at first blush, the finding that children do not say *big he* is really not that exciting. To put the challenge crudely: of course children do not say *big he* -- they never hear adults say such things. You don't need to appeal to innate categories to explain this, just imitation.

This challenge needs to be modified. In its current form, it rests on the assumption that children only produce an utterance if they have heard adults produce it. This is clearly false: Even the youngest of children are productive, they have to be productive if they are to learn the language. So children hear utterances like *nice coffee* and they (somehow) infer that *nice water*, which they have never heard before, is acceptable, but that *nice Fred* is not. (In fact, the problem is more difficult than this; children sometimes do hear utterances like *bad Mike*, and must somehow screen them out or recognize them as pragmatically marked). Further, children make the exact generalization predicted by theories of adult syntax; they do not infer that all words can follow adjectives, or that just those cases used in adult speech can follow adjectives, or that just names for objects/substances can follow adjectives, and so on. Instead, they draw the generalization that only *nouns* can follow adjectives. In the absence of an alternative theory of how to explain the child's behavior, and given the fact that adult grammars employ the categories "noun" and "NP" in their principles of word order, the most parsimonious view is that children possess these abstract categories from the very start.

2.7. Could pronouns and proper names be nouns?

I have been arguing here that children's grammars allow only nouns to follow adjectives, not NPs, and that children categorize some words and phrases as nouns and others as NPs. But an alternative is that *all* lexical nominals are nouns, including pronouns and proper names. For instance, Chomsky (1965) has proposed that proper names are nouns marked with the semantic feature [+Proper]. Under this view, there would be no grammatical difference between words like *dog* and *water* and words like *Fred* and *he*; they all belong to the same grammatical category (nouns), and the reason why adjectives and determiners cannot precede words in the latter class is due to semantic factors concerning the role of modifying adjectives.

To examine this proposal, it helps to look at the exceptional cases, noted at the beginning of this chapter, where apparently adjectives *can* precede pronouns and proper names. We can exclude nicknames, like *Big John* and *Proud Mary*, since these are frozen forms, in the sense that they are not produced by combinatorial syntactic devices, but rather are coined as wholes, similar to NP idioms. A more relevant class of cases are forms like those in (9):

(9) a. I saw short Fred, not tall Fred.
b. You used to be little John, now you're big John
c. The Fodor of 1964 would have agreed to the semantic decomposition hypothesis.
d. She's a real Einstein.

These seem to be best explained as the result of lexical rules inducing category change; specifically, proper names (possibly NPs) are transformed into count nouns. In (a), *Fred* is transferred into a count noun meaning "person named 'Fred'", while (b) and (c) are cases where the count noun is quantified over "time-slices" of an individual -- the idea being that there are many Johns or Fodors, each occupying a unit of time, and the adjective or PP specifies the reference of the noun. Finally, (d) is an example of where a lexical rule transforms proper names referring to X to

count nouns meaning (roughly) "the general type of person that X is/was". In this regard, it helps if person X has some sort of culturally agreed characteristic -- Aronoff (1980) notes that this is same condition that holds when proper names are transformed to verbs, making *He Bruce Lee'd the piece of wood* easy to make sense of.

Another pattern of examples are those like *lucky you, silly me,* and *bad Mike.* These appear with marked intonation to make statements about a person: You are lucky, I am silly, Mike is bad. There are severe restrictions on which adjectives can appear in this construction (in fact, there may just be a small list, memorized by rote). Thus *fortunate you, irrational me,* and *unpleasant Mike* are all extremely odd.

I have no account of why these forms exist, but it is worth noting that the adjective serves a predicating role, not a modifying one. This relates nicely to the generalization mentioned at the beginning of the chapter: predicate adjectives appear with NPs, not nouns. The fact that strings like *bad Mike* are semantically construed as statements is then explained by viewing nominals like *Mike* as the subjects of subject-predicate relations. If so, then they should not be analyzed as NPs of the form [A - NP], but as sentences of the form [AP - NP].

To sum up, there are two relevant cases where pronouns and proper names are modified by adjectives and in each they support the analysis that these expressions are standardly NPs. For cases like *She's a real Einstein* they do appear as nouns, but they share the hall-marks of being converted into nouns by a productive rule of word-formation, akin to when proper names are used as verbs. And cases like *bad Mike* have the semantic and pragmatic properties of sentences where a predicate AP is modifying an NP, and differ radically in meaning and force from cases where adjectives modify nouns. Both considerations support the NP analysis. Aside from these arguments, there are three further considerations.

First, I have been focusing here on the referential usages of pronouns, but they are also, of course, anaphors. Thus in *Mary told the dog that he should sit, he* can be co-referential with *the dog.* Most pronouns, despite their name, are actually Pro-NPs, and they can only substitute for NPs, not nouns (compare **Mary told the*

dog that the he should sit). Under the plausible assumption that pro-forms belong to the same category as the constituents that they are co-referential with, pronouns would thus be NPs. Similarly, proper names also participate in co-reference relations and they pattern in exactly the same way as full NPs (e.g., *John cut himself/The man cut himself*).

Second, there is nothing particularly unusual about NPs that are not productively formed and which are listed in the lexicon as single items (see Di Sciullo and Williams, 1987 for discussion). There are uncontroversially other non-compositional NPs that serve as definite descriptions (*the man in the middle, the straw that broke the camel's back, the cream of the crop,* etc.). Children could categorize *Mary* as a lexical NP the same way that they could acquire these other non-compositional NPs (see Chapter 3 for discussion).

Finally, not all languages forbid modifying adjectives occurring with pronouns and proper names. Fukui (1987: 230) argues that Japanese has no NP pro-forms, and thus one can always add any semantically appropriate modifier to a pronoun in Japanese. He gives the example of how one might respond to the question, *Did you meet with Taro yesterday?*.

un, demo kinoo-no kare-wa sukosi yoosu-ga
Yes, but yesterday-GEN he-TOP somewhat state-NOM

hendat-ta
be strange-PAST

lit.: 'Yes, but yesterday's he was somewhat strange'

(This is Fukui's example (47b), with minor modifications; GEN means "genitive", TOP means "topic", NOM means "nominative".)

Fukui argues that, unlike English, Japanese has no specifiers. Since specifiers "close" phrases (e.g., in English, *the* closes off the NP *the dog*), Japanese NPs can always be further modified. Regardless of whether this explanation is correct, the

Japanese phenomena show that the unacceptability of *big it* in English is not because of universal properties of cognition or pragmatics. Instead, the English/Japansese difference is most likely due to syntactic differences between the two languages.

For all of these reasons, I will assume that pronouns and proper names are NPs, not nouns. Further evidence will be presented in the following chapter.

2.8. A learnability issue

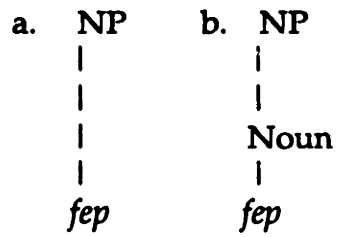
Since the 14 children I studied appear to possess rules that order nouns and NPs, then they had to somehow *learn* which words are nouns and which are NPs. For instance, even the youngest of children appear to categorize *dog* as a noun and *it* as an NP. How do they do this? This issue will be the focus of the following chapter; here I will present a consideration regarding the acquisition process that must be taken into account.

One proposal is that children start off categorizing all names -- including pronouns and proper names -- as nouns, and then use distributional information to recategorize words that are actually NPs (Maratsos, 1988). This proposal has the advantage of simplicity, as it only posits a single form-function mapping. But there is a serious problem with it.

Even 1- and 2-year-olds appear to correctly categorize the words, but the distributional cues that would help them recategorize a noun as a lexical NP just do not exist. So there are a lot of clues whether something is a noun -- if a child hears *a fep* or *many feps* or *a lot of fep*, he or she can infer that *fep* is a noun. But NPs have no special marking in the input and they often pattern the same ways as mass nouns and plural count nouns. So if you hear *I like fep*, you cannot tell if *fep* is a lexical NP like *John*, a mass noun like *water*, or a plural count noun like *children*.

There is a principled reason for this dilemma, born out of the nature of X-bar syntax. The internal and external arguments of verbs are often NPs, but never nouns. Thus if the child hears a single novel word as part of a full sentence (e.g., *I*

like fep), she knows that *fep* is an NP, but cannot distinguish between the two possibilities below, (a) where it is a lexical NP and (b) where it is a noun that is projected into an NP.



This is a classic subset problem (see Berwick, 1986). The cases where nouns without an overt determiner can appear are (from the standpoint of syntactic information) a superset of the cases where lexical NPs can appear, and thus if the child guesses incorrectly that a token like *fep* is a noun, there is no linguistic input that could inform the child that he or she is mistaken. In other words, there is no syntactic usage of *fep* that can make it clear that *fep* is a lexical NP and not a mass noun or plural count noun: NPs can only appear without determiners, and nouns can appear both without determiners and with determiners. Because of this, the learnability problem only goes one way; if children mistakenly categorize *fep* as a lexical NP and it is actually a mass noun, there is abundant information that could cue them that they are wrong (e.g., *there is a lot of fep*).

No special knowledge about adult input to children is necessary to realize this; the reader is invited, as an exercise, to attempt to use *fep* in a structural context such that it is impossible for *fep* to be anything other than a lexical NP. No such syntactic context exists. This motivates the following observation about the acquisition process:

- **Primacy of NPs:** Any word or string of words that is not clearly marked as either a noun or an NP must be categorized by the child as an NP.

This will play an important role in the acquisition theory presented in the next chapter.

Chapter 3

Nominal semantics and the acquisition problem

> Language ... must have its perfectly exclusive pigeon-holes and will tolerate no flying vagrants. Any concept that asks for expression must submit to the classificatory rules of the game ... It is almost as though at some period in the past the unconscious mind of the race had made a hasty inventory of experience, committed itself to a premature classification that allowed of no revision, and saddled the inheritors of its language with a science that they no longer quite believed in nor had the strength to overthrow. Dogma, rigidly prescribed by tradition, stiffens into formalism. Linguistic categories make up a system of surviving dogma--dogma of the unconscious. -- Edward Sapir, *Language* (1921), pp. 99-100.

3.1. Introduction

All languages have a distinct class of nominals, which share certain semantic and syntactic properties (see Schacter, 1985 for a review). Semantically, one rough description of nominals is that they "cluster around concrete concepts" (Sapir 1921: 119); these concepts include objects and substances and people. Thus in all languages, words like *table* and *water* and *she* will fall into the class of nominals; they will not be determiners or modals or adverbs (see also Greenberg, 1966; Hopper and Thompson, 1984; Macnamara, 1982). From the perspective of syntax, these words are constrained by certain universal principles of grammar; for instance, they are subject to subjacency restrictions and structural conditions on

case-marking (Chomsky, 1981). Some manifestations of this relationship between cognition and grammar are strikingly obvious; given a novel name for an object or substance, a speaker of English will categorize the name as belonging to the grammatical class "noun". The cross-linguistic evidence strongly suggests that this capacity exists for speakers of every language, and there is developmental evidence that this competence is also present in very young children. This implies that the relationship between cognition and grammar is rooted in biology, and not cultural evolution or historical accident.

The precise nature of the relation, however, is a matter of considerable debate. On one extreme, it could be that the grammatical categories "noun" and "NP" can be semantically defined; this has been argued within the framework of Montague Semantics (e.g., Partee, 1987) and Cognitive Grammar (e.g., Langacker, 1987). Alternatively, it may be that the meaning-form correlations are the result of correspondence rules between cognition and grammar, but that this relationship is not one-to-one, and thus not definitional (Jackendoff, 1983). Finally, it has been argued that the semantic properties of nominals are due (at least in part) to special mechanisms of language acquisition (Grimshaw, 1981; Pinker, 1984).

This final suggestion is known as "semantic bootstrapping", and in earlier work (Bloom, 1990a), I adopted a version of this to explain the finding that 1- and 2-year-olds categorize some words as nouns and others as NPs (see Chapter 2). It was suggested that children apply the following mapping rules:

- Names for kinds of objects and substances are nouns
- Pronouns and names for individuals are NPs

The first rule captures a universal property of language; as noted above, words like *dog* and *water* always appear as nouns. The second rule may not be universal, though it does hold for almost all pronouns and proper names in English. If Fukui (1987) is correct, then languages such as Japanese may encode pronouns and proper

names as "open" categories, and thus they are more like nouns than NPs. Moreover, some pronominals, such as the English word *one*, are not NPs. Nevertheless, the second rule will suffice, since if children initially miscategorize some nouns as NPs, there is abundant positive evidence that could lead the children to make the re-categorizations (see Chapter 2 for discussion).

Although these mapping rules may describe the inferences that children make, they do not constitute an adequate theory of the acquisition process. For one thing, they are purely stipulative. They do not explain why these mappings, and not others, occur. More generally, they do not relate to the more general relationships that hold between cognition and grammar. For instance, it is not only names for material entities that tend to be nouns, it is also names for times (*week, moment, eternity*), names for emotions (*anger, sorrow, joy*) and names for types of energy (*electricity, light, heat*). Under the bootstrapping theory above, the fact that *dog* is a noun and the fact that *week* is a noun are totally unrelated -- the first is the result of an acquisition device, the second is due to some other (unexplained) relationship between meaning and form. It would be theoretically desirable (though not necessary, of course) if the explanation for how children learn the grammatical categories that new words belong to "falls out" from a theory of meaning-form relations in general -- this is what I have dubbed the Semantic Competence Hypothesis (SCH). Outlining such a theory is the goal of this chapter.

In the next section, I will sketch out a preliminary theory, a rough attempt to explain the acquisition process in terms of semantic competence. This will turn out to be inadequate in several ways, and its problems are reviewed in Section 3.3. The rest of the chapter will spell out a more detailed attempt to extend the SCH to explain how children categorize novel words as nouns and NPs. The revisions are largely motivated by Pinker's (1989) analysis of the general relationship between language and thought, by Jackendoff's (1983, 1987, 1990a) theory of semantic features, and by Langacker's (1987, 1990) theory of the cognitive structure of nominals -- though the revised theory will differ in certain ways from all of these

views. It will not be without its flaws, but it will hopefully contain enough insights to provide the framework for further research.

3.2. A preliminary theory: Kinds and individuals

There are two reasons for starting off with an inadequate theory. First, the discussion of the problems that it runs into will make it clear what sort of phenomena a successful acquisition theory must account for. Second, the view I end up with will basically be an extension and elaboration of this theory, rather than an entirely different proposal.

3.2.1. Representation

The fundamental claim is that for both children and adults, nouns are conceptualized as words that encode natural categories, what are often called kinds, while NPs are conceptualized as words or phrases that refer to entities or individuals. Thus *dog* itself does not refer, but it can be quantified and modified to form an NP, which does refer. For instance, *a dog* picks out a particular individual that happens to be a dog, *big dogs* picks out those dogs that have the property of being big, and so on. Pronouns and proper names are lexical NPs because, unlike words such as *dog* and *water*, their role is to pick out individuals.

Note that this is a psychological theory, not a metaphysical one. Thus the hypothesis is not about the actual relationship between linguistic categories and objective reality. Rather, it concerns the relation between different levels of mental representation. In particular, I will use expressions like "refers to", "picks out", and so on (perhaps idiosyncratically) to characterize certain relations between mentally represented linguistic structure and non-linguistic mental models of the world, not between linguistic structure and the world itself. This is important to stress, because a "psychologized" theory of the sort here is immune to many of the

criticisms that might be levied against a metaphysical hypothesis.[9]

The notion that nouns themselves do not refer is a familiar one from Montague Semantics, where nouns are viewed as predicates, which must combine with determiners to establish reference (Barwise and Cooper, 1981; see Bach, 1989 for discussion). More generally, the notion is that nouns are semantically incomplete (or "unsaturated"; see Higginbotham, 1983); only NPs are semantically complete. As such, only NPs can participate in certain forms of semantic interaction, such as being subjects or predicates or having thematic-roles. Most relevantly, only NPs can refer to individuals.

3.2.2. Acquisition

From the standpoint of language acquisition, this semantic characterization of the noun/NP distinction can allow the child to draw inferences from word meanings to linguistic categories, and vice-versa. Since *Fred* picks out an individual person, children could categorize it as an NP. Since *chair* picks out a kind of object they could categorize it as a noun. Similarly, children should be capable of drawing inferences about word meaning from syntactic type; knowing that a word is an NP should tell children that it denotes an individual, while knowing that it is a noun should tell them that it denotes a kind.

There is some evidence supporting these claims. Using different methodologies, Gelman and Markman (1985) and Waxman (1990) found that 3-year-olds are sensitive to the contrast between adjectives and nouns; nouns facilitate contrasts

---

[9]Chomsky (1981: 324) makes a similar point: "If I say 'the flaw in the argument is obvious, but it escaped John's attention,' I am not committed to the absurd view that among the things in the world are flaws, one of them in the argument in question ... Suppose now that we make a rather conventional move, and assume that one step in the interpretation of LF [Logical Form] is to posit a domain D of individuals that serve as values of variables and as denotata. Among these individuals are specific flaws that can appear in arguments (cf. 'the same flaw appears in both arguments')." See Heim (1982) for a detailed theory of the type of cognitive domain in which different sorts of NPs are interpreted. It is interesting, incidently, that like Lewis (1972), Chomsky insists that the enterprise of linking up grammar with non-linguistic cognition is not "real semantics", but rather is "an extension of syntax".

between basic-level kinds (e.g., between dogs and cats), while adjectives draw children's attention towards within-kind contrasts (e.g., between big dogs and small dogs). There is also evidence suggesting that even younger children interpret new nouns as referring to taxonomic relations (i.e., to kinds) (Markman and Hutchinson, 1984; Waxman and Gelman, 1986). For instance, in one study, Markman and Hutchinson (1984) showed 2- to 3-year-old children a novel object and told them (e.g.,) *See this? It is a dax. Find another dax that is the same as this dax.* They found that the children would tend to point to another object of the same kind a significantly greater proportion of the time than when they were shown a novel object and told *See this? Find another one that is the same as this.* These results suggest that young children are sensitive to the semantic implications of nouns.[10]

In another study, Katz et al. (1974) taught young children a new word by pointing to an object and saying either *This is a wug* or *This is wug*. Thus in the first context it is a count noun and in the second it is either a mass noun or a lexical NP. When the new word was used to describe a doll, even 18-month-old children were sensitive to the grammatical difference; when they heard *This is a wug*, they took *wug* to be the name of the kind, when they heard *This is wug*, they took it to be the name for the individual (i.e., as an NP; interestingly, children apparently did not entertain the mass noun interpretation, presumably because the word referred to an object). Moreover, there was an interaction with what the nominal referred to; a word denoting a block was always treated as a kind term, regardless of the syntax. These studies suggest that even young children are sensitive to the relationship between meaning and grammar, as predicted by the theory above.

---

[10]These word/non-word experiments should be interpreted cautiously, however. Although they are often taken as showing early grammar-cognition links (e.g., Markman's (1990: 37) conclusion that "nouns help focus children's attention on categorical relations"), it is not clear that this result has anything to do with nouns *per se*, as opposed to words. These experiments – unlike the ones by Gelman and Markman (1985) and Waxman (1990) – do not determine whether children are sensitive to a noun/non-noun contrast or to a word/non-word contrast. Though it would be nice if the first conclusion were true, these experiments do not favor it.

3.3. Problems

Unfortunately, there are several serious flaws with this theory; these are summarized below.[11]

3.3.1. Other lexical and phrasal categories

One problem is that the theory above is not so much about the difference between nouns and NPs, as it is about the difference between lexical categories and phrasal categories. Thus it is unclear why *dog* refers to a kind and *runs* does not, or why *five dogs* refers to an individual and *runs around the house* does not. At best, then, the account above explains how a child, once he or she knows that a given word is a nominal, determines whether it is a noun or an NP. But how do children determine which words are nominals?

3.3.2. NPs that do not refer to individuals

Many NPs do not refer to individuals. Indeed some may not refer at all. Some troubling exceptions are as follows:

- **Expletives.** Some NPs mean nothing at all; these are known as "expletives" or "pleonastics". Thus the *it* in *it is raining* exists solely to satisfy a grammatical condition stating that all clauses must have subjects, even if these subjects have no semantic role (i.e., the "Extended Projection Principle" posited by Chomsky, 1981; Rothstein, 1983 and Fukui, 1987 discuss alternative formulations). Thus *it* is an NP but does not refer. (See Postal and Pullum (1988) for a review of semantic and syntactic criteria distinguishing referential NPs from expletive NPs.).

---

[11] I am grateful to Susan Carey, Lila Gleitman, and Jane Grimshaw for discussion on these matters. If it wasn't for them, this chapter would be much shorter.

- **Idiom chunks.** Another class of "meaningless" NPs are found within idioms, which are phrasal expressions whose meaning cannot be determined compositionally. In some cases there may be some sort of metaphorical reference for an NP component of an idiom (e.g., there is usually some entity that gets transferred when one *passes the buck* or *lets the cat out of the bag*). But there are many cases where an NP within an idiom has no conceivable interpretation; when someone *buys the farm* or *kicks the bucket*, the NPs *the farm* and *the bucket* are clearly not referential, even in a metaphorical sense.
- **NPs with the semantics of other phrases.** Grimshaw (1979) discusses NPs that have interpretations similar to concealed questions and exclamations. Thus *John asked Mary the time* appears to be synonymous with *John asked Mary what time it was* and *It's amazing the big car he bought* appears to be synonymous with *It's amazing what a big car he bought*. In these contexts, the NPs *the time* and *the big car* appear to share the semantic properties of sentences, and sentences (whatever their interpretation) surely do not refer to individuals. Similarly, Pustejovsky (1989) discusses a range of contrasts such as *Mary likes books/Mary likes reading books*, where apparently the NP *books* has a similar semantic role to the sentence (or VP) *reading books*. If NPs can freely adopt the semantic properties of other constituents, this would be a problem for any strong claims about syntax-semantics links.[12]

[12]Another case that one might bring up is so-called bare NP adverbs (Larson, 1985), such as *that way* in *John was going that way*. However, Emmonds (1987) convincingly argues that an analysis such as Bresnan and Grimshaw's (1978), where these phrases are actually PPs with null prepositions, is preferable on many grounds. One argument for this analysis that is not discussed by Emmonds is Jackendoff's (1983) observation that PPs canonically have the semantic feature PATH, and PATH seems to be exactly the semantic role that bare NP adverbs have.

- **Predicational NPs.** Perhaps the most important objection concerns predicational NPs (see Williams, 1983 for discussion). Thus the NP *a dog* in *Fred is not a dog* has a different semantic role than in *a dog walked into the room*. Under one analysis, only the second sentence actually refers to a dog; in the first sentence, the NP serves as a predicate and has much the same semantic role as the AP *happy* in *Fred is not happy*. If it turns out that NPs can adopt the same semantic role as predicates such as APs, this would seriously weaken any strong claims regarding the specific semantic nature of nominals.

3.3.3. Word learning

The theory above does not explain how children learn the meaning of words like *Fred* and *dog*; it simply assumes that they have acquired their meanings prior to exploiting the grammar-cognition links to acquire syntactic structure. The same assumption is made in the semantic bootstrapping literature, by Grimshaw (1981) and Pinker (1984). This might seem unproblematic; while it is indeed a neat question how children learn the meaning of words, a theorist concerned with grammar acquisition may reasonably view this as someone else's problem.

This assumption behind this, however, is that it is possible to learn the meaning of a word prior to learning its syntax, presumably through observation. And in a recent paper, Gleitman (1990) presents an extended argument against perceptually-based theories of meaning acquisition. Part of it goes as follows:

> The rapidity and accuracy of vocabulary acquisition are jewels in the crown of rationistically oriented developmental psycholinguists (see particularly Carey, 1978). So just as in the case of syntax, we have initial grounds for claiming that a limit on the hypothesis space must be a critical source of sameness in the learning functions. Bolstering the same view, languages seem to be as alike in their elementary vocabularies as they are in their syntactic devices (Talmy, 1975, 1985). But surprisingly enough, all the telling arguments invoked for syntax to restrict the interpretation of the input -- that is, constraints on representation -- that are to explain the sameness in form, context, and learning are thrown out the window in most theorizing about the lexicon. *There* it is usually maintained that the child considers many complex, varying, cross-cutting, subtle conjectures about the scenes and events in view so as to arrive at the right answers, comparing and contrasting probabilities across many events, properties, discourse settings and so forth ... The very richness of perception guarantees multiple possibilities at many levels of abstraction for single scenes, but the problem of word learning is to select from among these options the single interpretation that is to map onto a particular lexical item (pp. 12-13).

Gleitman's solution -- what she calls "syntactic bootstrapping" (see also Landau and Gleitman, 1985) -- is that "semantically relevant information in the syntactic structures can rescue observational learning from the sundry experiential pitfalls that threaten it". Thus, learners can infer properties of word meaning from syntactic structure. But the theory defended in this thesis shares with semantic bootstrapping the premise that learners infer syntactic structure from properties of word meaning. If one took Gleitman's critique to the logical extreme, and adopted the position that children cannot learn the meaning of a word without first knowing its syntactic category, then this is inconsistent with the Semantic Competence Hypothesis.

Gleitman's critique is directed towards the puzzle of learning aspects of verb meaning -- does it also apply to nominals? To put it another way, is it possible to learn the meaning of a word like *dog* without first encoding it as a noun? It is clear how knowing that it is a noun could at least *help* children learn its meaning.

Suppose a child sees a black dog running under a table and hears the word *dog*. Knowing that *dog* is a noun could suggest to the child that *dog* denotes a physical object, such as the dog (or the table), and preclude the interpretation that it denotes an action (running) or a spatial relation (in) or a property (big). Thus knowing the syntactic category that a new word belongs to might aid considerably in the acquisition of its meaning (see also Brown, 1957; Carey, 1982; Waxman, 1990).

Moreover, syntactic support might be of critical importance if it turns out that children cannot be assured of hearing nominals in isolation (as in the above example) but rather must acquire the meanings of words that embedded in sentences. It is commonly assumed that children hear (e.g.) the single word *dog* used several times in the context of dogs, which could permit the abstraction of relevant features common across the situations (i.e., some version of Skinnerian stimulus discrimination; see Richards and Goldfarb, 1986 for a recent example). But what if children are unfortunate enough to be stuck with parental utterances like *those are dogs, do you want a dog for your birthday?, a dog is in the room, the dog is chasing the cat* and so on? If so, then the problem of acquiring word meaning without syntactic support becomes considerably more difficult. But it must be possible if the Semantic Competence Hypothesis is correct.

It will not be necessary here to present a theory of how children acquire the meanings of words that are presented in sentences. But it is necessary to show that children can extract enough of the meaning of words to map at least some words onto at least some linguistic categories at least some of the time. This modest accomplishment will be sufficient for the child to enter the grammar. I will return to this issue in Section 3.5.3.

All of these problems motivate some major changes in the kind/individual theory posited above. In the sections below, I will present a modified theory of how children use meaning-form relations to categorize novel words as nouns and NPs. While the major assumption of the above account -- that NPs can refer and

nouns cannot -- will be retained, there is a great deal that needs to be added. In particular, it will be argued that there is a semantic feature particular to nominals, and that children can exploit the mapping from this feature to non-linguistic cognition when determining the linguistic status of novel words and phrases.

3.4. Representation

Below, I will sketch out a theory of "semantic competence". This representational account, along with certain properties of the acquisition process, should be sufficient to explain how children grammatically categorize new words and phrases as nouns and NPs.

3.4.1. Semantic features

Recent work in the acquisition and representation of verbs (e.g., Pinker, 1989) suggests that the link between non-linguistic cognition and grammar is mediated by a semantic level, as outlined below.

Cognition
|
|
Semantics
|
|
Grammar

The motivation for this semantic level, and its precise relationship with cognition, are discussed in detail in the next chapter. Perhaps the major consideration for positing such a level is as follows: It is often noted that grammars only employ a small range of all of the possible meaningful distinctions and, more to the point, children only hypothesize among this range of possibilities (Pinker, 1984, 1989; Slobin, 1985). For instance, no language would make a grammatical distinction between nouns for red things and nouns for green things and no language-learner would hypothesize such a distinction in the course of language

acquisition. Children expect (and languages employ) grammatical distinctions on the basis of shape and animacy but not on the basis of color and weight, despite the fact that they can *think* fully well about color and weight. All of this suggests that there is a semantic level intermediate between non-linguistic thought and grammar, a level which employs primitives that are a proper subset of the vocabulary of cognition and which relates to certain grammatical contrasts.

One could also explain the subset phenomenon (the fact that the meaningful notions expressible in language are a subset of all psychologically accessible notions) by positing correspondence rules that map directly from a subset of cognitive primitives to grammar (e.g., Jackendoff, 1983, 1990a). This might make a distinct semantic level unnecessary. But there are two further arguments that favor such a level. First, the similarity relations that hold among semantic properties often run contrary to those relations in cognition. Thus *dog* and *opinion* are very different cognitively, but (under certain assumptions about which features exist) identical semantically, while *garments* and *clothing* are much the same cognitively, but fall into different semantic classes and thus are categorized differently within the grammar (see Pinker, 1989, and the discussion in Chapter 4). And second, there appear to be principles that apply solely at the semantic level. For instance, the semantic property of plurality in *the dogs* results from the movement (or "percolation") of the semantic property expressed in the plural morpheme up to the NP. This movement appears best characterized at the level of semantics, not cognition or grammar.

Semantic features are used to express certain grammatical contrasts, such as between different classes of verbs, and are also expressed cross-linguistically as closed-class morphemes. Candidate semantic features within the verbal systems include PATH, DIRECTION, LOCATION, MANNER, TENSE, ASPECT, and CAUSATION (Pinker, 1989; Talmy, 1985). In the nominal system, features such as DEFINITENESS, PROXIMITY, MAGNITUDE, and NUMBER define the semantic properties of determiners. Some of these features, such as NUMBER, select for the type of accompanying noun; others, like DEFINITENESS, do not.

Semantic features must link up with non-linguistic cognition. Otherwise, they are not semantic at all, but rather arbitrary grammatical diacritics, such as gender marking in some Indo-European languages. This might seem like an obvious point, but it is quite common to fail to take it into account. Thus Gazdar et al. (1985) and Sag et al. (1985) explain the grammaticality of conjunctions such as *John is a Republican and proud of it* by proposing that (i) the conjunct *and* can link phrases that share common features and (ii) both *a Republican* and *proud of it* share the feature of PRD (predicate). This is a reasonable view, but it is left to be explained how PRD interacts with cognitive notions of predication -- which will explain why an NP like *a Republican* gets marked as PRD, but not an NP like *Fred* or an Adverbial Phrase like *very carefully*. Similar issues come up for other features they discuss, such as PAST and MANNER. If these features are not posited as linking up with non-linguistic cognition, then their names are misleading; it would be better to call them F1, F2, and F3, or GRAPE, BANANA and ORANGE. Oddly, Gazdar et al. (1985) insist as a matter of principle that their grammatical theory is unrelated to psychology, which, if true, would render their account considerably less interesting. In Chapter 5, I discuss some similar confusions in theories of the acquisition of the count-mass distinction.

Many features cross-cut different linguistic categories. For instance, INDIVIDUAL appears for both nouns and verbs, affecting the construal of both entities or events as being individuals or portions (see Bach, 1986; Jackendoff, 1990b). Similarly, though the feature NUMBER is usually associated with nouns and determiners, some languages have both singular and plural verbs (Bybee, 1985). But some semantic features may be particular to a grammatical category. I will argue below that there is a feature special to nominals -- what I will call, following Jackendoff (1983, 1987, 1990a), "THING" -- and that children can infer the presence of this feature through non-linguistic cognition. The restriction that this feature appear on nominals and the relationship between cognition and grammar are both part of semantic competence, and are not the result of an acquisition device.

Semantic feature hypotheses are held in low esteem in certain circles, and it is worth distinguishing the sort of hypothesis above from a different view, one that is probably wrong. This is the position that lexical items such as *dog* and *run* completely decompose into semantic primitives, some which are linked to perceptual or motoric mechanisms. Based on this theory, Clark (1973) proposed that the acquisition of word meaning occurs through the accretion of these primitives, and applied this theory (the "Semantic Feature Hypothesis") to the child's acquisition of certain lexical domains.

This sort of theory has been extensively criticized, most notably by Fodor (1975, 1981), who argues that that there are virtually no successful semantic decompositions of lexical items within natural language (i.e., "most lexical items are undefinable"; Fodor, 1981: 292). And there is by now a general consensus that the Semantic Feature Hypothesis does not adequately explain the process of lexical development; children's "immature" understanding of word meanings cannot be explained in terms of incomplete or incorrect collections of hypothesized semantic features (Carey, 1982; Clark, 1983). Despite the similar names, however, the Semantic Feature Hypothesis is very different from the version of the Semantic Competence Hypothesis proposed in this chapter. Unlike the semantic primitives posited by Clark and others, the "semantic features" discussed here are posited as only part of the representation of a lexical item; they only express the aspects of meanings relevant to grammar. From the standpoint of acquiring the meaning of words like *dog* and *opinion*, then, this hypothesis is fairly uninteresting -- these words are identical at the semantic level. They only differ at the cognitive level, and a theory of cognitive representation might turn out to be very different from a theory of the semantic level. As for acquisition, Clark (1983) notes that a major problem with her theory was that the semantic features hypothesized were not those that play a grammatical role in language use; they were posited more as basic conceptual units. Indeed, in her critique of the Semantic Feature Hypothesis, Carey (1984: 369) concludes that "component-by-component acquisition may hold only for

components motivated syntactically as well as semantically". Such components are precisely what the SCH is concerned with.

3.4.2. Evidence for a semantic feature special to nominals

I will argue below that nominals are associated with a semantic feature, which we can call "THING", that is linked to certain aspects of non-linguistic cognition. The claim is not that the category "nominal" can be replaced with THING, as there are some NPs that do not have this feature (in certain cases, nominals can have features such as PROPOSITION or EVENT; in other cases, they can have no feature at all, e.g., expletives). Rather, the argument is (i) all THINGS are nominals and (ii) every word or string of words that is categorized as a nominal can, in at least some contexts, have the feature THING. Thus this view does not forbid non-referential usages of (for example) the NP *it*, as in *it is raining*, so long as this NP can, in at least some circumstances, have the feature THING, as in *it is broken*.

It is the first condition that does all the work from the standpoint of acquisition theory. If we assume that for young children THING links up with the cognitive notions of material entity, we can explain the acquisition of nominals in roughly the same terms as Bloom (1990a) -- children take names for objects, substances, and people to be nominals. Thus this revised view accepts one of the major insights of the semantic bootstrapping theory developed by Grimshaw and Pinker -- that it is the uni-directional mappings from meaning to syntax that guide the acquisition process. The difference is that the existence of these contingencies will be motivated within the general framework of linguistic theory.

There are three general arguments for the position that there is a systematic relationship between the syntactic category "nominal" and the semantic feature THING.

3.4.2.1. Correlations between meaning and form

Most obviously, the semantic theory predicts a non-random relationship between linguistic category (nominal vs. non-nominal) and non-linguistic cognition. Such a relationship clearly exists; the most obvious example of this is, as discussed at the beginning of the chapter, the phenomenon noted by Greenberg (1966), Sapir (1921), and others: names for objects and substances and people tend to be nominals. While obviously "nominal" cannot be defined as "name for object/substance" (consider words like *week* and *race*), this correlation shows that there is some sort of systematic interaction between non-linguistic cognition and grammatical structure.

If the view defended here is correct, this correlation reflects two distinct relations: between cognition and semantics, and between semantics and grammar. There are, of course, alternatives: Hopper and Thompson (1984) explain the correlation in terms of shared discourse roles, not semantic properties. Other scholars, such as Jackendoff (1983) and Lakoff (1987), posit that grammatical categories relate directly to cognition, and reject the notion of an independent semantic level. Sapir (1921) suggests that the reason that nominals tend to denote things is that nominals are canonically subjects and things are the characteristic referent of subjects. Still another alternative is to explain these correlations as the by-product of biases in language acquisition (Grimshaw, 1981). But the general point here is that these correlations are out there and an adequate theory of language must explain them.

The reason that I am stressing this is that some generative grammarians take a strange stance towards the meaning-form correlations. To my knowledge, nobody has ever denied that they exist, but there are some arguments in the literature that merit comment, if only because they come up over and over again in discussions of these issues. Newmeyer (1983: 9) provides a typical example: to show the irrelevance of semantics to syntactic categories, he cites a few cases where there is no obvious reason why a word is an adjective or a verb (e.g., *likes* vs. *be fond of*) and

takes this as showing that the categorical distinction between adjective and verb is "nonsemantically based". But the form of the argument is hardly reasonable, i.e.: there is no off-the-cuff explanation for why word X is a verb and word Y is an adjective, thus the verb/adjective distinction is not semantically based. All that this argument shows is that the relationship between grammar and meaning is not *obvious*. Further, Newmeyer overlooks the fact that in languages that have both verbs and adjectives, actions like *run* are always encoded as verbs, and properties like *red* are always encoded as adjectives (Hopper and Thompson, 1984). This suggests that there is some sort of conceptual contrast between adjectives and verbs, though its precise nature is not clear.

Newmeyer's point is in the context of a discussion of the autonomy of syntax; the motivation for his argument is presumably that the autonomy claim precludes any deep relation between syntax and meaning. As argued in Chapter 1, there is no reason to believe this to be true; a more rational version of the autonomy thesis is merely that syntax has its own autonomous principles. In fact, the relation between nominals and entities such as people is simply taken as given in much of current syntactic theory, especially Government-Binding theory. To take some examples, binding theory presumes that NPs can refer and co-refer; the theory of thematic roles presumes that NPs can be "agents" and "themes"; the theory of Logical Form presumes that NPs can serve as quantifiers which have scope; predication theory presumes that NPs can serve as subjects to predicates. None of this could apply if (say) modals were used to refer to people and if NPs were used to make distinctions between modal properties such as necessity and possibility. Indeed, if it turned out that the relation between nominals and things was merely a matter of blind luck -- historical accident, not innate knowledge of language -- then one must say the same of all the other aspects of grammar that rest on it.

In fact, we can turn this argument around: The very success of a syntactic program of research that tacitly assumes that NPs refer to people and modals do not (consider binding theory, for example) is a strong argument that the

relationship between NPs and people is itself an integral part of grammar. This point is fairly banal as a matter of linguistic research; to my knowledge, for instance, no linguist has worried about working out co-reference relations for modals on the off-chance that some unknown language might use modals to refer to people. As usual, it is only in the meta-theoretical pronouncements that these issues sometimes get muddled.

In any case, the argument here about the meaning-form relations is that (i) these correlations exist and (ii) syntactic theories require that they exist. Not a bad beginning, but if syntactic theory provides some indirect support for certain grammar-meaning links, much more evidence comes from the theory of lexical semantics itself. In the following section, I will present evidence from the semantic theory of verbs that suggest a relation between NPs and the semantic category THING. This is followed by a discussion of how the semantic feature THING relates to non-linguistic cognition. If it turns out that both links exist (i.e., between grammar and semantics and between semantics and cognition), then this will suggest that the grammatical class of nominals is systematically (although indirectly) related to cognition. This in turn has implication for the question of how children map words and phrases onto linguistic categories.

3.4.2.2. Semantic theories of verbs and prepositions

The argument in this section goes as follows:

- Verbs have semantic representations associated with them. Parts of these representations include a slot which must be filled by a linguistic category bearing the semantic feature THING.
- The only syntactic category that can appear in a THING position is an NP.
- Every word or string of words that belongs to the category NP must be able to appear in slots that select for THING. (Note: This does not mean that a given NP (e.g., *it, the time*) cannot sometimes appear without the feature THING -- just that it must be capable of encoding the feature THING at least some of the time.)

Semantic theories of verbs (and other categories, which I will ignore for the purposes here) posit that they select for phrases that fulfill certain semantic roles. Consider, for example, Jackendoff's analysis of the verb *put*.

> "Put" expresses a semantic function that maps three arguments into an EVENT. The arguments, corresponding to the readings of the subject, the object, and the PP, are two THINGs and a PLACE or PATH.

This translates into the following semantic structure (modified from Jackendoff (1983: 67), see Pinker (1989: 180) for a similar analysis, also involving THING and PATH).

EVENT

"PUT": ([THING], [THING], [PLACE/PATH])

This provides an illustration of the difference between semantic representations and cognitive representations. The structure above does not by any means constitute a complete description of what people know about *put*. Plenty of non-synonymous verbs map two THINGs and a PLACE or PATH onto an EVENT

(*returned, threw, lobbed, positioned,* etc.). Thus one should not equate acquiring the semantic structure of a word with acquiring its meaning. Nevertheless, if Jackendoff's structure is the correct description of the semantic representation of *put,* then it provides all the lexical properties of the verb that are relevant to grammar, with all the rest of the cognitively relevant properties (which are irrelevant to grammar) represented at a distinct cognitive level.[13]

The representation specifies that *put* takes three arguments and not two or four (**John put the book,* **John put the book on the table to Bill*). More to the point, it specifies the type of argument, that the direct object must be a THING and not a PROPERTY or PROPOSITION. So one cannot use *put* with PROPERTY direct objects, as in **Santa put happy into his children's hearts* (compare *Santa put happiness into his children's hearts*), or with PROPOSITIONS as direct objects, as in **John put Frege is wrong into his book* (compare *John put the argument that Frege is wrong into his book*). Importantly, it also turns out that the very same slots that select for THINGS also select for NPs. This holds when a verb is used to encode some sort of physical event, as in *John put the book on the table,* but also when it is used to encode something other than transfer of a material entity, as in *John put ideas in Mary's head* or *John put his financial future in great peril.* The phrases *John, ideas,* and *his financial future* are all NPs.

Is it true that all and only NPs can have the feature THING? Immediately after presenting his analysis of *put,* Jackendoff suggests that the relation between semantic features and syntactic categories is not one-to-one; not all nominals are THINGs. He gives *destruction* as the example of a noun that maps onto the feature EVENT and *mile* as mapping onto AMOUNT. Jackendoff (1990a) gives the further examples of *the war,* which he categorizes as an EVENT, and *redness,* which he categorizes as a PROPERTY. If Jackendoff is correct, this would not seriously affect

---

[13]Jackendoff takes his representation to be a proposal about conceptual structure, as he does not adopt a distinction between semantic and cognitive levels. The difference between Jackendoff's view and the SCH is more clear in his analysis of the verb *drink* (e.g., in Jackendoff, 1987: 386) which has the semantic primitive MOUTH. Under an analysis which distinguishes between a semantic representation and a cognitive representation, this aspect of the meaning of *drink* would be encoded at the cognitive level.

the learnability solution presented here, since the child could still safely infer that everything with the cognitive correlates of the semantic feature THING is a nominal. The more damaging examples would be cases where constituents which are not nominals can be THINGS, since then children might sometimes categorize words with the feature THING as nominals when they actually belong to some other linguistic category. To my knowledge, there are no counter-examples of this sort. Nevertheless Jackendoff's counter-examples are worth consideration, since I wish to make the stronger claim that not only are all phrases with the semantic feature THING NPs, but that all NPs can have the feature THING, at least some of the time.[14]

It is true that it is awkward or ungrammatical to use some of Jackendoff's counter-examples as the direct object of *put* (e.g., *?put destruction ...* or *?put a mile ...*). It could be that Jackendoff is correct and these are cases of semantic ill-formedness; the direct object of *put* must be a THING and *destruction* and *mile* are not THINGs. Alternatively, though, the problem may be because such words are cognitively nonsensical in these contexts. That is, they are unacceptable at the cognitive level, not at the semantic level, because it is hard to make sense of the conceptual notion of "putting" with regard to entities like miles and destruction. This is supported by the fact that all of Jackendoff's counter-examples can appear in other contexts which clearly select for THING, as follows:

The alcohol put redness into her eyes. / The alcohol put a stain in the rug.
John gave the war his full attention. / John gave Mary his full attention.
John hated destruction. / John hated snakes.
A mile is longer than a foot. / The stick is longer than the pencil.

---

[14]It is worth pointing out why I am arguing for this position since, strictly speaking, all that is necessary to solve the acquisition problem is the weaker claim that all THINGS are NPs. Nevertheless, the notion that all NPs must be *capable* of having the feature THING is related to the acquisition theory discussed in Section 3.5. It is argued there that the route to learning that a word or string of words is a NP is by inferring (through observation or linguistic structure) that the word or string of words has the feature THING. If this was the only way that children could come to use a given NP, then all NPs should be THINGS. But language learners, once they have acquired a given nominal, can then use productive properties of language to extend it to fulfill other semantic and grammatical roles, such as expressing PROPOSITIONS or appearing as expletives. This leads to the weaker conclusion that all NPs should be able to have the feature THING at least some of the time, because otherwise they could not have been acquired.

Thus *redness* can appear in the same context as *a stain, the war* as *Mary, snakes* as *destruction,* and *a mile* as *the stick,* and these replacements are all THINGS even in the more intuitive sense that Jackendoff is relying on. The toughest case is *a mile;* while it is clearly an NP, and can appear in the same contexts as other NPs, it is not easy to find contexts where it can be inter-changed with phrases describing material entities. It is worth considering why this is the case. One might argue that it is because *a mile* does not really have the semantic role of THING, but rather (something like) AMOUNT, which can be either an amount of space (*a mile*) or of time (*an hour*) (AMOUNT, like THING, might also be restricted to NPs.) But there is another explanation, which is that both *an hour* and *a mile* are THINGs, just as much as *a dog* and *John,* but because of their cognitive representation -- they pick out bounded regions of space and time -- they often cannot felicitously appear in the same contexts as names of physical objects. To put it another way, it is difficult to find a cognitive notion that can sensibly be applied to both a dog and a mile. You can change the location of a dog, for instance (e.g., *John put the dog in the house*), but cannot change the location of a mile, as in ?*John put the mile in the house.*

The deciding factor here is that we can find some contexts where *a mile* can appear with predicates that apply to THINGs. These can occur in the context of measurement (e.g., *A mile is longer than a foot; the stick is longer than the pencil*), and in contexts where the THING is the object of a certain type of action, (e.g., *If you divide a mile in half, you get half a mile; If you divide an apple in half, you get half an apple.*) The second pair of sentences is important because here *a mile* cannot be replaced by a PP with a similar meaning, (e.g., **If you divide from Boston to Montreal in half, you get half from Boston to Montreal*). Thus these examples show that (i) *a mile* can appear in positions that select for THINGS, and (ii) *a mile* can appear in positions that only take NPs. These considerations suggest that *a mile* is both an NP and has the semantic feature THING.

The argument so far is that a decompositional semantic treatment of verbs

involves a semantic feature THING and that this feature always corresponds to a NP. If correct, then this is half of the necessary argument -- there exists the relevant mapping from grammar to semantics. In the next section, I discuss the mapping from the semantic feature to non-linguistic cognition.

Of course, if Jackendoff and others are wrong and verbs are semantic monads (as argued by Fodor, 1981, for instance), then the argument falls flat. But there is considerable reason to believe that verbs do have the sort of semantic structure assumed here. This gains support not only from detailed linguistic analyses (e.g., Jackendoff, 1990a), but from a set of experimental studies of language acquisition (see Pinker, 1989 for a review).

Note also that the relationship between semantic features and grammatical categories is non-trivial: it is not as if Jackendoff's semantic features are merely re-descriptions of syntactic subcategorization frames. For one thing, some (probably most) semantic features do not always map onto a single grammatical category. As Grimshaw (1979) points out, the semantic role of PROPOSITION can be filled by either a sentence (e.g., *John asked me what time it is*) or an NP (e.g., *John asked me the time*). Similarly the semantic role of PATH can be filled by either a PP (e.g., *John walked from Boston to New York*) or an NP (e.g., *John walked the entire distance from Boston to New York*).

In fact, for better or worse, the claims about the relation between THINGS and NPs are clearly falsifiable. For instance, if it turned out that Jackendoff was correct and *a mile* could not appear in any contexts that selected for THINGS, then it would be false that all NPs could be THINGS and the claims above would have to be softened considerably. The relationship between a semantic feature and a grammatical category is an empirical hypothesis, and it might hold in some instances (THINGS and NPs) but not in others (PATHS and PPs).

The argument so far has gone as follows: Verbs and other categories select for phrases that fulfill certain semantic roles. For instance, the direct object of *believe* is a PROPOSITION, the indirect object of *throw* is a PATH. While it is very unclear how

these notions are related to our cognitive conceptions of propositions and paths, there is no doubt that some relationship exists. It is our understanding of *birds fly* that leads us to view it as a PROPOSITION and our understanding of *towards the wall* that leads us to view it as a PATH. Further, these selectional properties cannot be reduced to selection for syntactic categories; *the story that John told* and *the house that John liked* are both NPs, but only the former can appear as the direct object of *believe*, presumably because only the former has the semantic role of PROPOSITION. The argument in this section has concerned another semantic role: THING. Just as with PROPOSITION, GOAL, PATH, MANNER, and so on, the feature THING was hypothesized to capture certain phenomena within natural language, e.g., that the direct object of *put* can be a phrase like *a book*, but not a phrase like *on the floor*. I have argued above that *as a matter of empirical fact*, this feature is strongly related to the syntactic category NP, i.e., all THINGS are NPs and all NPs can appear with the feature THING. If this is correct, then this relationship may explain the child's capacity to syntactically categorize new words and phrases.

Two red herrings that should be put aside: First, none of this precludes the possibility that the relation from semantic features to syntactic structure is a many-step process. I have been pretending here that Jackendoff's semantic structure for *put* gets directly instantiated into strings like *John put the book on the table*. But nothing rests on this; the semantic structure of a verb and its arguments might first get transformed into some other representation before it finds its way into a sentence -- all the arguments still apply. And second, the theory above does not imply that the selectional properties of a verb can be entirely determined by the verb's semantic representation (which is claimed by Chomsky (1986), who cites Pesetsky (1983) as supporting it). This may be the case, but it could also be that there are strictly syntactic requirements at work, as argued by Grimshaw (1979). Thus, there may be two requirements that a constituent must satisfy in order for it to be selected by a verb: syntactic and semantic. To take one of Grimshaw's examples, both *ask* and *wonder* have the semantic property of selecting for indirect

questions, but syntactically only *ask* selects for an NP. Thus we get the contrast between *John asked Mary the time* and **John wondered the time.* None of the arguments here would be affected if Grimshaw's analysis is correct, which is a further illustration of how a syntactic category can relate to semantics without violating the autonomy of syntax hypothesis.

3.4.2.3. Extension of nominal semantics to non-physical "spaces"

As noted above, the physical/spatial relationships encoded in the semantic structures of lexical items can extend outside the physical domain. Thus spatial prepositions can be extended to the temporal domain and to other more abstract domains (*from Boston to NY, from Monday to Tuesday, from rich to poor, from fresh to rotten,* etc.). Verbs of physical transfer and physical contact can be similarly extended (*John gave Mary a book, John gave Mary a headache, John gave metaphysics all his attention; John made Fred sit, John made his son go to college, John made punk music what it is today*). This notion that languages productively and systematically extend physical notions across more abstract domains is known as the "Thematic Relations Hypothesis" (first developed by Gruber, 1965) and it plays an important role in the semantic theory that I assume here.

I am inclined to agree with Pinker (1989) that this theory is "rather profound". Among its other explanatory benefits, it can give us some insight as to how THING can extend to non-material domains. The idea is this: when the semantic structure of the verb gets extended (e.g., from transfer of a material entity through space to transfer of a non-material entity from one person to another), its arguments, most relevantly those that are THINGS, undergo similar extension. Just as the verb need not refer to perceptually salient manipulations of the material world, its arguments need not refer to material entities. For instance, as the verb *gave* is extended outside of the material, it can select for THINGS such as events (*The movie gave John a nightmare*), units of time (*The boss gave me a day to finish the big project*), mental states (*He gave me a headache*), and abstract properties (*He gave professional boxing a bad reputation*).

We are now in a position to take a stab at a critical question: What is the cognitive correlate of THING? What is the nature of this particular link between semantics and cognition? The necessary claim for the acquisition argument to fly is that it be initially related to representations of material entities, and that it be extendible to other mental spaces, such as time and possession. A detailed theory of this is presented by Langacker (1987, 1990) and the brief discussion that follows is based largely on his work, although it differs in some regards.

Here's the wrong way to think about how grammar relates to cognition, one that is very tempting: Entities a,b,c in the world are named with nominals, entities x,y,z are not, and the task of an adequate theory is to explain what properties a,b,c have that x,y,z do not. Once you start on this road of analysis you will find at best probabilistic generalizations, such as "nominals tend to be names of objects" or "verbs tend to be names of actions". These relations are correct and non-trivial, but they do not lead to very much insight. Moreover, there are frequent counter-examples, both within English and other languages. One immediate problem that comes up is that the same entity (or slice of space/time) can be described both with nominals and with other categories. An event can be described as *a race* or *racing*; a patch of color can be described as *redness* or *red*, and so on. These are the sorts of phenomena that led Bloomfield (1933: 266) to the cheerful assessment that "to accept definitions of meanings ... is to abandon scientific discourse".

What the *race/racing* cases actually suggest, however, is not that linguistic categories are semantically arbitrary but just that their roots are to be found in cognition, not ontology. This is entirely unsurprising from the perspective adopted here, as we are studying the links between grammar and thought, not grammar and the world. We are likely to find, then, that the differences between those words that do and do not get assigned the feature THING have more to do with the way that the referents of these words are *construed* than with the nature of the actual real-world referents themselves.

In particular, one might suggest that the difference between *a race* and *racing* has to do with cognitive structure. The first phrase encodes a bounded unit of time, a region of what we can call "event-space". Thus it is a THING. In contrast, the verb *running* is relational; it focuses on a process involving THINGS, and is thus an EVENT. In general, THINGS are static regions of some mental space. These include regions of the mental space corresponding to the material world (*a dog*), property-space (*redness*), event-space (*a race*), interval-space (*a week*), emotive-space (*rage*) and so on. One way to view this proposal is with the analogy of a flashlight; the notion of THING corresponds to the light of the flashlight shining on one section of a mental space. In the most literal case, it is the spotlight of attentional focus on the output of the visual system, but it can also focus on regions of more abstract spaces, such as those expressing properties, events, intervals, and so on.

The acquisition theory presented in this chapter does not rest on this type of account being correct, but only that there be *some* systematic semantic-cognitive relation that holds between the feature THING and non-linguistic thought. And there is a fairly knock-down argument that such a mapping exists, not only for material names like *dog*, but also for words like *week*, *idea*, and *thirst*.

The argument is this: Determiners with certain semantic properties interact with nouns to form NPs, and the cognitive construal of the NP reflects the interaction between determiner semantics and noun semantics. The interesting fact is that the semantic force of a given determiner is consistent across all nouns, no matter how abstract their meanings. For instance, the contrast between *many* and *few* is the same whether we are talking about the *many dogs* vs. *few dogs* or *many problems* vs. *few problems*; the cognitive import of the determiner does not change as we move from (to keep with the terminology introduced above) one mental space to another. The simplest view is to say that the effect of MAGNITUDE can apply to all nouns and has a consistent effect on NP semantics. Similarly for DEFINITENESS (*a race* vs. *the race*); PROXIMITY (*those ideas* vs. *these ideas*), EXISTENCE (*no trouble* vs. *trouble*), and so on.

The fact that the semantic notions encoded in these determiners can readily interact with all nouns to produce clear and understandable construals is evidence that there are some aspects of meanings shared across all nouns. More specifically, just as verbs extend notions of causation and contact from the physical domain to more abstract domains, determiners interact with nouns to extend spatial notions such as magnitude and proximity from the material domain to more abstract domains. Just as *more dirt* has greater magnitude (i.e., takes up more actual space) than *less dirt*, there is a similar interval-space effect (*more time* vs. *less time*), event-space effect (*more jogging* vs. *less jogging*), property-space effect (*more pain* vs. *less pain*), and so on. Just as one would use *this dog* to talk about a dog that is close and *that dog* to talk about one that is far, there is similar contrast when one talks about *this idea* vs. *that idea*, *this war* vs. *that war*, and so on. The simple fact of these extensions suggests a consistent relation between nominal semantics and cognitive structure, one based to some extent (no pun intended) on the notions of space, though the details are, of course, almost a total mystery.

3.4.3. NPs that are not THINGS

I have been careful to repeatedly stress that a given NP need not always have the semantic feature THING, but under certain circumstances can encode other semantic features. Here's why: Nominals can sometimes fulfill other semantic roles, such as when *the time* can mean *what the time is* (e.g., *I asked Mary the time*) or when *the book* can mean *reading/writing the book* (e.g., *Mary finished the book*) (see Grimshaw, 1979, and Pustejovsky, 1989, for discussion). Obviously, these NPs can also have the semantic role of THING -- *the time is 8:00*, *the book is heavy* -- but in cases like *I asked Mary the time* both the phrases' interpretations and the linguistic contexts that they appear in suggest that they can also adopt semantic features such as PROPOSITION or EVENT.

I have no theory for why some lexical items can encode semantic features other than THING and others cannot, or why some semantic roles appear to be more

accessible than others. The explanation may lie in what Pustejovsky (1989) calls "qualia structure", properties of conceptual structure that allow for the extension of a primary word meaning into other domains. So part of the qualia structure of *book* may be the characteristic function of a book -- to read it or write it, and this might allow the EVENT interpretation in *John began the book.* A similar phenomena may account for how nouns can shift from "concrete" to "abstract" construals, as in the shift from *This book weighs half a pound* to *This book is deeply confused* (both readings correspond to the semantic feature of THING, however). The study of what kinds of extensions are permitted may provide considerable insight into conceptual representation and its development; it extends Keil's (1979) program of describing ontological structure in a surprising direction.

But in any case these extensions do not detract from the solution of the learning problem. Children can still infer that THINGS are nominals, because they are -- it is just that, in some contexts, nominals are not THINGS.

3.4.4. On so-called predicate nominals

Consider the contrast between *A dog is in the room* and *Fred is a dog.* However one chooses to describe the contrast, there is a difference; in the first sentence it refers to a single dog, while in the second it says something about Fred. The first sentence entails that there exists a specific dog; the second does not.

One common view is that NPs can sometimes appear as predicates, with their status licensed by syntactic structure (see Williams, 1980, 1983) and determiner choice (see Partee, 1987). The shift from referential to predicative status is nicely explained within the framework of Montague Semantics, where the change from a referential reading (*e*) to a predicate reading (*e, t*) is the natural result of a type-shifting rule. Grammatical support for the view that some NPs are predicates comes from the fact that these NPs appear in the same structural position as other phrases, such as APs, that are more obviously predicates (e.g., *Fred is a dog/happy*). Further, they can co-join with such categories (e.g., *Fred is a Republican and proud of*

it; Sag et al., 1985), also suggesting that they share semantic properties with these predicates. Dowty (1989), Partee (1987), and Partee and Rooth (1933) provide further evidence for this analysis.

I will not be adopting this view here, however. Type shifting can elegantly describe a range of linguistic phenomena (such as conjunction and the selectional properties of verbs) and some version of it may capture interesting generalizations in language development and representation. Nevertheless, it does not mesh well with the semantic framework proposed above, and in certain regards is patently unrealistic as a psychological theory. In particular, the notion of common nouns as predicates is quite problematic from a grammatical point of view (e.g., Gupta, 1980) and there are serious difficulties in extending a model-theoretic analysis as part of a psychological theory (e.g., Lakoff, 1987).

In any case, there is a plausible alternative proposed by Jackendoff (1983). This is the view that NPs like *a dog* in *Fred is a dog* are not predicates, but rather are categorized as TYPES. Jackendoff argues that NPs can be either TYPES and TOKENS, both which are sub-classes of THING, with their particular interpretation a result of syntax, determiner choice, and cognition. So *Fred is a dog* should not be viewed as equivalent to saying that Fred has a certain property, that of doggishness, but rather that Fred belongs to the type "dog". We construe TYPES as referring to classes of entities, perhaps with some sort of "essence", while PROPERTIES, like *happy*, are construed differently, as picking out properties or attributes of TYPES and TOKENS. There are four main advantages to this approach.

First, the Montague Semantics approach predicts that *John is Republican* and *John is a Republican* should have identical interpretations, as both are cases where something is predicated of John. But there is an intuition that these sentences differ in meaning -- the first sentence attributes a property to John while the second sentence makes the stronger claim that John is the member of a certain group. There is some empirical support for this contrast; in an unpublished study, Smith and Markman compared adult's intuitions about NP-AP sentences vs. NP-NP

sentences. They contrasted *John is intellectual* with *John is an intellectual, John is male* with *John is a male,* and so on. Smith and Markman found that subjects believed that the NP-NP statements were "stronger", conveyed more information, and "seemed to refer to a more central, permanent feature of the person" (Markman, 1989: 123). This suggests that some distinction between NPs and APs should be preserved at the semantic level.[15]

Second, some utterances of the form NP-*is*-NP are clearly not predicative, such as *Samuel Clemens is Mark Twain* (see Williams, 1983 for discussion). This can be viewed in terms of two types of *is*, one predicative and one equative, but one might think that a generalization is being lost, especially since other verbs (e.g., *resembles*) also allow for both readings (Jackendoff, 1983). It would seem that the most parsimonious view would capture the two usages of *is* under a single semantic characterization, something like Jackendoff's (1983) posited relation "IS EXEMPLIFIED BY" (see Dik, 1987 for a similar proposal).

Third, the TOKEN/TYPE contrast can explain phenomena distinct from predication, such as generic reference. Consider the contrast between *dogs ran into the room* and *dogs bark*; the first says something about particular dogs (TOKENS), the second is generic, and says something about the category "dogs" (the TYPE). As Jackendoff (1983) argues, this is accounted for naturally in terms of the TOKEN/TYPE contrast.

Finally, there are contexts that only select for TOKENS (e.g., ... *walked in* or *John bought ...*; see Jackendoff, 1983: 89). There are also contexts that select for TYPES and TOKENS without selecting for PROPERTIES (e.g., *John is similar to ..., John resembles ...*). This suggests that one needs separate classes of TYPES and TOKENS (each distinct from PROPERTIES) in any case, so it is not superfluous to posit this sort of division within NPs.

---

[15]Palmer (1990: 232) provides another example: "... the noun *cripple* is today avoided and the adjective *disabled* is used instead. The purpose of this change is, I think, to avoid categorizing disabled people as a "kind" and so a class apart, but to concentrate on their "property", which implies the need for help and respect."

But we have not yet explained why it is that NPs and APs can sometimes pattern together (e.g., *John is a Republican and proud of it*). Perhaps we can retain some of the insight of the predicate approach by assuming that there is some semantic feature that encompasses both THINGS and PROPERTIES and that verbs such as *is* and *becomes* select for it. We could dub this feature PREDICATE or STATE, and would presumably expect to find some cognitive correlate relating to it. But in any case, the arguments above suggest that the existence of "predicative NPs" can be explained without positing that these NPs do not have the semantic feature THING.

Note that both TOKEN and TYPE are sub-classes of the semantic category THING. These features only show up in NPs, however, not nouns. This leads to the following conjecture: nouns are simply marked with the feature THING, while NPs must be marked as either TOKENS (as in **Several dogs** *are on my lawn*) or TYPES (as in **dogs** *bark* or *Fred is* **a dog**). This is an extension of the claim made in Section 3.2 that NPs, not nouns, can refer -- and since reference must be either to a TOKEN or a TYPE, it follows that NPs that have the feature THING must be either THING-TOKEN or THING-TYPE.

3.4.5. Summary

Where does this leave us? While none of the problems could be fairly said to have been solved, the outlines of solutions have been presented. Based on evidence from noun and verb semantics, one could make the following conclusions:

- THING is a semantic feature selected for by some lexical items.
- All THINGS are NPs, and all words or strings of words that are NPs can, at least some of the time, have the feature THING.
- The feature THING corresponds to the cognitive notion of a material entity, but can also extend to non-perceptual mental spaces.
- NPs can also encode other features, such as PROPOSITIONS and EVENTS.
- THINGS can be either TYPES or TOKENS; this corresponds to the contrasts of referential vs. predicational and generic vs. non-generic.

We are still left with other puzzles, such as expletives and idiom chunks. These are best discussed in the context of the acquisition theory.

3.5. Acquisition

I will start by first considering older children, those who have already categorized some words as nouns and verbs, and then move to the more interesting case of children who have not yet acquired any word-category mappings.

3.5.1. What can children learn about one word from the semantics of another?

3.5.1.1. Learning about syntax

In Section 3.4.2.2 it was argued that certain predicates select for THINGS. If children map THINGS onto NPs then it follows that they should be able to categorize unfamiliar new words as NPs without observation at all. Suppose a child hears *I ate a cookie* and has never heard the phrase *a cookie* before. If the child knows the semantic structure of *ate*, then she knows that *a cookie* must be a THING, and if she has the relevant semantics-syntax mapping, she knows that it must be an NP. More generally, the child can learn the syntax of one word by exploiting (i) the semantic selectional properties of another word along with (ii) semantics-syntax links.

This is sometimes described as categorizing words through syntactic cues (e.g., Gordon, 1985), but if the view above is correct then this is the wrong way to look at it. It is not the syntactic properties of *ate* that license the inference; it is the semantic properties.

A similar procedure could aid in the categorization of new verbs. Suppose children know that *John* and *Mary* are words which have the semantic feature of THING and then hear them used with the novel word *kissed*, e.g., *John kissed Mary*. Assuming that children expect complete sentences to express propositions, the utterance might be semantically encoded as follows:

PROPOSITION

([THING] ??? [THING])
John kissed Mary

One likely candidate for a universal of language is some constraint to the effect that all NPs with semantic content must get their semantic role from some predicate. This constraint shows up in all linguistic theories in one form or another (e.g., the theta-criterion; Chomsky, 1981), though its precise formulation is a matter of some debate (see Jackendoff, 1987). If children possess this constraint, then they can categorize *kissed* as a predicate (since there must be something licensing these NPs and *kissed* is the only candidate) and give it the following representation:

EVENT

"KISSED": ([THING], [THING])

Assuming that there is some mapping from EVENTS to verbs, this could explain how children who only know NPs can use semantic structure to categorize novel words as verbs.

3.5.1.2. Learning about meaning

Returning to our *I ate a cookie* example, the child can use properties of the verb meaning not only to infer the syntactic category of *a cookie*, but also certain aspects of its meaning. From the standpoint of word learning, there are actually two different sorts of inferences that can go on here.

One is based on the semantic selectional properties of the verb; this tells the child that *a cookie* is a THING. It also tells the child something about the meaning of *a cookie*; it could refer to an object, for instance, as there is some link between the semantic feature THING and the cognitive notion of a material entity. Thus the direction of this inference is from the semantic structure of *give* to the semantic feature THING and from the semantic feature THING to its cognitive correlate. Here the sort of cognitive knowledge gained is rather rough-grained; since there are basically two semantic classes of nouns in English, count and mass, there really isn't that much one can learn from attending to the semantics -- the same semantic features that *dog* has are also shared by *week, headache,* and *divorce.* In the verbal system, however, there may be a large number of possible semantic configurations, each with its own cognitive correlate, and thus this sort of semantic inference might play a considerable role in the acquisition of word meaning. This is one of the core claims of the "syntactic bootstrapping" theory presented by Landau and Gleitman (1985) and Gleitman (1990), though they put it in a rather different framework (see below).

A second very different sort of inference is at the cognitive level. Consider what an adult might infer about the new phrase *a zoop* upon hearing the sentence *John ate a zoop.* From the context, you can infer that *zoop* is likely to be a physical entity, tasty, smaller than a house, non-toxic, with mass, etc. These sorts of inferences have nothing to do with the semantic feature THING; they are inferences based on the cognitive representation of *ate.* The vast majority of NP meanings are probably learned in this way, by hearing them used in context and making cognitive inferences about their meanings. Whatever the merits of perceptual learning as a

theory of the acquisition of words describing objects and substances, it is not going to explain the meaning of nominals that refer to unobservable entities, like *Russia, math, idea,* and just about every nominal other than *dog* and *water* and a few thousand others.

Thus there are two ways to learn about word meanings through language – inferences from the semantic structures of other words and inferences from the cognitive structure of other words. The notion of acquiring aspects of word meaning solely through linguistic cues (e.g., learning something about *zoop* from hearing *John ate a zoop*) is sometimes called "syntactic bootstrapping" but the term is somewhat misleading, at least if the picture above is correct. There is no reason to believe that children attend to the syntax at all when inferring word meaning (or that they attend to prosody or phonology, for that matter). The theory above predicts that semantic structure and conceptual structure are relevant, and phonological structure and syntactic structure are not. Of course, syntactic contrasts do affect inferences about word meaning (as found by Naigles, 1990, for instance); the suggestion here, however, is that they do not do so directly, but only because they reflect semantic contrasts.

One could speculate on the role of "semantic inference" on the development of cognitive domains. It is likely that children start off with several mental spaces; at minimum, they have one corresponding to the material world -- this allows for the initial mapping between THING and material entities. Nevertheless, some spaces must be acquired and it might be that children can use information from the semantic structure of verbs to structure these new spaces. So consider an older child who hears *Protons have mass.* If she represents the appropriate semantic structure of *have,* then she should infer that *protons* and *mass* are THINGS, even if she knows nothing else about them. So children know, if we follow Langacker's (1990) analysis, that *protons* and *mass* are profiled regions in some mental space. This is knowledge at the cognitive level, not the semantic level. Knowing that *protons* and *mass* are NPs, then, might lead to the establishment of some cognitive structure, and

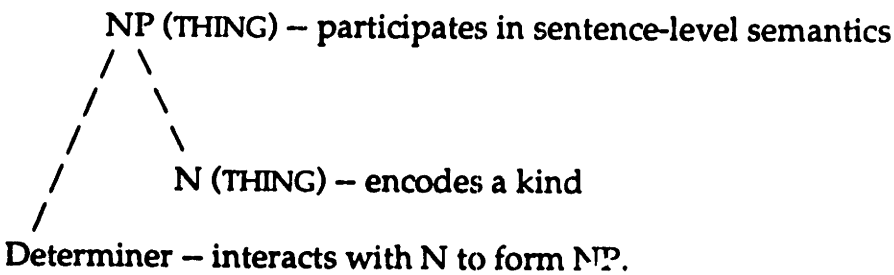
in some non-trivial sense, language can affect the way we think. I will pursue this issue in some detail in the following chapters, with regard to the count-mass distinction.

3.5.2. Categorizing new words through cognition-semantics mappings

It is tempting to say that children syntactically categorize nominals through the semantic structure of verbs in the manner discussed above and since this thesis isn't about how children learn verbs, stop right here. Unfortunately, children learn nouns before verbs (Gentner, 1982) and theories of verb acquisition (e.g., Pinker, 1989) tend to presuppose that children already have access to the category NP. Moreover, we have not yet explained the puzzle that we started off the chapter with -- how do children learn that some words are nouns and others are NPs?

3.5.2.1. The structure

Following the discussion above, I assume that children encode the following semantic structure of NPs. We can view this as a template onto which children map novel phrases. All of these constituents can have semantic features in addition to THING, such as INDIVIDUALITY, NUMBER, DEFINITENESS, and so on. The determiner always co-occurs with the noun, as it maps its semantic features onto the NP. In some English NPs, however, the determiner is phonologically empty, i.e., in some contexts with mass nouns and plural count nouns.



3.5.2.2. The learnability condition

The basic premise is that children start off parsing the sentences that they hear into phrases, such as NP and VP, and use principles of X-bar syntax and semantics to infer which words are nouns and verbs. The logic of this is simple; nouns and verbs are not the primary components of meaningful utterances; they are not subjects or predicates; nor can they have thematic roles. It is *phrases* that are initially semantically accessible to children. So, contrary to Pinker (1984), for instance, the claim is not that children use meaning-form relations to categorize new words as nouns and then use the principles of X-bar syntax to make inference about phrases. Rather, children use meaning-form relations to categorize novel words and strings of words as NPs, and they infer that some of them are lexical NPs and others reduce to determiners and nouns. This leads to the question of how they know which phrases are lexical NPs (like *Fred* and *she*) and which expand to a noun preceded by a null determiner (like *dogs* and *water*).

Here we return to the arguments presented at the end at the last chapter, where it was noted that if children miscategorize a noun as an NP (suppose they hear *water is good* and think that *water* is a lexical NP), then they can quickly change their categorization through positive evidence; hearing *some water* is enough to tell the child that he or she was mistaken. But if a child mistakenly categorized a lexical NP as a noun, he or she would be in a bind, as there would be no syntactic evidence that could show the child that the word was actually a lexical NP, since there are no contexts where only NPs can appear and nouns (preceded by a non-overt determiner) cannot. And such instances might come up in cases such as the acquisition of pronouns and proper names in languages such as Japanese. All of this motivated the "Primacy of NPs" hypothesis, which is basically that if the child is in doubt as to whether a word or phrase is a noun or NP, she should categorize it as an NP.

One additional consideration, however, is that there are semantic cues that could provide relevant information about re-categorization. So the principle should be described as follows:

- **Primacy of NPs**: Categorize all words and word combinations that have the semantic feature THING as NPs. Only decompose them into [determiner - N] sequences if there is further semantic evidence for such a decomposition. Otherwise, store them in the lexicon as NPs.

The existence of such a principle gains evidence from a surprising source. Idioms, by definition, are not fully compositional. Consider what happens according to the "primacy of NPs" principle when the child encounters the phrase *the big cheese* used to refer to somebody. Since the child cannot semantically decompose the phrase (since the referent need not be big and is usually not cheese), it is stored as a lexical NP, i.e., $[\text{the-big-cheese}]_{NP}$.

This makes the prediction that idioms will only be syntactically decomposed to the extent that they are semantically transparent – a prediction that has some evidence supporting it. Cutler (1982) reviews evidence on how frozen different idioms are (i.e., how accessible they are to syntactic movement, such as passivization). She noted a strong trend that was summed up as: "the colder the older". In other words, older idioms are more resistant to syntactic transformations that newer ones. One explanation is that older idioms may be harder to make sense of (especially to children) and thus may be acquired as syntactically unanalyzed elements, and hence would be "colder". In other words, children only do a syntactic decomposition of a phrase if they can do a semantic decomposition; older idioms are less semantically transparent to children than newer ones, ergo: older idioms should be more "frozen" than newer ones. Supporting this, there is experimental evidence showing that degree of semantic "analyzability" is strongly correlated with how syntactically decomposable a given idiom is (Gibbs and Nayak, 1989). These findings suggest that syntactic decomposition is in some sense parasitic on semantic decomposition.

This can also explain the existence of NP chunks within idioms, such as *the bucket* in *kick the bucket*. The problem with these from the standpoint of the semantic

theory proposed in Section 3.2 is that they are counter-examples to the view that all NPs refer. This view is not adopted here; it is possible under the revised theory for an NP not to have a semantic feature (see the discussion of expletives below). But in any case it might turn out that cases like *the bucket* are not really counter-examples, because they are not analyzed as NPs. The studies cited above suggest that *kick the bucket* may be acquired as a lexical VP (*kick-the-bucket*$_{VP}$) – or, if partial decomposition can occur, as *kick*$_{V}$-*the-bucket*$_{VP}$, since the meaning of *kick* in this context may be transparent. Thus *the bucket* in *kick the bucket* might not be an NP at all, any more than *cat* is a count noun in *catatonic*.[16]

3.5.2.3. Learning situations

How do these constraints apply in word learning? For all the examples below I assume that the child already knows that a word or string of words corresponds to a THING, either because of the semantics of some other part of speech or because the child is aware that these words are used to denote a material entity and the semantics-cognition links lead them to construe names for material entities as having the feature THING. Further, I assume that children know that some words and strings of words refer (and are thus "referentially complete"), while others are names of kinds (e.g., *dog*) and can only refer when they co-occur with a determiner. For the sake of exposition, I will use known English words to express words whose meanings children already know and nonsense words to express words that are unfamiliar to the children.

Case 1: (*Mary*)

The child knows that *Mary* picks out a THING and that it is "referentially complete". Thus the child will categorize it as an NP.

---

[16]Note, however, that the sole claim is that children do not do a *syntactic* decomposition; there is no reason to doubt that they know enough about word boundaries and morphology to understand that (e.g.,) *kicked the bucket* has three separate words.

Case 2: (*water*)

The child knows that *water* picks out a THING but that it is not "referentially complete". Here the child will categorize it as a noun and will posit a phonetically empty determiner as being adjacent to the noun.

Case 3: (*fep dog*)

The child knows that *dog* picks out a THING but that it is not "referentially complete". Since all nouns must co-occur with a determiner, the child categorizes the unfamiliar word *fep* as a determiner.[17]

Case 5: (*this zoop*)

The child knows that *this* is a determiner and since determiners must always co-occur with nouns to form NPs, *zoop* must be a noun.

Case 6: (*this one*)

This is a special instance of Case 5 above; if the child has mistakenly categorized *one* as an NP; the determiner will cue the child to re-categorize it as a noun.

Case 7: (*dax moop*)

The child knows that this set of words picks out a THING but doesn't know whether it decomposes into a determiner + noun or whether it is a lexical NP. The default is to store it in the lexicon as an NP, i.e., as $dax\text{-}moop_{NP}$. (Note that the same procedure would apply if there was just a single word; if the child cannot decompose it, the default is to store it as a lexical NP.)

Case 8: (*the big cheese*)

This is a special instance of Case 7 above; since the child cannot semantically decompose it, the string is stored as a lexical NP.

---

[17]A problem with this is how children exclude the possibility that *dog* takes a null determiner and that *fep* is an adjective, as in in *big dog*. One gap in this account is that I have no concrete proposal of how young children distinguish adjectives from determiners -- though they can clearly do so (Gordon, 1987).

The only serious learnability problem that remains (besides the Gleitman critique, see below) is expletive NPs. It would seem that these are best viewed as on a par with other non-THING NPs, such as *the time* as in *I asked Mary the time.* Just as children who first acquire an NP as a THING can later extend it to encode other semantic features such as PROPOSITION, perhaps they can later use it with no semantic feature at all. This leaves us with:

Case 9: (*it* as in *it is raining*)

If a phrase that has been previously categorized as a THING is used without any cognitive role, children will assume that it is an expletive and exists to satisfy some grammatical constraint.

This leads to two predictions: (i) children should only use an NP as an expletive after they have acquired it through its having the feature THING (ii) no language should have expletive NPs that are solely expletive and have no referential double-life. In other words, no language should have a word that is used solely in expletive contexts, because, according to the above acquisition claim, children would never be able to acquire such a word. Nishigauchi and Roeper (1987), who present a similar theory of the acquisition of expletives, argue that both predictions hold.

Note that all of the nine cases of learning discussed above fall out from the representation theory; there are no assumptions about the relationship between grammar and non-linguistic cognition that are not otherwise motivated by the demands of a theory of adult competence. The only partial dangler is the "primacy of NPs" hypothesis, and while this may be special to acquisition, it might also be a more general principle of language parsing. In any case, it is not a principle regarding the semantic nature of linguistic categories and as such its existence does not go against the SCH.

3.5.3. How do children learn which words and phrases are THINGS?

I want to return to the last of the problems mentioned in Section 3.3. Some of the learning proposals above assume that children have the capacity to associate a word or string of words with the feature THING in the absence of semantic support from other words. The child must also be capable of realizing which words and phrases are "referentially complete" and which are not. This posits some (however minimal) ability to link up words with these semantic properties prior to learning their syntactic category, and thereby forces us to consider Gleitman's arguments about the difficulties of such a procedure.

It should be emphasized that Gleitman's learnability arguments are directed at the acquisition of subtle aspects of verb meaning; she is not making the claim that the child needs access to syntactic structure before learning that *dog* names a kind and *Fred* refers to an individual. Such a claim would be overly strong and there is good reason to believe that it is false. Names for objects, substances, and people are used appropriately by children who are younger than one-and-a-half years old (e.g., Huttenlocher and Smiley, 1987) and they can be understood by children less than one-year old (e.g., Benedict, 1979). Though it is always risky to make claim about children's ignorance, it seems very unlikely that children this young are parsing the speech stream into NPs and VPs – they show no evidence of any syntactic knowledge at all. Given this, it seems that at least some word learning can proceed in the absence of syntax. But it is worth addressing the question of how this is done.

One might think the problem has already been solved, at least to the extent necessary for the purposes here. If THING is cognitively related to material entities then it might seem simple for children to determine the semantic status of at least a few words. It doesn't work for the vast majority of nouns, which refer to non-material entities, but this doesn't matter -- children can learn these nouns later, through the semantic information provided by other parts of speech. For now, all children have to do is hear a word (e.g., *table*), follow an adult's point (or use some

other means to determine the implied reference), and note that the adult is referring to an object. The child could then categorize *table* as a nominal. If the child has the additional knowledge that people and some animals get their own names and other entities do not, this can account for why the child would take *table* as the name of a kind, and thus as a noun, and *Mommy* as the name for a person, and thus an NP.

This may happen, at least for children growing up in North America, but consider a very different scenario, where children are not assured of having entities named for them. One case in point may be the acquisition of language by Kaluli children, whose mothers do not engage in naming behavior (with one notable exception, see below). Instead, Kaluli mothers provide their children with extensive training in conversational interaction, often modeling sentences for them (Schieffelin, 1985). Despite this difference in input, children acquire words at much the same rate as children acquiring languages such as English (Gentner, 1982; Schieffelin, 1985). To see the problem here, imagine such a child seeing a bird in the sky and then being told (say) *zav goop wicket mep*. By hypothesis, the child knows no words at this point, no inflections, and nothing about the syntax. Which word (if any) does she assign the semantic feature THING to? How do Kaluli children learn *any* word meanings?

There are two possible solutions. The first is to bite the bullet and assume that children perform the sort of extensive analysis that Gleitman worries about. They break the sentence up into words or phrases (perhaps using prosodic cues to do so) at the same time as they cognitively divide the world around them into entities and processes, and attempt to determine which words and phrases relate to which aspects of cognition. Indeed, if it turns out that (i) children younger than a year of age do not know the syntactic categories that words belong to, (ii) children only hear words in context, and (iii) children learn the meanings of words, then something like this *must* be going on.

There is an alternative, however, and this is to deny one of the premises above. Perhaps all children hear *some* words in isolation, and these can be syntactically

categorized and thus get children into the language system, allowing them to categorize other words through linguistic inferences of the sort discussed earlier. One class of words that all children might hear in isolation are proper names -- referentially-complete names for individual THINGS.

There is some evidence that proper names are privileged in lexical acquisition. First, they appear among the very first words of children learning a range of different languages (Gentner, 1982) -- including Kaluli children. Second, proper names might be the one class of words which all children are guaranteed of being exposed to in isolation, or at least in some special stressed context. The motivation for this is the intuition that regardless of the ideology of a society concerning the child's acquisition of names for kinds of entities, it is unlikely that any society assumes that the child is born knowing the proper names of the people surrounding him or her. These must be taught.

This predicts that even in cultures where adults do not standardly label objects for children, proper names will be taught to children. If we turn to the Kaluli children, we find that this is the case. Schieffelin (1985: 534) claims

> ... this aspect of language, the saying of names, is downplayed in Kaluli families. There are no labeling games to facilitate or encourage the learning of object names. This is primarily due to the linguistic ideology of the culture. It is only in families who are acquiring literacy that one sees any attention paid to saying the names of objects, and this activity is initiated by the child when the mother is looking at books. When extended by the child to other contexts, the mother's response is disinterest.
>
> In contrast, because of the cultural importance placed on learning the proper names and kinterms of the individuals with whom they interact. Kaluli children are consistently encouraged to master a large number of proper name, kinterms, and other relationship terms ...

This supports the view that all cultures have some class of nominals that are special with regard to interaction with children. In Western societies, this is a very

broad class, including virtually all object and substance names; in the Kaluli society it is much more narrow, restricted to proper names and relationship terms. It might be that all cultures will treat proper names as special in this regard.

The alternative is that children are capable of somehow learning the meanings of words that are embedded in sentences. The force of the Gleitman (1990) critique, then, is not that syntax is necessary to learn word meanings (it clearly is not) but rather that an adequate theory of lexical acquisition must assume either stronger extrinsic constraints (all cultures use some nominals in isolation) or a more powerful learning mechanism (one that can learn words that are not presented in isolation). Either possibility would be of considerable interest.

3.5.4. Summary

The key claims about the acquisition process are the following:

- Children can infer the linguistic categories that new words belong to through cognition-semantics mappings.
- Children can infer the linguistic categories that new words belong to through the semantic selectional properties of other words.
- Children can infer some aspects of word meaning from the semantic structure of other words and considerably more from the conceptual structure of other words.
- Children's initial hypothesis is that a new word or combinations of words that has the feature THING is an NP; only if there is further evidence is it decomposed into a determiner followed by a noun.

One frequent complaint about semantic theories of language acquisition is that they are so vague as to be unfalsifiable. I am sensitive to these concerns, since at points -- for instance, in the discussion of abstract mental spaces -- I had to settle for making general speculations. Nevertheless, this chapter has made some reasonably

strong claims, all of which could be proven false. These include claims about lexical semantics (e.g., the semantic slot THING can only be filled by NPs), about children's ability to learn the grammatical categories that new words belong to (e.g that they can attend to both verb semantics and non-linguistic cognitive cues), and about children's capacity to acquire word meanings through linguistic information (e.g., that they can draw inferences from both the semantic structure and the cognitive structure of other words). Less integral claims concerned the acquisition of idioms and expletive subjects.

In the next three chapters, I extend the semantic theory presented in this chapter to a more modest domain -- the distinction between count nouns and mass nouns. Experimental evidence will be presented showing sensitivity to subtle properties of semantics-cognition links on the part of naive adults and even more naive children.

Chapter 4

Semantics and the count/mass distinction

> No differentiation was made between the standby cars and the ones they often replaced. In the jargon of the precinct they were *all* called "the junk". Cops would pile into the junk when their tour of duty started and would drop off the junk when the tour ended. The junk was both singular and plural. One patrol car was the junk. Six patrol cars were the junk. When a car broke down, it was called "the fuckin' junk". To listen to the motorized cops of Eight-Seven, you'd have thought they were narcotic dealers. – Ed Mcbain, *Eight Black Horses* (1985), pp. 99-100.

4.1. Introduction

In many languages, including English, there are two types of nouns: count nouns and mass nouns. Only count nouns can follow *a, another,* and numerals, and only they can be pluralized (e.g., *a dog, five rocks*). Mass nouns cannot be syntactically individuated, but they can follow determiners that pick out an indeterminate quantity of what the noun refers to (e.g., *much water, a little beer*), and they can appear with classifier constructions (e.g., *a piece of furniture*).

The nature of the count/mass distinction has engendered considerable controversy. It is clear that objects tend to be described by count nouns and non-solid substances tend to be described by mass nouns. This is unlikely to be a coincidence, and it leads to the hypothesis that the only entities that are described by count nouns are those that have perceptually salient boundaries, and thus those we are able to count, while mass nouns describe everything else. It makes sense to

speak of *a car* or *a hundred dogs* since cars and dogs are the sort of things that fall into discrete units. In contrast, non-solid substances and groups of homogeneous particles have no perceptually salient boundaries; it is not obvious, for example, how to make sense of locutions like *a water* or *a hundred clays*. You cannot count water or clay. One plausible claim, then, is that the count-mass distinction within a language directly reflects the properties of what the nouns refer to or, to take a more temperate view, at least reflects the human conception about what the nouns refer to.

I will argue in this chapter that some version of this claim is essentially true -- that the count/mass distinction is fundamentally semantic and related in a critical way to the perceived individuatability of what the nouns refer to. However it is equally clear that the count/mass distinction is not defined either in terms of objective properties of entities in the world or in terms of perceptual classes. In particular, the count/mass distinction cannot be defined as a distinction between objects and non-solid substances or between objects and everything else. There are three main reasons for this, each of which has been noted by several scholars (e.g., Bloomfield, 1933; Gathercole, 1986; Gleason, 1969; Gordon, 1985; McCawley, 1975; Quine, 1960; Ware, 1979; Weinreich, 1966; Whorf, 1956)

Abstract nouns. The vast majority of nouns don't refer to either objects or substances but nevertheless they get either count or mass syntax. For example, *opinion* is a count noun and *advice* is a mass noun. But it doesn't make any sense to say that opinions are objects and advice is a substance. In fact, it is not at all obvious that opinions and advice are entities *in the world*, existing independently of human cognition. In any case, the fact that such nouns are categorized as count or mass suggests that the count/mass distinction cannot be reduced to objects vs. non-solid substances or to any other material contrast.

Mass nouns that name objects. There is the puzzle of superordinate mass nouns, an apparent anomaly which we will be concerned with throughout this chapter. These are nouns that refer to higher-level categories and which have mass syntax,

despite the fact that they appear to describe discrete objects. So for example, consider the nouns below:

(1) animal, building, flower, tool, toy, vegetable, vehicle, weapon
(2) clothing, footwear, furniture, headgear, jewelry, linen, money, silverware

The nouns in (1) are count superordinates, you can say *an animal* or *two weapons*. But some other superordinates, like those in (2), are mass nouns; you cannot say *a furniture* or *two moneys*. On the face of it, this looks arbitrary. Why are *animal* and *building* count nouns and *clothing* and *furniture* mass nouns? The very existence of object names that are mass nouns is further evidence against reducing the count/mass distinction to a contrast between objects and non-solid substances. But a stronger implication one might draw is that the count/mass status of these words has absolutely nothing to do with what they mean. This would suggest that the count/mass distinction is arbitrary, perhaps random.

Multiple categorizations. Perhaps the most puzzling cases are those where the very same entity in the world can get two names, one count and the other mass. This shows up cross-linguistically; for some words, languages differ as to whether they get categorized as count or mass. For instance, the French equivalent of *furniture*, which is a mass noun in English, is a count noun: *meuble*. Food terms like *rice, beans* and *peas* also differ across languages (Wierzbicka, 1985), as do abstract nouns like *information, advice* and *news* (Mufwene, 1984), and superordinates like *furniture* and *jewelry* (Markman, 1985).

Note that this is not the same issue as the fact that some language do not encode the count-mass distinction at all (see Greenberg, 1966). This does not matter as much, since a distinction does not need to show up in the syntax and morphology of all languages in order for it be counted as semantic. English lacks the shape-classifiers that Navajo has and the dual plural marking found in Irish -- but this is hardly an argument that these aspects of language are arbitrary. The point here is that for languages that *do* have the count-mass distinction, there are some

words that are categorized differently than in English. And if the distinction is semantic, wouldn't we expect it to be applied the same way across all languages?

This phenomena -- of the same entity being named with both count and mass nouns -- also shows up within languages; sometimes synonyms (or near-synonyms) differ as to their count-mass status. Consider the contrast between *poetry* and *poems, clothes* and *garments, humanity* and *people* (examples from Mufwene, 1984). In all cases, the first word is a mass noun and the second is a plural count noun. But the difference in meaning, if it exists, is surely not obvious.

Ware (1979: 22) sums up all these objections to a semantic theory -- abstract nouns, mass superordinates, and multiple categorizations -- in the following quote:

> [There is] a tremendous amount of variation that appears unnecessary and inexplicable. I have already pointed out the variation of non-concrete terms like *truth, hope,* and *justification.* If counting and measuring have anything to do with count and mass nouns, it is certainly not here. ... There is a count/mass difference between *fruit* and *vegetable* but they apply to things that for all accounts and purposes seem to be alike. Nor can I see anything that would explain the count/mass difference between *footwear* and *shoe, clothing* and *clothes, shit* and *turd* or *fuzz* and *cop.* These are normally count nouns and mass nouns for basically the same thing. It is also difficult to understand why *knowledge* is a mass noun while *belief* is normally a count noun when our theories tell us they are about such similar material. ... *furniture* is a difficult case. Although we usually count the pieces in a set, we still talk about how much or how little furniture someone has. Counting is much easier when we are dealing with furniture, but we put up with the mass noun. It would seem more sensible to use a count noun as do the Germans and French. ... Faced with evidence and examples like these, I become sceptical of finding a general distinction between count and mass occurrences with respect to individuating or anything else.

Ware might be right in being skeptical of semantically defining the count/mass distinction. But it is mistaken to view the phenomena he discusses as *refuting* a semantic theory of count-mass. They only refute a theory that would directly link

up grammatical categories with classes of entities in the external world. This is what Gathercole (1986) calls the "Ontological View", and it has been proposed by scholars such as Grandy (1973) and Moravcsik (1973). It is also assumed to be the semantic view in virtually all critical discussions of this position -- it makes for an easy target.

There is abundant reason to reject the ontological view; the considerations brought up by Ware and others definitively show that the count-mass status of a word cannot be directly predicted by the ontological status of the entity that the word picks out. But it is mistaken to view this as an argument against the sort of semantic theory that a psychologist would be interested in, because any theory requiring direct links between linguistic categories and ontological categories falls outside the domain of psychology anyway.

One premise of a materialist cognitive science is that our mental representations, including our linguistic ones, do not make direct contact with objective reality. Rather, our understanding of the world is mediated by our perceptual and (in some cases, at least) cognitive mechanisms. Thus the right sort of semantic theory would link up linguistic types with our *construal* of what entities there are in the world, not with the entities themselves. And grammatical categories, to the extent they are semantic, will systematically relate to cognition, not ontology. To get back to specifics, then, a semantic theory of the count/mass distinction is not embarrassed by the fact that the same entity can be described with both a count noun and a mass noun, because this alternation can be viewed as the result of shifting from one construal of that entity to another. Nor is it refuted by the failure to match up linguistic categories directly to ontological categories, because this sort of semantic theory denies that such a direct relationship exists. Mufwene (1984: 204) puts it as follows:

It should be made clear that although communication is about objects and states of affairs in some world, how speakers talk about these is determined more by how they as individual speakers or a community of speakers wish to conceive of/perceive them than by their ontological structures.

Since there is no psychologically direct mapping from psychological (or psycholinguistic) categories to entities in the world, no linguistically sophisticated semantic theory will require such a relationship. This is not really as controversial as it sounds. Indeed semanticists working within the framework of "model-theoretic semantics" (which critically involves links between linguistic categories and non-mental entities like possible worlds and situations) have long argued that their theories are *not* to be viewed as falling within the domain of psychology (e.g., Lewis, 1972; Montague, 1974; see Partee, 1978 for discussion). Indeed, Lewis (1972), in his critical remarks on Katz and Fodor (1963), has argued that the sort of program followed here isn't really "semantics", since it is not concerned with meaning in the proper sense of the term (Chomsky, 1981 makes a similar point). While this terminological imperialism seems non-productive, Fodor's (1987) phrase "psychosemantics" or Jackendoff's (1990a) "I-semantics" may be better descriptors of the sort of program carried out here.

Nevertheless, there are scholars who argue that I-semantics is inappropriate as a theory of cognition (e.g., Burge, 1979). In fact, Macnamara and his colleagues (see Macnamara, 1986; Macnamara and Reyes, 1990) have presented a theory that covers almost exactly the same domain that I am interested in, but which explicitly involves notions such as *truth* and *reference*. Leaving aside more general worries about the role of these notions in psychology, they are unlikely to be fruitful in the domain of count/mass for exactly the reasons summarized by Ware. Ontology is the wrong place to look for the difference between *shoes* and *footwear*, *clothing* and *garments*. A cognitive theory of semantics seems more promising in this domain.

This rejection of notions like reference can be taken too far. Some scholars who share my view of what an adequate semantic theory for psychology is (e.g.,

Jackendoff, 1983; Lakoff, 1987), also reject the notion that there is any scientifically interesting relationship between mental states and entities in the world, e.g., between the thought *Lo, a dog!* and a specific dog. Such a radical form of methodological solipsism seems ill-advised. It is not hard to think of scientific domains where the notion of reference would be essential. For example, it is likely to play an important role in any serious theory of the evolution of language and cognition. It is hard to see how one could explain (for instance) how the innate concept of "object" has evolved without specifying some sort of relationship between this concept and actual objects. Similarly the origin of the count/mass distinction may be indirectly rooted in the fact that the physical world contains a distinction between kinds like water and kinds like dogs; if the world (or human cognition) were very different, such a cognitive distinction might not exist. One should distinguish the (rather reasonable) claim that language interfaces with our construal of the world with the (overly strong) view that there is in fact *no* relation at all between language and reality. These issues get blurred because of a failure to distinguish a theory of the current workings of the mind with a theory of how the mind evolved (see Bloom, 1990c for discussion). So while the actual structure of the world does not directly determine the count/mass categorization of a word (because of the mediating role of perceptual-cognitive systems), it may be part of an explanation of how humans have come to evolve the distinction in the first place.

Given all of this, the fact that the same entity can be described with both a count noun and a mass noun does not in itself refute a semantic theory of count-mass, and neither do the existence of mass superordinates and abstract count and mass nouns. However, Ware does have some genuine concerns. The phenomena he discusses are important, and we should not take a theory of count-mass seriously unless it can provide some explanation for them.

In this chapter, the foundations of a semantic theory of the count/mass distinction will be laid out. I will argue that the grammatical difference between count nouns and mass nouns maps onto a contrast at a semantic level; this semantic

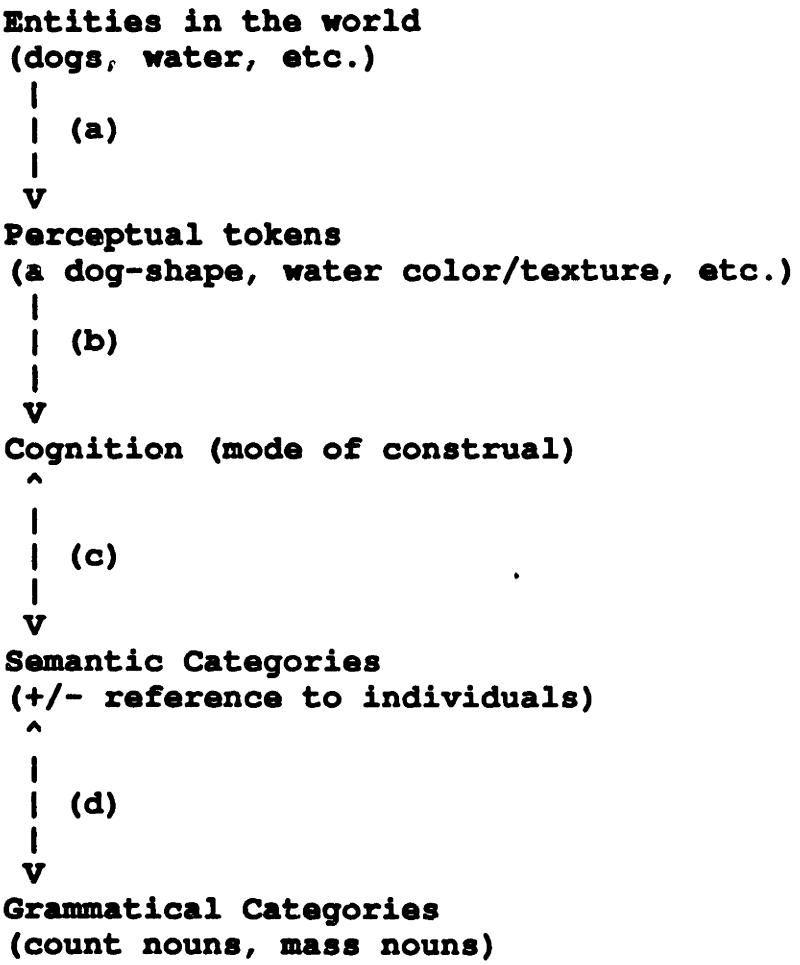
level relates to, but does not reduce to, non-linguistic cognition. In the next section, I defend the notion of an autonomous semantic level and discuss some proposals of how it might interact with the rest of cognition. Section 4.3 concerns Jackendoff's semantic theory of count/mass, which will provide the starting point for a somewhat different theory that I will outline in Section 4.4. In Section 4.5, I discuss the link between the semantic count/mass distinction and non-linguistic cognition. In Sections 4.6 to 4.9, some "hard cases" will be discussed: mass superordinates, cross-linguistic differences, and the categorization of nouns describing non-material entities. Empirical studies will suggests that these phenomena can be rather nicely explained through the right sort of semantic theory. In Section 4.10, some speculations about abstract nouns, like *knowledge* and *belief*, will be presented.

4.2. Cognitive architecture

4.2.1. Preliminaries

To set the framework for explaining these phenomena, we need to make it clear what it is to say that the count/mass distinction is semantic. What I have in mind is shown below in (3), the letters (a) - (d) denote the connections between different levels.

(3)



There are objective entities in the world and they leave perceptual tokens; dogs leave dog-percepts, water leaves water-percepts, etc (this is link (a)). The perceptual mechanism can be viewed as developing through maturation, not learning, and for the purposes here, we can assume the modularity position (Fodor, 1983). This is the view that perception operates independently of other aspects of cognition -- it is "informationally encapsulated". As a result, the link between perception and cognition (b) is one-way arrow from perception; it is not bi-directional.

Perception is indirectly reflected in grammar; a novel object (one that is perceived as a coherent bounded unit) is more likely to be named with a count noun than a novel non-solid substance (which is not perceived as a coherent bounded unit). And work by Soja and her colleagues (1987, 1990; Soja et al., 1990)

suggests that this distinction between objects and non-solid substances is also reflected in children's very first words -- a word for an object will be systematically applied along different dimensions than a word for a non-solid substance. Note that the distinction between objects and non-solid substances is not itself perceptual; it exists at the cognitive level (see Spelke, 1988 for discussion; see also Chapter 6). It is, however, tightly linked to perception -- one can usually tell whether an entity is an object or a non-solid substance just by looking.

The link between perception and grammar (or language in general) is mediated by cognition. With regard to words, we can describe the very same percept using words that have different meanings (the same dog-percept can correspond to *animal* or to *dog*). For count/mass, the same percept can be described with either a count noun or a mass noun (*a cake* vs *cake*, or *un meuble* vs *furniture*). Furthermore, it is clear that language can express thoughts that are not directly linked to perception. We can talk about *countries*, *conferences*, and *calculus*. Even the youngest of children can express thoughts about entities that are not present. All of this shows that language does not directly link up with perception, but does so only via cognition -- which is evidence for links (b) - (d).

The nature of link (c), the link from cognition to semantics, will be discussed below.

Finally, languages differ as to how semantic features are expressed in grammar (link (d)). Assume for the moment that the semantic nature of the count/mass distinction is a contrast between [+INDIVIDUAL] and [-INDIVIDUAL] (see Section 4.4). In English, this contrast is expressed through determiner choice; in Arabic it is marked with distinct numerical classifiers (Greenberg, 1978); and in some Amerind languages the distinction is not marked at all. Thus link (d), between semantic and grammatical structure, must be in part language-specific and therefore must be learned or triggered (see Chapter 6). It is, however, considerably constrained; as far as I know, all languages that convey the count/mass distinction do so through some sort of marker within the NP (determiners or classifiers). This sort of

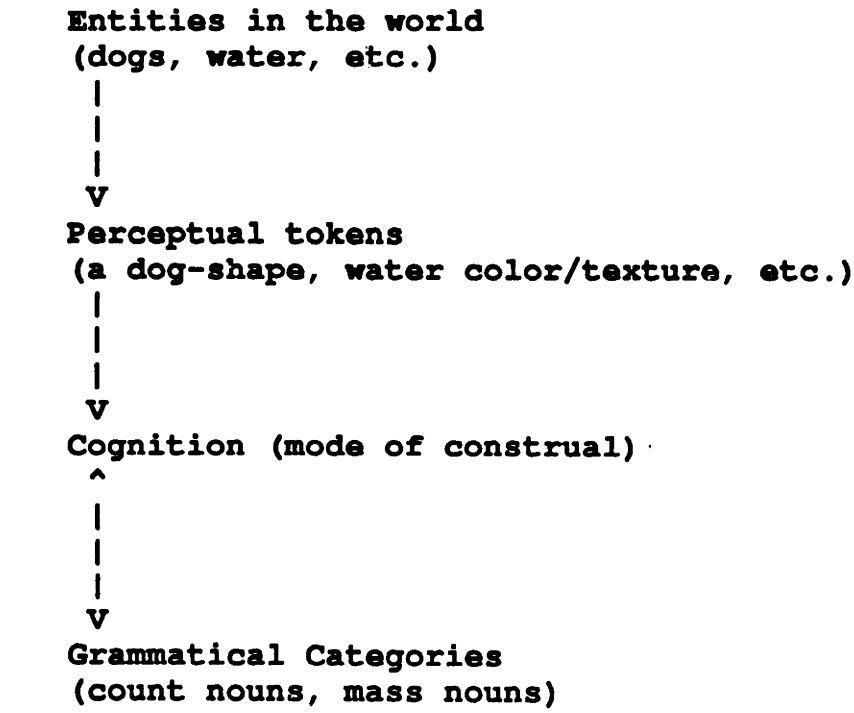
"constrained variation" appears in other domains where semantics and grammar interact; notions like TENSE, ASPECT, AGENCY, NUMBER, DEFINITENESS, MODALITY, and so on, are expressed in subtly different ways across different languages, but they also vary only within a relatively narrow range.

Two of the links are bidirectional, (c) and (d). The links between grammar and semantics, and semantics and cognition, go both ways, to capture the fact that the grammatical categorization of a noun can affect how we construe its referent. Some experimental evidence showing such an effect with young children is presented in Chapter 6, but it is not hard to see how this effect can work for adults. There is a considerable difference between hearing *there is cake on the table* vs. hearing *there is a cake on the table*; the grammatical contrast changes the meaning of the sentences and this affects our construal of the world. In general, hearing a novel word used with count syntax conveys the information that the word, whatever else it means, has the semantic feature of appearing in NPs that can refer to individuals. And this in turn has implications for cognition.

4.2.2. The autonomy of semantics

I have not yet defended the controversial part of the structure. Clearly the idea that there are perceptual mechanisms is not debatable (the modularity claim is, but it's not essential here). And nobody doubts that there are grammatical distinctions and categories, at least some of which relate to thought. To take another grammatical distinction having to do with nouns, the contrast between *dog* and *dogs* is strongly related to whether we are talking about one dog or more than one dog. And we not only need perception and grammar, we also have to propose that grammar links to non-perceptual cognition as well; consider the contrast between *day* and *days*. So far none of this is very exciting. The part which is debatable is the notion of an intermediate level between cognition and grammar. Other, more ontologically parsimonious scholars, such as Jackendoff (1983, 1990a) and Hale and Keyser (1987), would posit a structure much like the following:

(4)



In contrast I am adopting a version of what Pinker (1989) calls the "autonomy of semantics" claim, which is that there is no direct link between cognition and grammar; rather there is an intermediate level of semantics, which interacts with both of these other levels but which also has its own special properties and primitives. (Similar proposals have been made by Bierwisch, 1981 and Schlesinger, 1981.) Pinker has proposed this level to account for certain phenomena in the domain of verb classes, phenomena which also show up in the domain of count/mass syntax. I will review Pinker's (1989: 356-360) arguments, showing how they relate to the concerns that were discussed at the beginning of this chapter.

Pinker first notes that the sorts of grammatical contrasts between different verb classes cannot be described solely in terms of syntactic structure, but rather have some "cognitive content". For instance, the distinction between the selectional properties of *spray* vs *pour* (compare *spray the flowers with water* and **pour the flowers with water*) is related to how we think about spraying and pouring (see Pinker, 1990: 228-239 for discussion). Similarly, the nominal distinction between *dog* and *water*

(compare *a dog* and **a water*) has a lot to do with how we think about dogs and water. Thus for both verb classes and noun classes, there is some sort of link between grammar and non-linguistic cognition. The question now is whether this link is direct.

Pinker argues that it is not. For one thing, there are cross-linguistic differences. Thus "even in languages as closely related to English as Dutch, the equivalents of *say* and *suggest* are grammatical in the closest translation of the double-object form." (This is different from English, where **say me something* and **suggest me an idea* are ungrammatical). The same phenomena shows up for the count/mass distinction. In French, the word for furniture is acceptable as a count noun, while in English it is not. These differences are surprising. Since grammatical contrasts such as the double-object dative and the count/mass distinction are readily characterized in terms of cognitive content, it is puzzling to find this cross-linguistic variation. Do Dutch speakers think differently from English speakers about saying? Do the French think differently about furniture than the English? Moreover, there also exist within-language differences. Pinker notes that he disagrees with Green (1974) about the acceptability of *I dragged him the box*. He adds: "Yet surely Green and I do not have different conceptions about what dragging is". Compare the case of usages of count/mass syntax; Allan (1980) finds *admiration* acceptable as a count noun (his example: *Hermoine's is an admiration that I value greatly*). I find it very marginal. Some speakers use *sperm* as a count noun (e.g., *sperms*), others find this unacceptable.[18]

---

[18]These cases are much rarer for the count/mass distinction than for verbs, where there are many such cases of differing intuitions. This is probably because there are a lot of different verb classes, and someone can go a long time before learning the community standard for all of a verb's syntactic privileges. Nouns are much simpler -- a single utterance like *a pea* is enough to tell the child that *pea* has the semantic feature of being a count noun. Given that children have the capacity to determine semantic features through positive linguistic evidence (see Chapter 6) and given that the relevant evidence is so abundant, there is less possibility of intra-community differences, because anybody who differed from the community standard would quickly be guided to the norm through positive evidence. (The same sort of frequency effect applies to intra-community pronunciation differences -- all other things being equal, there is likely to be a lot less variation in how people pronounce a frequent word like *pliers* than an infrequent word like *Pleistocene*.)

Without a semantic level mediating between cognition and grammar, one is forced to claim that these differences in grammar reflect differences in cognition, something which Pinker views as "a very strong and implausible Whorfian claim". Pinker argues that there are no cognitive differences among speakers of these different languages and idiolects and that the way to explain the grammatical differences, while still accounting for the cognitive generalizations that alsc exist, is to posit the distinct level of semantics.

For the most part, I agree with Pinker's argument; cases where "cognitively motivated" grammatical alterations vary across and within languages motivate a semantic level between thought and grammar. If there was no semantic level, then every shift that is important in grammar should be important in thought, and the fact that there exists cross-linguistic variation suggests that this is not the case. But I disagree with his view that there are no cognitive differences corresponding to these grammatical contrasts. I will argue in the next section that there are excellent reasons for assuming that (to some very limited extent), the English word *furniture* has a different meaning (i.e., is thought about differently) than the French word *meuble*. If this is so, then the argument above must be weakened. It is incorrect to say that there are semantic contrasts that have no effect on cognition; rather, their effects are sometimes negligible, but they nontheless exist.

This relates to Pinker's next point. He notes that the grammatical subclasses of verbs are based on criteria that have little (but not nothing) to do with how people cognitively divide up classes of events for the purpose of induction, contemplation, etc. Thus *handling, carrying,* and *taking* are all counted -- from the standpoint of semantics -- as falling into three separate classes, despite their cognitive similarity, while *throwing, kicking,* and *rolling* all fall into the same linguistic class, despite their cognitive differences. In the domain of nouns, perhaps the most salient cognitive difference, that between material and non-material entities, does not appear at all as a grammatical contrast, while the more subtle contrast based on individuation puts *clothes* and *garments* into different classes, and collapses *justice* and *water* into the

same class. Once again, this suggests that grammatically relevant shifts in semantics may have little effect on thought, just as large differences in cognitive construal may have no effect on the semantic representation.

There is one final consideration. It is often noted that grammars only employ a small range of all of the possible meaningful distinctions and, more to the point, children only hypothesize among this range of possibilities (Pinker, 1984, 1989; Slobin, 1985). Thus no language would make a grammatical distinction between nouns for red things and nouns for green things and no language-learner would hypothesize such a distinction in the course of language acquisition. Children expect (and languages employ) grammatical distinctions on the basis of shape and animacy but not on the basis of (for example) color and weight, despite the fact that they can *think* fully well about color and weight. This suggests a further dissociation between the level of non-linguistic thought and the level of conceptual distinctions expressible in language (i.e., the semantic level).

4.2.3. Does semantics influence cognition?

It was proposed in Section 4.2.1 that the link between semantics and cognition is bi-directional. This is necessary if we are to account for the relations between grammatical input and non-linguistic cognition. Even if I don't know what *zav* means, I will make different inferences when told *I saw a zav* vs. *I saw zav*. And even young children are capable of making inferences about noun meaning on the basis of grammatical contrasts (e.g., Brown, 1957; Gelman and Markman, 1985; Gelman and Taylor, 1983; Katz et al., 1974; Soja, 1990; Waxman, 1990; see also the studies in Chapter 6). This suggests that grammatical encoding is systematically

linked (via semantic structure) with non-linguistic cognition.[19]

But there is a problem here, because there are also cases where speakers of different languages differ on how they grammatically encode a word, such as the count/mass difference between *furniture* and *meuble*. It may seem obvious that these cases show that some grammatical changes have no effect on cognition -- our intuition is that *furniture* and *meuble* mean exactly the same thing, despite their grammatical difference. This is the puzzle: on the one hand, there is all this evidence that grammar can affect cognition; on the other, there are all these obvious cases where it seems that grammar cannot affect cognition. How can we reconcile this?

One solution, advanced by Pinker (1989: 262), is that children who learn word meaning's through attending to grammar are engaged in "non-linguistic cognitive inference" and "riddle-solving". With regard to the count/mass distinction, one could suggest that children have inferred certain generalizations regarding the relation between count/mass syntax and word meaning. If it turned out that (i) these inferences are limited to a small domain (for instance, basic-level names for objects and non-solid substances) and that (ii) within this domain, there are no cross-linguistic differences in grammatical categorization, then our problem would be solved. Children *can* determine word meaning from grammar, but they can only do so in cases where the grammar does not differ across languages. So the child's ability is harmless; it does not lead to meaning differences across languages.

This proposal has serious empirical problems, however. It is incorrect that the

---

[19]More precisely, the developmental findings fall into two classes. Brown (1957) found that children are sensitive to syntax in determining which aspect of a scene to focus on; if they hear *a zav*, they look for an object, if they hear *zav*, they look for a non-solid substance. The other studies show a more radical sort of effect -- children will think differently about the meaning of a word for the *same entity* as a function of the word's syntax. For instance, some 2-year-olds will take a word for a given entity as having one meaning if it is presented in a count noun context and another different meaning if it is presented in a mass noun context (Soja, 1990). It is this second type of effect that is relevant with regard to cross-linguistic differences, because we know that the English word *furniture* and the French word *meuble* are used to describe the same entities in the world -- the question is whether or not speakers of the two languages construe the entities in different ways.

only role for syntactic inference is for material domains. For example, a computer scientist would construe *blicket* differently in *This program doesn't have any blickets* vs. *This program doesn't have any blicket*; the first sentence can refer to some discrete abstract entity like symbols, program lines, etc., while the second cannot. (Importantly, the scientist can make these inferences even if *blicket* does not mean the same as a word that already exists in the scientist's language). The studies presented in Chapter 6 also militate against the proposal that count/mass syntax can only influence cognition in material domains. Finally, there are cross-linguistic differences even for basic-level categories, (e.g., for food terms like *peas* and *beans*), which is the very domain where it has been shown that grammar *can* affect cognition in young children. It looks like we are stuck with the seemingly outrageous claim that if one person learns a word as a count noun and the other learns it as a mass noun, then they think differently about what that word means. Is this right?

Perhaps it is, but it is not as troubling a phenomena as it first appears. Thus [+/-INDIVIDUAL], which is, by hypothesis, a primitive at the semantic level, systematically interacts with induction and inference at the cognitive level. But there is a twist: whether or not a word is [+INDIVIDUAL] or [-INDIVIDUAL] *matters more for some domains than for others.* Thus for some basic-level material entities it matters a great deal whether to construe the entity as individuated or non-individuated. It would be impossible for a language to describe dogs with a mass noun because a critical aspect of the meaning of *dog* is that it refers to discrete individuals. Similarly, it is integrally tied to the meaning of *water* that it describes unbounded homogeneous portions, not countable entities. This sort of restriction also applies to some words for non-material entities; if another language has a word that means the same thing as the English word *day*, then this word must be a count noun, because it is part of the very meaning of *day* that it refers to bounded temporal individuals. In contrast, however, for words such as *opinion* and *advice*, it really doesn't matter that much whether you view the entities as individuals or

portions. As a result, there can be cross-linguistic variation as to whether the word for opinions is a count noun or a mass noun, and this grammatical differences would correspond to a cognitive difference -- but it would be largely irrelevant.

When thinking about the cross-linguistic differences, we naturally turn to the very cases where the differences in construal are the least. This is because if count noun *X* in one language and mass noun *Y* in another language differed a great deal in conceptual structure, then we would not be tempted to view *X* and *Y* as being greatly affected by count-mass syntax *because we would not think of them as synonyms in the first place.* So we are drawn towards the very cases where the shift in meaning is the smallest, cases like *furniture* vs. *meuble* which almost entirely overlap in how we think about them, and for these cases it is hardest to see the effects of count-mass semantics on our construal of what a word means. Nonetheless, the intuition is mistaken; findings about the effect of grammar on thought suggest that the English word *furniture* and the French word *meuble* are actually not synonyms -- even though they refer to same entities in the world, they map onto subtly different cognitive structures.

Some people find it hard to imagine any cases where such a shift in construal exists across languages. This reluctance may be due to a failure to distinguish a strong Whorfian claim (that speakers of different languages think very differently from one another) from the milder view proposed here (that grammatical structure can cause subtle shifts in construal, and that there is limited variation in the grammatical categorization of words across languages). As an aid to the imagination, here is an example. As noted before, the English word *cake* can appear in both count and mass syntax, with a corresponding change in meaning as a function of its syntax. *Chicken* and *lamb* are other examples, but some words only appear in one frame. *Cow* is only count, referring to the animal; *beef* is only mass, referring to the meat of the animal. Returning to *cake*, then, it is entirely possible for there to be two languages, one which has *cake* only as a count noun and the other which has *cake* only as a mass noun. (In each language, other words might exist for

the missing construal). If we think differently about *a cake* vs. *cake*, it follows that *cake-count* has a different cognitive representation than *cake-mass*, even though we might easily be misled into thinking that they are synonymous. It is this sense that *furniture* might differ from *meuble*. Thus it is fully possible for speakers of different languages to mentally represent the meaning of a word differently as a function of its syntax. In the absence of an alternative explanation for the semantics-cognition links that we know exist, the best bet is that cross-linguistic differences of this sort are not only possible, but actual.

In this light, some of Pinker's arguments for the autonomy of semantics need to be modified somewhat. Grammatical changes will always have the potential to affect cognition, because they change semantic features which can have inferential consequences. Thus aspects of grammar that are linked to semantic features cannot vary independently from cognition. But there is still the relevant dissociation; a major cognitive distinction (e.g., *dog* vs. *opinion*) is not reflected in grammar, while a major grammatical distinction can be trivial from the standpoint of thought (e.g., *clothes* vs. *garments*). Moreover, as discused earlier, the cognitive notions encoded in grammar (i.e., in closed-class words, in inflections and classifiers, and in verb classes) are a proper subset of the cognitive notions that play a role in non-linguistic cognition. These considerations are sufficient to motivate a semantic level within cognitive architecture.

4.2.4. Summary

There is a set of phenomena that any adequate theory must explain. These include:

1. The obvious cognitive motivation for at least some categorizations of count/mass (basic-level names for objects are count; basic-level names for non-solid substances are mass).

2. The correlation between meaning and form even for non-material domains (e.g., *day* as a count noun, *chaos* is a mass noun; see Study 3 below).

3. The cross-linguistic (*furniture/meuble*) and idiolectal (*sperm, admiration*) differences.

4. The fact that the count-mass distinction does not divide words on the basis of cognitive similarity (e.g., even though we intuit that clothes are more similar to garments than to mud we put *clothes* and *mud* in one linguistic category and *garments* in another).

5. The effect of grammar on cognition for at least some cases of material entities (e.g., *a cake* is construed differently than *cake*).

6. The effect of grammar on cognition for at least some cases of non-material entities (e.g., *a noise* is construed different from *noise*).

All of these considerations motivate the semantic theory outlined in this chapter. Any alternatives must also take these facts seriously. That is, the views that the count-mass distinction is "fundamentally arbitrary" or "an object/substance distinction" are wrong, and any theorist who merely describes it as "a semantic contrast", without further elaboration, is being radically incomplete.

What we need to do to support the framework above as empirically correct is to posit semantic definitions that systematically relate to both grammar and cognition. The test of such a theory is the extent to which one can predict which entities in the world will be named with count nouns and which will named with mass nouns, which cases will vary and which cases will not. Such phenomena are the result of semantics-cognition links. Further, the theory must capture linguistic distinctions within the nominal system (for instance, it should explain why mass nouns and plural count nouns pattern the same with regard to determiners like *more*). These are the result of semantics-grammar links, which are the topic of the following two sections.

4.3. Jackendoff's theory

4.3.1. Introduction

Jackendoff (1990b) presents a unified theory of certain aspect of noun/verb semantics -- here I will focus entirely on his discussion of nouns. He posits two semantic distinctions: boundedness (+/-b) and internal structure (+/-i). Individual objects and groups are bounded, and substances and aggregates are unbounded (-b). Boundedness is related to divisibility; for a noun marked (+b), like *apple*, "one cannot divide the referent up and still get something named by the same count noun", while for a mass noun marked (-b), like *water* "one can divide its referent up and still get something describable as water". For (-b) words, some (+i) refer to a "multiplicity of distinguishable individuals", others (-i) do not. This (-i/+i) distinction also applies to (+b) words; individuals have no internal structure and are marked as (-i), and groups do have internal structure and are marked as (+i). This is summarized below (from Jackendoff's (5); the examples are Jackendoff's):

(5)
+b, -i individuals (*a pig*)
+b, +i groups (*a committee*)
-b, -i substances (*water*)
-b, +i aggregates (*buses, cattle*)

Very roughly, the (+b) cases are singular count nouns, the (-b,-i) cases are mass nouns, and the (-b,+i) cases are plural count nouns. In addition, Jackendoff posits several correspondence rules; the descriptive success of these rules is the bulk of his argument for the existence of these features. Such an analysis is important, because there has been virtually no serious discussion of how nouns shift categories from count to mass and vice-versa. There is a tendency to ignore the phenomenon entirely and assume that nouns are fixed as either count or mass (e.g., Levy, 1988a), which obscures the relation between cognition and count/mass. Alternatively, some scholars take the existence of count-mass category shifting as evidence that

there is no "real" count-mass distinction: all nouns are indefinitely flexible (Pelletier, 1979). This is also mistaken; for example, *day* is a count noun in English and can never be used as a mass noun.

Jackendoff's functions can be summarized as follows. (I will merely describe them here and reserve critical comments for the following section):

PL ("Pluralization feature"). This changes the representation of an individual or group to an aggregate. Thus it changes *dog* to *dogs* and *committee* to *committees.* In other words, it takes a lexical entry that is (+b) (and either -i or +i) and changes it to (-b,+i).

ELT ("Element of"). This is (to some extent) the inverse of the pluralization feature PL; instead of taking a individual/group and outputting an aggregate, ELT takes an aggregate or substance and outputs an individual. Some expressions of this in English are terms such as *grain of* and *drop of.* This function takes a word for an unbounded entity (such as *water* or *rice*) and outputs something with the feature (+b,-i) -- an individual as in *a drop of water* or *a grain of rice.*

COMP ("Composed of"). This function is at work in phrases like *a house of wood* or *a house of bricks;* it takes a substance or aggregate and outputs an individual. Note also that this function does not have to be lexically encoded; it can be provided by a "rule of construal", where context requires that a substance be construed as an individual. This shows up in "restaurant talk", as in *I will have a coffee/three coffees* (see also Langacker, 1987).

GR ("Grinder"). This is the inverse of **COMP**. It takes an individual as an argument and outputs a substance. There is no phrase in English which does this -- it is, once again, only introduced by a rule of construal, as in *After the car hit the dog, there was dog all over the street.* (Langacker, 1987).

PART ("Part of"). This function takes a bounded entity (an individual or group) and usually outputs an identifiable bounded part, as in *a leg of the table* or *a part of the group*. It differs from **GR** in that it usually outputs a bounded individual, not an unbounded substance, but it can sometimes output an unbounded entity, as in *the blood of a pig*, It differs from **ELT** in that it takes an bounded entity as input, not an unbounded one.

CONT. ("Containing"). This function takes a part and outputs an entity containing a part (thus it's the opposite of **PART**). This may show up in phrases like *table with a drop-leaf* and *house with an orange-roof*. A second realization is in compounds such as *drop-leaf table*. (There is a similar form where a substance-part of a substance serves as the value -- as in *beef stew* or *blood pudding*).

With regard to the semantic vocabulary, we can summarize all of this as follows:

(6)
PL (PLURAL)
individual or group -> aggregate
(+b) -> (-b,+i)

ELT (ELEMENT OF)
aggregate or substance -> individual
(-b) -> (+b,-i)

COMP (COMPOSED OF)
aggregate or substance -> individual
(-b) -> (+b,-i)

GR (GRINDER)
individual -> substance
(+b,-i) -> (-b,-i)

PART (PART OF)
individual -> individual or substance
(+b,-i) -> (-i)

CONT (CONTAINING)
individual or substance -> individual or substance
(-i) -> (-i)

4.3.2. Criticisms

Jackendoff's proposal involves the right sorts of features and analysis, but there are some problems with it. My complaints are of the standard sort -- Jackendoff's formalism makes distinctions that are unnecessary, and as such obscures the real nature of semantic features within the nominal system.

For one thing, Jackendoff draws a distinction between words like *forest* and words like *dog*, because only the former has "internal structure". But there is no grammatical contrast distinguishing the two, e.g., there is no grammatical construction where *forest* can appear and *dog* cannot. Similarly, he draws a distinction between words like *water* and words like *rice*, because the latter has internal structure and the former does not. Yet they are both mass nouns and, once again, there is no aspect of grammar that distinguishes between the two.

Jackendoff also dismisses some real distinctions, such as between plural count nouns (*buses*) and some mass nouns (*furniture, rice*) -- both are classified as -b,+i (aggregates). But this leaves to be explained why only words in the first group can follow numerals and *many*, while only those in the second group can follow *much*. It also leads to problems with one of Jackendoff's correspondence rules: ELT. He proposes that this rule takes unbounded entities as arguments. But in fact, it only takes mass nouns, not plural count nouns (i.e., *a grain of* can only be followed by a mass noun like *rice*; a count noun like *peas* cannot appear in any such construction). And Jackendoff has no combination of semantic features that can define the category "mass noun". These difficulties motivate a somewhat different theory.

4.4. An alternative theory of the semantic level

4.4.1. The phenomena

To a first approximation, the relevant categories are those which are defined in terms of selectional privileges of determiners: singular count nouns, plural count nouns, and mass nouns. Thus *a* selects only singular count nouns, *many* selects only plural count nouns, and *much* selects only mass nouns. What also must be explained is the existence of determiners such as *the* which can co-occur with all nouns, and determiners such as *more* which select for both plural count nouns and mass nouns. Furthermore, plural morphology occurs with plural count nouns and singular morphology occurs with both singular count nouns and mass nouns. Ideally we should account for all such patterns; these are summarized in (7) (see also Gordon, 1982: 10 for a similar list).[20]

[20]Interestingly, there seems to be no aspect of English grammar that selects for both singular count nouns and plural count nouns, but not mass nouns. I am not sure why this is, or whether this pattern is universal.

(7)

	Singular Count Nouns	Plural Count Nouns	Mass Nouns
the	ok	ok	ok
(Also: *some, any*)			
a	ok	*	*
(Also: *one, another, each, every, either, neither*)			
many	*	ok	*
(Also: *several, few, both,* numerals other than *one,* plural morphology and plural anophora)			
much	*	*	ok
(Also: *a little,* classifiers such *a piece of*)			
more	*	ok	ok
(Also: *all, enough, less, most, plenty of, a lot of*)			
Sing. morphology	ok	*	ok
(Also: singular anaphora)			

4.4.2. The theory

While notions like "boundedness", "internal structure", "arbitrary divisibility", and "countability" are all related to the count-mass distinction, the proposal here is that the fundamental semantic contrast between count nouns and mass nouns is in terms of reference to individuals. That is, count nouns interact with determiners to form NPs that denote individuals, and mass nouns interact with determiners to form NPs that denote portions. By "portions", I mean that they refer to entities that are not specified as being composed of individuals (this is similar to Jackendoff's "-internal structure").

This is a fairly orthodox view, and has been defended (in somewhat different versions) by scholars such as McCawley (1975) and Mufwene (1981, 1984). It differs from other views in that it makes no claim about some special status for mass nouns (e.g., that they refer to "ensembles"; see Bunt, 1979). Mass nouns are viewed here as those nouns that simply lack the special property that count nouns have, that of participating in acts of reference to individuals.

We can express this through the distinction [+/-INDIVIDUAL]. This assumes that the default is mass -- that mass nounhood is just the absence of count nounhood. It should be stressed, however, that this is a claim about semantic primitives, not semantics-cognition links. In particular, it does not imply that there are no links from non-linguistic cognition to [-INDIVIDUAL]; the existence of mass nouns for solid substances (e.g., *wood, metal*) strongly suggests that one *can* categorize a new word as a mass noun on the basis of positive criteria. The motivation for viewing mass nounhood as the default is motivated instead by semantics-grammar links. As Allan (1980: 545) points out, in English, singular count nouns must appear with a determiner and plural count nouns must be appear with the plural morpheme -- but there is no necessary marking for mass nouns. Moreover, while there are some languages that have no count-mass distinction at all, there are no languages that mark mass nouns while leaving count nouns unmarked. All of this suggests that the best way to encode the contrast is in terms of [+/-INDIVIDUAL], rather than [+/-PORTION].

There is more to explain, however. For some purposes, plural count nouns and mass nouns fall into the same class. They can appear in bare NPs (i.e., without an overt determiner) and can follow determiners such as *more* and *a lot of*. How can we explain this? Obviously a simple binary distinction between count nouns and mass nouns cannot capture the required generalizations; we need something that is more fine-grained.

One proposal by Mufwene (1984) is that the feature distinguishing singular count nouns from mass nouns and plural count nouns is the feature of singularity [+/-SINGULAR].[21] Thus the contrast would be as follows:

(8)

	Singular	**Non-singular**
Individuatable	singular count	plural count
Non-Individuatable	????	mass

One minor problem here is that there is a gap in the feature system -- there is no clear candidate for a type of word that is singular and non-individuatable. But the major problem is that mass nouns are not plural; they are singular. They take singular verb inflection (*Water is wet*, not **Water are wet*) and they can enter co-reference relations with singular anaphors (*I saw the water and it was wet*, not **I saw the water and they were wet*). Thus the breakdown should be more like the following:

[21]By the same logic as above, it would be more accurate to view this feature as [+/-PLURAL], since singular is the grammatical default within language (Greenberg, 1978). But since nothing turns on this difference, I will adopt here the more standard [+/-SINGULAR] feature.

(9)

	Singular	Non-singular
Individuatable	singular count	plural count
Non-Individuatable	singular mass	plural mass

One advantage of the framework in (9) is that we can fill the missing gap -- there *are* plural mass nouns, nouns that cannot be individuated but which are syntactically plural. Examples of such words are the body part terms *guts, bowels, brains* and *hemorrhoids* (McCawley, 1975; Wierzbicka, 1985). These are mass nouns (*He doesn't have much brains* vs. **He doesn't have many brains*), but they are plural (*Brains are in short supply around here* vs. **Brains is in short supply around here*). Wierzbicka (1985: 332) provides an interesting suggestion about how our construal of these words explains their odd syntax:

> Each of these words designates an object .. which is clearly internally complex, having many similar parts, parts, however, which are not separate and which one couldn't, or wouldn't care to, count.

But this revised analysis leaves us without any explanation for what plural count nouns and mass nouns have in common. Gordon (1988) presents a theory that is restricted to asking the more specific question of why only these two parts of speech can appear as bare NPs. He asks this question from a learnability perspective: Given that children get no explicit negative evidence, how do they come to know that utterances like **look at dog* are ungrammatical? His explanation is as follows. Children learn that bare plural count nouns and bare mass nouns already have other semantic roles. Bare plural count nouns are used to establish generic reference to a kind of individuated entity (e.g., *dogs bark*), while bare mass nouns are used to establish generic reference to a kind of non-individuated entity (e.g., *water is wet*). Gordon argues that (i) children assume that the role of a bare

noun is to establish generic reference and (ii) children assume a uniqueness of form-feature mappings (e.g., Berwick, 1986; Pinker, 1984; Wexler and Cullicover, 1980). Thus children know that the only semantic role that bare singular count nouns could have is to establish generic reference (from (i)) but since that function is already taken over by plural count nouns, there is no semantic role that bare singular count nouns could fulfill and thus they are unacceptable (from (ii)).

The logic is right, but there is a generalization being missed. By only explaining why singular count nouns cannot occur as bare NPs, Gordon cannot account for why singular count nouns also do not occur with determiners such as *more* and *a lot of*. It is unlikely that "capacity to establish generic reference" *per se* is the solution; the potential to use *water* and *dogs* generically seemingly has nothing to do with the fact that we can say *more water* and *more dogs*, but not **more dog*. The solution is probably more general. (Though, of course, any more general solution must also solve the learnability problem -- see Chapter 6).

Jackendoff's formalism might be able to account for the phenomena, in that it categorizes plural count nouns and mass nouns as forming NPs that refer to unbounded entities (i.e., they are (-b)), and singular count nouns as forming NPs that refer to bounded entities (i.e., they are (+b)). This leads to the suggestion that only nouns marked (-b) can appear as NPs without an overt determiner, and can co-occur with determiners such as *more*.

It is not clear if this works, however. Plural count nouns do sometimes form unbounded NPs; thus *those dogs* is presumably unbounded in the same sense as *that water* is. But they can also appear within bounded NPs, such as *two dogs*. To use Jackendoff's own criteria; if you divide up two dogs, you no longer have two dogs, just as if you divide up a dog, you no longer have a dog. Thus two dogs seem to be just as bounded an entity as one dog, at least in the absence of any more detailed theory of the cognitive implications of boundedness. And this suggests that boundedness is not the proper way to distinguish singular count nouns from mass nouns and plural count nouns.

An alternative involves the semantic notion of extent. Consider the case of *more*. Very crudely, *more water* means more of the stuff that is water while *more dogs* means more of the individuals that are dogs. Nouns such as *dogs* and *water* form NPs that describe entities with some sort of EXTENT; to return to the theory discussed in the previous chapter, they are associated with regions in mental space that can be sub-divided. In contrast, we can view singular count nouns as forming NPs that describe atomic entities, indivisible points in mental space. It could be that some determiners (those that cause shifts in magnitude, like *most* and *a lot of*) must apply to nouns which are marked [+EXTENT]. More tentatively, one could speculate that the null determiner, which establishes generic reference, co-occurs with such nouns because the construal of generic NPs is associated with an extended region in mental space (see Chapter 3 for discussion). That is, the construal of *water* in *water is wet* or *dogs* in *dogs bark* might involve the essentialist intuition that there is "something out there", some real essence that the NPs refer to.

In sum, we can describe the phenomena in (7) by positing the following selectional privileges:

(10)

the [___noun]

a [___noun, +SINGULAR, +INDIVIDUAL]

many [___noun, -SINGULAR, +INDIVIDUAL]

much [___noun, -INDIVIDUAL]

more [___noun, +EXTENT]

Sing. morphology [_noun, +SINGULAR]

4.4.3. Correspondence rules

Returning to the correspondence rules discussed by Jackendoff, we can capture some of the same generalizations under the system above. I will ignore PART and CONT here, as there is no evidence that they are actually correspondence rules, as opposed to instances of the general capacity of language to productively combine words to form new meanings. So while it is likely that the shift from *dog* to *dogs*, or the shift from *coffee* to *a coffee*, are the results of mappings within the semantic level, there is no reason to think the same of the shift from *orange roof* to *house with an orange roof*.

PL (Pluralization) is accounted for as a shift, for [+INDIVIDUAL] words, from [+SINGULAR] to [-SINGULAR]. The function ELT (Element of) is viewed as a shift from [-INDIVIDUAL] to [+INDIVIDUAL], e.g., from *rice* to *grain of rice*. COMP (Composed of) is described the same way, but outputs phrases like *a house of wood*. While COMP is equivalent to ELT with regard to its semantic description, these rules differ conceptually, with ELT picking out an individual that is part of the stuff that the noun initially described, and COMP picking out an individual that is composed out of the stuff that the noun initially described.

GR (Grinder) takes a singular individual [+INDIVIDUAL,+SINGULAR] and outputs a mass noun referring to the stuff that this individual was made out of (e.g., *there was dog all over the street after the car accident*).[22] There is a puzzle with GR, however, that shows up in my account but not in Jackendoff's. While it may be marginal to use this rule of construal for singular count nouns like *dog*, it is terrible to use the rule for singular count nouns like *committee* (e.g., **after the committee was hit by a truck, there was committee all over the street*). Jackendoff's theory can account for this, because he draws a semantic distinction between words like *committee* (+i) and

[22]Some find these utterances fully acceptable, but there is some dialectal difference -- I have always found them rather marginal, only acceptable in a jocular sense. Pinker (1989), following Aaronoff (1983), notes that certain lexical rules are used only with special pragmatic effects, and there are differences in how willing people are to accept them.

words like *dog* (-i). I have argued earlier that this is the wrong contrast, because no aspect of language distinguishes the two. Is this a counter-example? Does the rule of construal that Jackendoff calls GR only apply to non-collective singular nouns, thus motivating his contrast between collective and non-collective nouns?

Perhaps, but there is another explanation for the phenomenon that does not require positing semantic features. It is simple: a dog is composed of "stuff", but a committee is composed of people. Thus GR cannot apply, since it must output something that is [-INDIVIDUAL] and there is no construal of *committee* that allows this. Thus the unacceptability of **committee all over the street* is due to how we construe committees and does not have to be expressed at the semantic level itself.[23]

Finally, there a correspondence rule that should be added: KIND. This is expressed in English as *kind of* or *type of*. It takes a word that is either [+INDIVIDUAL] or [-INDIVIDUAL] and outputs a [+INDIVIDUAL] word which describes subclasses of the extension of that word. Like GR, it need not be overtly present; *a good computer to buy is a Mac IIfx* is acceptable. Many cases where mass nouns are used as count nouns are due to the application of this rule, e.g., *Evian is a fine mineral water*. Finally, Gil (1987) notes that Indonesian has a special inflection for this interpretation, which is further evidence that KIND is a likely candidate for a rule at the semantic level.

To sum up, I have posited here a semantic theory that involves three binary semantic features: [+/-INDIVIDUAL], [+/-SINGULAR], [+/-EXTENT]. These semantic features can be used to define the linguistic categories "singular count noun", "plural count noun", "singular mass noun" and "plural mass noun", and some properties of determiners such as *the, a, much, more,* and *many,* as well as

---

[23]This assumes that the psychological notion corresponding to GR to not transitive, i.e., just because a committee is composed of people and people are composed of stuff, this does not imply that a committee is composed of stuff. A similar relation clearly holds in the linguistic expression *part of* -- even if Fred's arm is part of Fred and Fred is part of the committee, Fred's arm is not part of the committee.

singular/plural morphology and anaphora. Finally, the correspondence rules discussed by Jackendoff can be accounted for.

In Chapter 6, I will return to this feature calculus in the context of a theory of how children learn the selectional properties of English determiners and how they categorize new nouns. Below I discuss some ways in which the semantic feature [+/-INDIVIDUAL] interacts with cognition and perception.

4.5. On the link between [+/-INDIVIDUAL] and non-linguistic cognition

Not surprisingly, the proposal is that the semantic feature [+INDIVIDUAL] maps onto a corresponding cognitive notion of "individual". This cognitive notion is related to aspects of thought such as countability, divisibility, boundedness, and spatial extent, and is roughly equivalent to "discrete bounded entity". Boundedness is a property of single individuals, not necessarily of groups of individuals, and thus there is no requirement that plural count nouns -- which describe multiple individuals -- pick out a bounded set.

The mapping between the semantic feature and cognition is most transparent for words that denote material entities -- names for bounded objects are count nouns, names for unbounded substances are mass nouns. In a sense, this brings us back to the observation made at the beginning of the chapter: words like *dog* and *table* are count nouns and words like *water* and *sand* are mass nouns. But the explanation is *not* that there are direct links between count/mass and the cognition notions of object and non-solid substance. Rather, there are direct links between count/mass semantics and the cognitive contrast between individuals and portions. With some notable exceptions (see Chapter 6), this cognitive contrast shows up *in the material domain* as a contrast between objects and non-solid substances. But it also shows up in cognition about non-material entities; thus *day* is a count noun because it denotes a bounded unit of time, and *race* is a count noun because it denotes a bounded event.

Wierzbicka (1985) presents a similar account of the cognitive correlates of the count/mass distinction and shows how it can explain some of the cases that are frequently taken as arbitrary. Much of her analysis turns on names for foods. Why are the English words *pea* and *noodle* count nouns, while the words *rice* and *flour* are mass nouns? Wierzbicka's answer is fairly simple: individual peas and noodles are *bigger* than individual particles of rice and flour, and thus peas and noodles are more readily conceived of as being composed of individuals (i.e., as discrete bounded entities) than are rice and flour.

Wierzbicka notes the cross-linguistic difference here; in Russian the words for peas and beans and rice and flour are all mass nouns, yet nevertheless the word for broad beans (lima beans) is a count noun. One way to characterize the cross-linguistic phenomena is that languages have different ranges in which they split up a domain into count and mass; very crudely, for English basic-level categories, everything the size of a pea and bigger is construed as an individual and gets described by a count noun, while in Russian something has to be as big as a lima bean to be described by a count noun.

In general, then, there are two somewhat different reasons why a word could be a count noun in one language and a mass noun in another. In one set of cases, the word's semantic status plays a relatively small role in how we conceptualize its meaning (see Section 4.2.2), and thus we can view the word as referring to either individuals or portions without it causing a drastic effect on our mental life. This is the case for abstract nouns like *information* and *news*. But a second set of cases is when there are two ways to construe a given entity in the world -- either as a set of bounded individuals or as an unbounded portion -- and people can differ in how they categorize it. This construal will then link up with the grammar and some languages will encode the word as a count, and others as a mass noun. Languages will *not* vary randomly in this second case; subtle factors such as the manner in which the entity is used will partially determine whether it is encoded as [+INDIVIDUAL] or [-INDIVIDUAL].

Two domains will be studied in the following sections: mass superordinate nouns and words for non-material entities. It will be argued that the intuitions of naive subjects regarding the properties of what these nouns describe -- in particular, their intuitions concerning properties related to individuatability -- will be strongly correlated with the noun's count/mass status.

4.6. Basic-level and superordinate nouns

Basic-level categories have a special status in language and cognition. The basic-level has been defined as the level at which members of a category are perceived of as being most similar to each other and are least similar to members of other categories (Rosch, 1979; Rosch et al., 1976). So *dog*, for instance, is a basic-level category; dogs are relatively similar to each other and are relatively different from cats, birds, etc. Virtually all of children's first nouns refer to basic-level categories (see Gelman and Baillargeon, 1983) and so do the nouns that parents use when naming things for children (Anglin, 1977; Callanan, 1985; Shipley et al., 1983; Wales et al., 1983; White, 1982). The semantic categorization of these basic-level nouns can be readily predicted by their perceptual properties; basic-level object names are count nouns and basic-level non-solid substance names are mass nouns. This is not surprising under this account; it is at the basic-level where perception interacts directly with cognition (Fodor, 1983), and thus the perceptual correlates of boundedness will interact most directly with semantic structure.

Superordinates are a different story. These refer to categories whose members are more dissimilar than the members of either basic or subordinate categories. Parents tend to use superordinate nouns to describe groups of objects, not individuals (Callanan, 1985; Horton and Markman, 1980; Shipley et al., 1983; White, 1982). Thus adults would very rarely call a single dog, *an animal*, but would reserve the superordinate for describing a group of different animals.

Since these words are often used to describe groups, the perceptual system does not provide an unambiguous categorization and therefore there are different ways

we can semantically represent them. A pile of laundry can either be construed as an unbounded portion (a mass noun: *clothes, laundry*), or as a set of individuals (a count noun: *garments*). Thus languages vary; in some languages, *clothing* is encoded as a count noun; in others it is encoded as a mass noun. This is exactly the same situation as with the food terms discussed by Wierzbicka (1985) -- *beans, rice, spaghetti,* and *peas* -- people have the capacity to conceptualize them as either discrete individuals or as substances with indeterminate boundaries. It is names for these entities whose syntax varies across different languages.

It would be evidence for this view if it turned out that even for cases that vary across languages, the variance was non-random. There should be systematic differences between which superordinate categories are named by count nouns and which are named by mass nouns. One predicted difference involves what we can call the "perceptual heterogeneity" of a category. The intuition is as follows: If the members of a category look very different from each other, individual members of the category may be more cognitively salient and thus there may be more of a tendency to describe the category with a count noun. Conversely, if all the members of a category look alike, the individual members of the category will be less salient and the name for it may be more likely to be a mass noun.

The assumption underlying this prediction is that a perceptually heterogeneous array of items is more likely to be construed as a set of individuals than a perceptually homogeneous array of items. Thus a superordinate that describes a group of dogs, monkeys, snakes, and so on (i.e., *animal*) is very likely to be categorized as [+INDIVIDUAL], because we are very likely to construe such a group as being composed of individuals -- the members of the group "stand out" from one another. Compare this to a superordinate describing knives, forks, spoons, and so on (i.e., *silverware*) or to a superordinate denoting entities like pillow-cases and sheets (i.e., *linen*). Since these items look pretty much alike, we are unlikely to view them as separate individuals (when they are in the context of a group) and therefore *silverware* and *linen* are likely to be categorized as [-INDIVIDUAL] and thus to be mass nouns.

It should also be stressed that the perceptual heterogeneity hypothesis is not a claim about how superordinates are acquired. While children may acquire some superordinates through perception -- and thus perceptual heterogeneity may play some role in acquisition -- children can also determine the status of a novel superordinate by attending to linguistic input (e.g., hearing *That's a piece of furniture* will tell the child that *furniture* is a mass noun). Rather, the prediction is that we are more likely to construe perceptually heterogeneous entities as individuals, and thus superordinates that name such entities are likely to be marked with the semantic feature [+INDIVIDUAL] when they enter the language.

If any of this is true, then English superordinates should differ along the lines suggested -- count superordinates should be more perceptually heterogeneous than mass superordinates. This was tested in the following study.

4.7. Study 1 -- Intuitions about count and mass superordinates

4.7.1. Subjects

Thirty-seven MIT undergraduates volunteered for the experiment.

4.7.2. Materials

The subjects were tested on a subset of the 25 superordinate categories used by Markman (1985), which were originally chosen from studies by Ashcraft (1978), Battig and Montague (1969), Rosch (1975), and Rosch et al. (1979).

For a variety of reasons, several of Markman's categories were not used in this experiment. As Markman points out, the findings of Rosch *et al* suggest that *tree* and *bird* are actually basic level categories, not superordinates. Four other categories--*reading material, sporting equipment, human dwelling,* and *musical instrument*--were not used in this study because they are not single words in English, and it's not clear whether their syntactic status should be explained in the same way.

This leaves us with 19 superordinates (see the first row of Table 4.1). Eight of them are count nouns, eight are mass nouns, and the remaining three -- *food, fruit,* and *people*-- do not fall clearly into either category. *Food* and *fruit* can appear in both contexts. Consider the difference between *some fruits* and *some fruit,* and between *some foods* and *some food.* Mass syntax is used to quantify over portions, while the count syntax is used to quantify over the individual kinds (types of fruit, types of food). This type of ambiguity does not occur with any of the other superordinates. Expressions like **three furnitures* or **three silverwares* cannot be used to felicitously describe three distinct types of furniture or silverware. Similarly, the count superordinates do not have optional mass usage (e.g., **There is animal at the zoo* is ungrammatical).

People is standardly used as an irregular plural, e.g., *those people* to describe several persons. It is not clear what the status of such a word should be with regard to perceptual heterogeneity. It is a count noun, unlike words like *furniture.* But like the mass superordinates, it cannot be used to describe a single individual (i.e., **a people, *one people*). Such an irregular plural is a special case and there may be specific conceptual reasons why it has this categorization (Pinker, 1990). Finally, it is not clear that *people* is a superordinate at all; it could be the name for a basic-level category.

Due to their special status, *fruit, food,* and *people* were not used in any of the analyses. Subjects were tested on them, however, and their mean ratings are reported in the Results section.

The mean frequencies of the two groups of nouns were virtually identical: 40 for the count nouns, 43 for the mass nouns, $t(14) < 1.0$, $p > .50$. (from Kucera and Frances, 1986; frequencies reported are per million). Therefore, any significant differences between count and mass nouns in this experiment cannot be due to a difference in frequency.

Each superordinate noun was printed on a separate piece of paper. Below the superordinate was a 9-point rating scale with the endpoints labeled "very similar"

(1) and "very different" (9). All subjects were tested on the 19 superordinates and the order of the categories was separately randomized for each subject.

4.7.3. Procedure

Subjects were told to think about each category and rate it on either the extent to which its members look alike (i.e., the category's perceptual heterogeneity or "PH"). The relevant part of the instructions are as follows:

> What I want you to do for each category is to judge how perceptually similar its members are to each other. Suppose the category is *cup*. For me, all cups look pretty much alike, so I would circle a number closer to the left of the scale, where "very similar" is marked. On the other hand, suppose the category is *sporting equipment*. For me, they all look relatively different from each other; skis look different from baseball gloves and squash balls look different from hockey sticks. So I would circle a number closer to the right of the scale, where "very different" is marked.

Note that *cup* (the example of a low PH category) is a count noun and that *sporting equipment* (the example of high PH category) is a mass noun. Thus, if the examples lead to a response bias, this bias would work in the opposite direction as the predicted effect.

The procedure was identical to that of Gelman (1984). Subjects were given a minute to look over all the superordinates and familiarize themselves with them. Then they went through the pages one at a time. For each word, they were given 15 seconds to think about the category and 5 seconds to decide on and circle an answer.

4.7.4. Results

The prediction was that count nouns will be rated as describing more perceptually heterogeneous categories than mass nouns. Subjects rated the words on a 9-point scale, with 1 as "very similar" and 9 as "very different". The mean scores were 6.11 for count nouns and 4.00 for mass nouns. The subject analysis proved to be significant, as shown by a one-way ANOVA with noun type (count vs. mass) as the independent variable and with each subject's mean rating on that noun type as the dependent variable: $F(1,36) = 171.40$, $p < .0001$. The item analysis was also significant; there was a significant difference between the mean scores for the eight count superordinates and those for the eight mass superordinates: $F(1,14) = 10.69$, $p < .01$. The similarity ratings for each word are shown in Table 4.1. In post-test interviews, subjects were questioned to see if they were conscious of the distinction that was being studied. None of them was aware that the experiment was about the count-mass distinction and many were surprised when they were told that half the words were mass nouns.

Table 4.1 -- Mean judgments of Perceptual Heterogeneity

Noun	Mean rating
Count nouns	
animal	7.97
tool	7.16
toy	6.54
vehicle	6.08
weapon	5.84
vegetable	5.59
building	5.35
flower	4.32
--	
Mean:	6.11
Mass nouns	
clothing	6.20
furniture	5.57
footwear	4.49
jewelry	4.41
headgear	3.54
silverware	2.89
money	2.73
linen	2.14
--	
Mean:	4.00
Other nouns	
food	7.19
people	6.86
fruit	5.46

These results suggest that the count-mass status of superordinate nouns is related to how people conceptualize superordinate categories.

4.8. Study 2 -- Cross-linguistic analysis

If perceptual heterogeneity is a factor in determining the count-mass status of superordinates, then it should play a role in languages other than English. This can be examined in two ways: (i) There should be similarities in how nouns are categorized across different languages; (ii) Intuitions about perceptual heterogeneity should predict how frequently a word will be lexicalized as a count noun across languages.

4.8.1. Is there any relation between the English count/mass distribution and the distribution in other languages?

In order to study this, I analyzed Markman's (1985) data concerning which superordinates are count nouns and which are mass nouns in 18 languages from a variety of different language families (see Markman 1985 for details about how the data were collected).[24] For each language, Markman collected bilingual subjects' intuitions about the count-mass subcategorization of 25 superordinate nouns and 48 basic-level nouns. The subjects' responses fell into three categories: "count", "mass", and "other"--which meant that either the subject could not remember the word or it has no translation.[25]

When analyzing this data, Markman used a Guttman scale to see if there is a systematic order in which categories became introduced across languages as mass nouns. The scale determines if there is a hierarchy such that if a language represents a category high on the hierarchy as a mass noun it will represent all categories below it as mass nouns. The data was clearly not scalable and Markman concludes: "Thus the pattern in which categories are mass nouns cannot be predicted from one language to another." (1985: 41).

---

[24]I am grateful to Ellen Markman for providing me with these data.

[25]Two other response types were "count/mass", which meant that the word had two usages, one count and one mass, and "mass (count)", which meant that the word is primarily a mass noun but can be used as a count noun in some contexts. They were coded as "other" and "mass" respectively. These responses were rare, however, and the choice of how to code them has little effect on the overall significance of the tests reported below.

The following analysis differs from Markman's in two ways. First, only 16 category names -- the superordinates used in Study 1, excluding *food, fruit* and *people* -- are included. Second, the hypothesis here is not that there exists a systematic order in which categories are introduced as mass nouns, but simply that there is a significant relation between the count-mass distinction among English superordinates and the count-mass distinction in other languages. When comparing languages, "count" responses were dummy-coded as 1, "mass" responses were dummy-coded as 0, and "other" responses were not included. The correlations between English and each other language, along with the number of "count" and "mass" responses given by the informant, are shown in Table 4.2. (the breakdown into Family is from Markman, 1985).

Table 4.2 - Correlations between noun type of English superordinates and noun type of superordinates in languages from Markman's (1985).

Family	Language	# of Nouns (Count/Mass)	r
Indo-European	Afrikaans	5/8	0.76 *
Indo-European	Dutch	4/9	0.72 **
Indo-European	German	7/8	0.88 ***
Indo-European	French	10/5	0.76 **
Indo-European	Greek	13/2	0.42
Indo-European	Polish	8/5	0.73 **
Indo-European	Ukranian	6/2	0.75 *
Indo-European	Urdu	9/6	1.00 ***
Uralic	Finnish	14/1	0.29
Uralic	Hungarian	7/5	0.85 ***
Altaic	Turkish	4/6	0.46
Independent	Japanese	11/3	-0.10
Independent	Korean	10/2	-0.38
Afro-Asiatic	Arabic	7/8	0.34
Afro-Asiatic	Hebrew	14/2	0.00
African	Guro	5/6	0.83 **
African	Nzema	4/3	1.00 ***
	A.S.L.	3/10	0.43

* $p < .05$, ** $p < .01$, *** $p < .001$. All significance levels are two-tailed.

There were 18 languages studied and for 10 of them the correlation between the count-mass distinction in that language and the count-mass distinction in English is significant (for 2 languages, $p < .05$, for 4 languages, $p < .01$, for 4 languages, $p < .001$). 7 of these languages are from the Indo-European family; the only Indo-European language that has a count-mass distinction that does not come significantly close to that of English is Greek. Out of the 10 non Indo-European languages, there are three -- Hungarian, Guro, and Nzema -- that also have a count-mass distinction significantly close to that of English. When we compare the proportion of Indo-European languages vs. non Indo-European languages that have count-mass distinctions significantly close to English, the difference is significant, showing that there is a clear historical effect at work, $\chi^2(1) = 5.95$, $p < .02$, one-tailed. However, the fact that there are significant correlations between English and languages like Hungarian suggests that the cross-linguistic similarities are not entirely due to historical factors.

Moreover, there is reason to believe that sampling difficulties might lead us to underestimate the similarities across languages. In 8 out of the 10 languages where there was a significant effect, the sample had at least four count nouns and at least four mass nouns. This was the case for only 2 out of the 8 languages where there was no significant effect. The relationship between low sample size and significant correlations is significant, $\chi^2(1) = 5.445$, $p < .02$, one-tailed. Also, for all but 3 of the 18 languages that were studied, there was a positive correlation between its count-mass distinction and that of English ($p = .0038$, binomial probability test). In sum, if there was a larger sample of nouns, particularly for the non Indo-European languages, more of the correlations might have been significant.

In any case, other languages, even some from other language families, are similar to English in how they categorize superordinate categories into count and mass nouns. The existence of these correlations is encouraging; they support the claim that there is a systematic difference between categories that become count superordinates and those that become mass superordinates.

4.8.2. Is there any relation between perceptual heterogeneity and the count/mass distinction in other languages?

A more direct test, however, is to see if intuitions of perceptual heterogeneity are positively correlated with how often a superordinate is a count noun across different languages. Table 4.3 shows the numbers of times that each superordinate appears as a count noun across the 8 Indo-European languages and the 10 non Indo-European languages. These counts were correlated with the mean PH rating as well as the actual count-mass distinction in English (count nouns dummy-coded as 1, mass nouns as 0). The correlations are shown in Table 4.4, and they are significant for both Indo-European and non Indo-European languages.

Table 4.3 -- Number of languages where superordinates appear as count nouns

Noun	Indo-European languages (N=8)	Non-Indo-European languages (N=10)
animal	8	10
building	7	10
flower	8	8
vehicle	6	8
toy	6	7
tool	6	7
weapon	6	2
vegetable	4	2
clothing	2	4
footwear	1	5
jewelry	2	4
silverware	1	4
furniture	1	3
linen	1	3
headgear	1	3
money	1	0

Table 4.4 -- Correlations between number of languages in which a word is a mass noun and mean scores for each word

	Indo-European languages (N=8)	Non-Indo-European languages (N=10)
PH ratings (see Table 4.1)	0.62 **	0.54 *
English count-mass distinction	0.92 ***	0.59 *

* $p < .05$, ** $p < .01$, *** $p < 0.01$, All significance levels are two-tailed.

There is an interesting asymmetry between the predictive power of the mean PH rating of a category (i.e., how well it correlates with the number of languages where the category is lexicalized as a count noun) and the predictive power of the noun's count-mass categorization in English. First consider the Indo-European languages. Here, whether or not a word is a count noun in English is a far better predictor (r = 0.92) than its mean PH rating (r = 0.62). An analysis of the difference in effect-size (see Rosenthal and Rosnow, 1984: 372-373) shows this difference to be statistically significant (Z = 1.162, $p < 0.05$, two-tailed.) Contrast this with the non

Indo-European languages, where the two measures are virtually equivalent as predictors (r = 0.59 versus r = 0.54). This supports the view that there are two effects at work -- an effect of historical relatedness and an effect of perceptual heterogeneity. The predictive power of the English count-mass distinction drops when moving across language families because both effects are present within the Indo-European language family but only the shared effect of perceptual heterogeneity holds across both Indo-European and non Indo-European languages.

In sum, the cross-linguistic data show two things: First, the count-mass breakdown in English superordinates is similar to that of other languages, even languages that are historically unrelated to English. Second, the extent to which a category will be a count noun across different languages is related to how perceptually heterogeneous its members are. These findings strongly support the above account for why some superordinates are mass nouns.

4.9. Study 3 -- Novel words for sounds and sensations

This study concerns adult's capacity to infer a word's count-mass status from aspects of its meaning. It is obvious that people can do this for words describing material entities; if an object is described as *the blicket*, an adult will categorize the novel word as a count noun, if a non-solid substance is described as *the blicket*, the word will be categorized as a mass noun.

This study will test a stronger prediction, which is that adults can draw inferences from the meaning of words for non-material entities to their count/mass syntax. In particular, the study examines how the *temporal* properties of a noun's meaning determine its syntax. The hypothesis is that a word describing an entity that is temporally individuated is likely to be categorized as a count noun, while one describing an entity that is temporally continuous is likely to be categorized as a mass noun. This may be obvious; it is hardly an accident that a word like *sleep* is a mass noun (as in *I didn't get much sleep*) and a word like *race* is a count noun (as in *I didn't see many races*). But an experiment could provide some independent support.

4.9.1. Subjects

Twenty-four MIT undergraduates volunteered for the experiment.

4.9.2. Materials and Procedure

Each subject was presented with four stories. In each story, a new word was used to name either a sensation or a sound, which was described as either occurring as a set of discrete temporal individuals or as a continuous event. Following an idea by Gordon (1985), the words were presented as part of a noun-noun compound (with the key word in boldface) so that it could be taught to the subject without any clue as to whether it was count or mass. The stories were as follows:

Sound-Individuated

> John attached the **moop**-producer to his stereo and switched it on. It started off very loud and then was silent. Then loud, then silent. This went on for many hours.

Sound-Continuous

> John attached the **moop**-producer to his stereo and switched it on. It started off very loud and stayed loud, the volume never changed. This went on for many hours.

Sensation-Individuated

> The mad scientist attached the **wug**-producer to John's head. John felt the oddest sensation, something that was uncomfortable but not really painful and then it was gone. Then the sensation returned and a second later it was gone. This went on for many hours.

Sensation-Continuous

> The mad scientist attached the **wug**-producer to John's head. John felt the oddest sensation, something that was uncomfortable but not really painful. The intensity of the feeling did not change. This went on for many hours.

After each story were two sentences, one where the word was used as a count noun (e.g., *The machine produced a lot of zoops*); the other where it was used as a mass noun (e.g., *The machine produced a lot of zoop*). The subjects were told to circle a number from one to seven, according to whether they thought the sentence was "terrible" (1) or "perfect" (7).

Before the subjects were tested on the novel words, they were given the examples of *John walked into a* **furniture**-*store* and *John walked into an* **animal**-*store*, each followed by two sentences: *John saw a lot of furniture* and *John saw a lot of furnitures* for the **furniture** example; *John saw a lot of animal* and *John saw a lot of animals* for the **animal** example. They were told that most people preferred *furniture* to *furnitures* and *animals* to *animal*.

The four stories were presented in four consecutive pages, separately randomized for each subject. The first page presented the *animal/furniture* examples and these were read aloud to the subjects, who were then told to look at the stories that followed. For half the subjects the continuous stories were with *moop* (sound) and *nug* (sensation) and the individuated stories were with *zoop* (sound) and *wug* (sensation); for the other half the assignment of words to stories was reversed. For the test sentences, half of the subjects had the plural noun (count) first, the other half had the singular noun (mass) first. These conditions -- word-assignment and choice of initial test sentence -- were counter-balanced.

4.9.3. Results

The preediction is that for both individuated conditions there would be a preference for the count noun reading of the novel word and for both continuous conditions there would be a preference for the mass noun reading. For each story the subject's preference for the count noun reading was calculated by subtracting the subject's score on the mass noun test sentence from the score on the count noun test sentence. This was used in separate ANOVA's for the sound and for the sensation, with story-type (individuated vs. continuous), word-assignment, and

test-sentence order as the independent variables and the preference score as the dependent variable. The results are shown in Table 4.5.

Table 4.5 -- Subject's intuitions about count and mass noun readings of novel names for sounds and sensations, presented as either temporally individuated or temporally continuous.

	Count	Mass	Difference	Effect of grammar
Sound (individuated)	4.458	3.333	1.125	$F(1,20) = 31.19$, $p < 0.001$.
Sound (continuous)	2.125	5.510	-3.375	
Sensation (individuated)	5.000	2.953	2.047	$F(1,20) = 19.32$, $p < 0.001$.
Sensation (continuous)	3.000	4.583	-1.583	

For both the sound and sensation stories, there was a preference for the count reading when the story denoted a temporally individuated event (mean difference for sound: 1.125; mean difference for sensation: 2.047). Also, for both the sound and sensation stories, there was a preference for the mass reading when the story denoted a temporally continuous event (mean difference for sound: -3.375; mean difference for sensation: -1.583). As shown in Table 4.5, the effect of story type was significant for both the sound condition and the sensation condition.

One unexplained finding was that for the sensation stories, there was a significant effect of the order of test sentences -- the count noun preference was much bigger when the first test sentence was a count noun ($F(1,20) = 8.507$, $p = 0.009$). But there was no interaction with story-type, and when each test sentence order was analyzed separately, the difference between the individuated stories vs. continuous stories was still significant (First sentence count: $F(1,10) = 13.74$, $p = 0.004$; First sentence mass: $F(1,10) = 7.53$, $p = 0.02$.)

The study above confirms the intuition that we are likely to construe names for temporally individuated entities as count nouns and names for temporally continuous entities as mass nouns.

4.10. Conclusion

What I have tried to do above is sketch out a semantic theory of count-mass and show how it can capture linguistic and cognitive generalizations, as well as make predictions about the linguistic intuitions of adults. Admittedly, this theory is no great shakes from a formal standpoint -- it is not at all explicit -- but the point of it was to present a theory that interacted with cognitive systems in the right way. Thus the theory was an attempt to apply the notion of a semantic level to explain the existence of clear generalizations in the material domains (objects tend to be named by count nouns, non-solid substances named by mass nouns), while at the same time capturing some subtle generalizations among superordinates -- accounting for patterns within English and across different languages that had never before been explained. Further evidence for a semantic theory of count/mass was provided by the finding that adults are sensitive to the semantic basis of the count-mass distinction even for words describing non-material entities. The studies in Chapter 6 will show a similar capacity for 3- and 4-year-old children.

Perhaps Ware would still be skeptical. I have not yet explored truly abstract nouns such as *knowledge, belief, truth, hope,* and *justification,* and thus cannot show that he is incorrect, that the grammatical status of these words is semantically

motivated. My own intuition is that the meanings of abstract count nouns are really quite different from the meanings of abstract mass nouns. Consider *belief* and *knowledge*. Beliefs are often taken as discrete units in folk psychology and philosophy. People have beliefs -- usually described as sentential, composed of symbols and having internal structure -- that combine with desires to form actions. Knowledge is more of a continuous state, and it has no natural grooves. Thus it is not surprising that *belief* is a count noun and *knowledge* is a mass noun. There is nothing circular about making this argument; if the semantic theory is correct then the linguistic difference between these abstract nouns can be predicted by our *non-linguistic* intuitions, at least in principle.

In any case, there is a point where the burden of proof shifts. There are very clear ways that the count/mass distinction relates to cognition, and this shows up in the acquisition of grammar, the acquisition of word meaning, in systematic form-meaning correlations across and within language for names of both material and non-material entities, and in the choice of determiners associated with the sub-classes of nouns. If we abandon a semantic theory we are left with no explanation for these phenomena -- Ware offers no alternative theory, nor do other skeptics such as Gathercole (1986) and Levy (1988a), whose views we will discuss in the following chapter.

All the data are certainly not in yet. But, empiricist biases aside, there is nothing implausible about the idea of extending semantic contrasts to abstract thought. When somebody tells us they have *problems*, we can be sure that they have more than one -- the singular/plural distinction does not suddenly fail us when we move outside the domain of material objects. Determiners such as *many* and *few* reflect differences in amount just as clearly when we are talking about prime numbers as when we are asking for cookies. Why should the semantic force of the count-mass distinction be limited to entities we can see and touch? Most likely, this grammatical contrast is not limited to drawing distinctions within the material domain; it can extend to the most abstract aspects of our mental life.

Chapter 5

Children's representation of the count/mass distinction: A literature review

De gustibus non est disputandum
– Latin Proverb

5.1. Introduction

The argument up to now is that the contrast between count and mass nouns is a semantic distinction, not an arbitrary formal one. Specifically, count nouns have the semantic feature [+INDIVIDUAL] and mass nouns have the feature [-INDIVIDUAL], and this semantic contrast is directly linked to non-linguistic cognition. Experimental evidence suggests that the semantic basis of the count/mass distinction holds not just for obvious examples like *dog* and *water*, but also for superordinates like *money* and *furniture*, and for words describing non-material entities such as sounds and sensations.

In the next chapter, evidence showing that young children also treat this distinction as semantic will be presented. I will argue that the links between semantic features such as [+/-INDIVIDUAL] and other levels of cognitive architecture are innate and will present a theory of how children map these features onto nouns and determiners. The goal of the current chapter is to critically review a set of studies in the developmental literature that are often taken as proof that children do not recognize the semantic nature of the count-mass distinction until quite late in

development. This is obviously incompatible with the nativist semantic view of the count-mass distinction presented here, and thus warrants discussion.

Even accepting the conclusions of the previous chapter, the nativist view is an empirical hypothesis, one that could be mistaken. It need not follow from the fact that adults represent a grammatical contrast as semantic that children must do the same. With regard to the count/mass distinction, there have been two alternatives proposed.

5.2. Non-nativist theories

5.2.1. Semantic change

One possibility is that children do represent the distinction as semantic, but in a different way than adults.[26] Macnamara (1982), for instance, suggests that children initially view the contrast between count nouns and mass nouns as marking the distinction between names for object kinds and names for substance kinds (but see Macnamara, 1986 for a different account). If so, and if the arguments in the previous chapter are correct, then the developmental trend is for count/mass semantics to becomes more and more abstract, eventually encompassing words for non-material entities. As Schlesinger (1981: 230) suggests,

> the child's earliest semantic categories are narrowly circumscribed and ... he extends them gradually into the broader categories of the adult grammar through semantic assimilation.

One could argue, following Fodor (1975), that the whole notion of "semantic assimilation" -- of a category becoming more abstract as a result of experience -- is

---

[26]This does not imply that defenders of this view distinguish between a semantic level and a cognitive level; to my knowledge, only Schlesinger does. I am using "semantic theory" here in a broader sense, to include any theories that posit some role for non-linguistic cognition in syntactic development.

conceptually incoherent, or at best hopelessly vague. However, there are well-documented cases where children's categories *do* change (e.g., the child's concept of "animal", see Carey, 1986), and whatever unknown mechanism responsible for cognitive change in these cases might also cause changes in the child's linguistic categories. Moreover, as I will argue at the end of the next chapter, the domain of count/mass semantics really does expand as a result of cognitive development (though, contrary to Schlesinger, this is best viewed as a change at the cognitive level, not the semantic level).

The objections to this view come from another direction. Evidence from language acquisition suggests that children access the semantic feature [+/-INDIVIDUAL] from the very start, and that this feature plays a role in the acquisition of determiners and nouns. This evidence will be discussed below and in Chapter 6.

5.2.2. Distributional theories

The second alternative is that children initially represent the distinction as semantically arbitrary, something akin to gender marking in an Indo-European language like French, and only later do they come to recognize that count nouns have systematically different meanings than mass nouns. Instead of attending to meaning, children initially categorize words according to their "distributional properties", such as what words they go before, what words they go after, and their absolute position within an utterance (see Maratsos and Chalkley, 1980). As a result of this analysis, children come to cluster words and phrases into appropriate linguistic classes, such as "noun" and "verb", "count noun" and "mass noun".

This was called a "discovery procedure" by linguists in the 1940's, and Fodor (1966) adopted it as a tentative theory in the first modern discussion of how children map words onto linguistic categories. For the count/mass distinction, then, children might initially take the distinction between (for example) the null determiner (which selects for mass nouns) vs. the determiner *a* (which selects for

singular count nouns) as being the same sort as the distinction between *le* and *la* in French, and they would therefore be insensitive to the semantic properties of this grammatical contrast.

As discussed in Chapter 1, however, there are serious problems with distributional theories (see also Pinker, 1979). For one thing, nobody has the foggiest idea about how such a procedure could work, and there have been no explicit theories proposed. Given the immense number of possible alternatives that an unconstrained analysis would have to sift through in order to converge on the correct adult categorizations (over 8 billion for the count/mass distinction - see Gordon, 1982), there is no reason to believe that such a procedure even exists. This problem becomes more vexing when one considers that children younger than 2-and-a-half appear to process syntactic rules and principles (Pinker, 1984; Valian, 1986; see Chapter 2), and that categorization errors are vanishingly rare (Brown, 1973; Maratsos, 1984).

Furthermore, this is considerable evidence that young children understand the semantic basis of contrasts within the nominal system. Katz et al. (1974) and Gelman and Taylor (1984) have shown that young 2-year-olds are sensitive to the semantic contrast between nouns and NPs when inferring the meaning of new words. Brown (1957) has shown that 3- to 5-year-olds understand that count nouns describe different sorts of entities than mass nouns, and Soja (1990) found that children younger than two-and-a-half are sensitive to count/mass syntax when acquiring the meanings of new words. Maratsos (1976) found that 3-year-olds are sensitive to the semantic distinction between definite and indefinite NPs. And Gelman and Markman (1985) and Waxman (1990) have found that young children are sensitive to the semantic differences between nouns and adjectives when determining the meanings of new words. None of this is explained by the distributional view.

Finally, proponents of the view that children's initially view the count/mass distinction as arbitrary have never presented any particular theory of how children

come to know that the distinction is semantic. So there is no debate over whether (say) children just notice one day that the distinction is semantic and switch their grammars accordingly, or whether there is an intermediate point where they have a hybrid representation of count/mass -- part semantic, part arbitrary. This omission is no accident; the view held by many developmentalists who have studied the child's understanding of count/mass is that the *adult* distinction is also arbitrary. The reason for this view should be familiar by now. It is assumed that a semantic theory of count/mass must predict that all objects are count nouns and all substances are mass nouns, but there are mass nouns like *furniture* and count nouns like *drink*, thus the distinction is arbitrary. The previous chapter discussed the problems with this argument, but the point here is that a distributional theory of the acquisition of the count/mass distinction is partially motivated by a distributional theory of the adult linguistic representation.

What is the status of the evidence regarding children's competence? In an article entitled "On the early learning of formal grammatical systems: evidence from studies of the acquisition of gender and countability", Levy (1988a) reviews the work of Peter Gordon and Virginia Gathercole and concludes as follows:

> Thus, Gathercole's conclusions are in complete agreement with the conclusions reached by Gordon (1985); namely, that children first learn the linguistic distinction as a morphosyntactic rather than a semantic distinction. Finally, similar to Karmiloff-Smith's gender data (1979), both Gathercole (1985) and Gordon (1985) found in their older subjects a move away from the priority of morphosyntactic considerations, towards the semantics of the count/mass distinction, which is in exact opposition to the predictions of the "meaning-first" hypothesis (p.186).

Others have reached similar conclusions. Thus Schlesinger (1988: 147), in his discussion of domains where his semantic assimilation theory does not apply, states:

> ... Gordon (1985) and Gathercole (1985) have shown that the count-mass distinction is acquired through formal clues rather than via the semantic object-substance distinction. The reason seems to be that, in English, there is not a very consistent correlation between these two distinctions.

These findings has been taken as strong evidence for the distributional view, Levy (1988a) argues, quoting Karmiloff-Smith (1979), that children view language as a formal puzzle, as "a problem space *per se*". This is obviously opposed to a central claim in this thesis, that much of language development is due to innate links between grammar and semantics, and semantics and cognition. It is necessary to explain why Gordon, Gathercole, and Levy seemingly have come to such different conclusions.

The major problem is that these scholars fail to distinguish between the type of semantic theory advanced by Macnamara and Schlesinger and the sort advanced here. Upon examination, most of the findings pose problems only for the first type of semantic theory. To the extent that the stated conclusions of Gordon and Gathercole do conflict with the claims made by the Semantic Competence Hypothesis, there are methodological problems with their studies that make accepting these conclusions unwise.

5.3. Two theories of semantics and language development

5.3.1. A discontinuous semantic theory (DST)

This is the view that Gordon, Gathercole, and Levy are trying to refute. Gordon calls it "the Semantic Category Hypothesis", and Levy calls it "the meaning-first theory". For the purposes here, I will call it the "discontinuous semantic theory" (DST) to stress the fact that it posits major representational differences between children and adults.

The DST has been defended by scholars such as Bowerman (1973), Braine (1987),

Macnamara (1982), and Schelsinger (1981). It posits that children start off using cognitive categories to acquire language, categories which are rooted in non-linguistic cognition. Since these do not accord with the language spoken by adults, at some point in development children must transform or abandon these categories. With regard to the count/mass distinction, one view is that children start off encoding count nouns as names for objects and mass nouns as names for non-solid substances.[27]

What predictions does this make about the capacities of young children? For one thing, it predicts that children should be incapable of categorizing words that do not name objects as count nouns and words that do not name non-solid substances as mass nouns. Thus children should have difficulty with count nouns like *nap* and *joke* and with mass nouns like *furniture* and *toast*. On the other hand, they should be capable of learning the category of new words solely through perceptual cues -- since they can tell whether something is an object or a non-solid substance solely by looking. Finally, they should be incapable of categorizing new words solely through linguistic input; hearing *a gav* should not tell the children that *gav* is a count noun.

Note that the distributional theory makes the inverse predictions: Children shouldn't find it any harder to learn *furniture* as a mass noun than to learn *water* as a mass noun, but they should be incapable of categorizing words through perceptual information (because they have no links between count/mass syntax and non-linguistic cognition). They should, however, be capable of learning the count-mass status of a new word by hearing it used with selective determiners,

---

[27]It is odd that at one point Levy (1988a: 180) characterizes this view as follows: "[It sees] young toddlers as tied to the here-and-now, possessing cognitive mechanisms that are dependent upon social, physical, and emotional states. It views the toddler as primarily driven by the effort to enhance communication, with an emerging verbal system derived directly from preverbal cognitions, and performing speech acts rather than learning distributional dependencies." It is not clear that anyone actually believes this. Surely, Macnamara (1982) does not; he explicitly rejects the idea that the child has a limited cognitive capacity, and makes no suggestion that the child's desire to communicate is relevant. In fact, Levy seems to have defined the "meaning-first view" as including every bad idea in developmental psychology, a sort of super-strawman.

because for these children, the categories "count noun" and "mass noun" are solely defined in terms of the linguistic contexts that they appear in.

5.3.2. Semantic Competence Hypothesis (SCH)

The claim in this thesis is that the mapping between semantics and cognition allows children to determine the semantic features of new words. For the count/mass distinction, this implies that children can use non-linguistic cognition to mark novel nouns as either [+INDIVIDUAL] or [-INDIVIDUAL]. They can also learn which determiners select for [+INDIVIDUAL] nouns (e.g., *many*) and which select for [-INDIVIDUAL] nouns (e.g., *much*). Having acquired some determiners and nouns in this manner, children can use the semantic features of one word to semantically categorize another. For example, if a child knows that the determiner *a* selects for nouns with the features [+INDIVIDUAL] and [+SINGULAR], then hearing *a gav* should lead her to infer that *gav* has the semantic features [+INDIVIDUAL] and [+SINGULAR], i.e., that it is a singular count noun.

This theory is "continuous", because it posits no difference between the child's semantics representation and the adult's. There is no "semantic stage" that the child uses as a stepping stone to later linguistic success; the semantics of the count/mass distinction is part of the human endowment for language. This is not to say that there are no relevant differences between children and adults. Children think differently than adults and this can affect language development in significant ways (see Chapter 6 for discussion). But the representational structure itself (the semantic feature and its link to cognition) does not change.

If the SCH is correct, then children should have very similar semantic capacities to adults with regard to count/mass. Like adults, they should be able to infer from cognition to semantic category (e.g., if they hear a bounded object called *gav*, they could infer that *gav* is a count noun; if they hear a non-solid substance called *gav*, they could infer that *gav* is a mass noun). They should also be able to infer from semantic category to properties of word meaning (e.g., if they are told *look for a gav*,

they should look for an object; if they are told *look for gav*, they should look for a non-solid substance). Finally, they should be able to make inferences on the basis of semantic features (e.g., if they hear *a gav*, and know the semantic properties of *a*, they should categorize *gav* as a count noun).

5.4. The Experimental evidence

5.4.1. Gordon (1985)

I will defer discussion of Gordon's theoretical claims until later; here I'll limit myself to his conclusion that only older children are capable of categorizing novel words solely on the basis of referential information. If he is correct, then both semantic theories are false. There are three relevant experiments, as follows:

In the first, 40 children aged 3;5 to 5;5 (mean: 4;3) were taught new words with both syntactic and semantic cues.[28] In this experiment, the syntactic cue was the linguistic context in which the new word was taught; whether it was presented following *a* and *another* (count condition) or whether it was presented following *any, some,* and *is* (mass condition). The semantic cue was the nature of the referent, whether it was a novel object (count condition) or a test-tube containing a novel substance (mass condition).

There were two test conditions: *Conflict*, where the two cues differed as to whether they suggested a count or mass interpretation, and *Accord*, where they were consistent. For example, in one of the *Conflict* conditions children would be taught a new word for a novel substance and the new word would be presented in a count linguistic context, while in an *Accord* condition the substance name would be used in a mass linguistic context. The goal of this study was to assess the

[28]This is the terminology used by Gordon and for the sake of convenience I will occasionally adopt it here, without accepting the implicit assumption that "semantics" should only apply to referential information. Translating to the perspective here, he uses "semantic cues" to mean "semantic information conveyed by the ontological status of what the child is exposed to" and "syntactic cues" to mean "semantic information conveyed by determiner choice, pluralization, etc."

relative role of linguistic and referential cues in children's categorization of new words.

After being taught a new word in either the *Conflict* or *Accord* conditions, children were encouraged to produce the word to describe a group of additional objects or substances through the following prompt: ... *So here, we have a/some garn, over there we have more .... what?*. The assumption was that if children use the word in the plural form (*garns*), they have categorized it as a count noun; if they use the word in the singular form (*garn*) they have categorized it as a mass noun. Recall that only count nouns can be pluralized (e.g., *more dogs* vs. **more waters*).

Each child was tested on four *Conflict* trials and four *Accord* trials. In the analyses, Gordon divided the children into a younger group (mean age: 3;9) and an older group (mean age: 4;9). He then categorized the children according to how many times (out of four) that they responded in accord to the syntactic cue vs. the semantic cue. For instance, suppose a child heard a substance described as *a gav*. If the child pluralized *gav*, she would be categorized as giving a syntactic response (SYN); if she responded with the *gav* in the singular form, she would be categorized as giving a semantic response (SEM).

The results from the *Conflict* condition are summarized below (from Gordon's Figure 1). SYN:4 and SYN:3 are the children who followed the syntactic cue for either all of the four trials or for three of the four trials; SEM:4 and SEM:3 are the children who responded to the semantic cue four times or three times, and EQ are the children who responded to each cue for half the trials. Some of the children either pluralized all the time or none of the time; these were excluded from the analysis.

Table 5.1 - Number of subjects by age and response type for Conflict condition

	Younger	Older
SYN:4	4	7
SYN:3	6	5
EQ	3	0
SEM:3	2	2
SEM:4	0	0
All plurals:	1	2
No plurals:	5	5

There is a significant tendency for both the younger children and the older children to attend to the syntactic cues over the semantic cue, though even in the older group, only half of the children did so for all of the four trials. In the *Accord* condition there were much cleaner results, with 37 of the 40 children responding correctly for either three or four trials (out of four) when the syntactic and semantic cues converged onto the same interpretation. Gordon concludes from these results that

> ... the count-mass distinction is *essentially* syntactic but not *exclusively* syntactic ... most children use syntactic cues as a basis for category assignment. However when discreteness of perceptual form conflicts with discreteness in quantification, this leads to a certain amount of inconsistency in responding (p.221).

Gordon's conclusion is unsupported. What he does have is evidence against the

DST (though it's worth pointing out that these children are fairly old even for this; as Gordon himself points out, it is possible that the earlier object/non-solid substance stage has already come and gone). His experiment is designed to contrast linguistic cues and perceptual cues; these are dubbed "syntactic" vs. "semantic", and it is assumed that if the child is biased towards the linguistic cues, then the count/mass distinction is not semantically based. But, as argued before, it is very likely that these are actually *both* semantic cues, just of different types. Both the prescence of the determiner *a* and the fact that noun refers to an object tells the child that a novel noun has the semantic property of heading an NP referring to individuals, i.e., it has the semantic feature [+INDIVIDUAL].

This study suggests that young children have the capacity to attend to both types of cues when categorizing new words. This is evidence against the DST (since it shows that children can attend to linguistic cues), and against a purely distributional theory (since the children are also attending to perceptual cues). But it is only evidence against *all* semantic theories if you make the unwarranted assumption that semantic information cannot be conveyed through linguistic structure. And, at least for adults, this assumption is false.[29]

Why do children favor the linguistic cue over the perceptual cue? (While Gordon did not test adults, my intuition is that the adult response would be the same as the children's). One explanation has to do with the asymmetry between perceptual cues and linguistic cues. The relation between a percept and count-mass semantics is mediated by cognition and as a result there is no one-to-one relationship between percept and grammatical category. Some entities that are perceived as objects are described by mass nouns (e.g., *wood* to describe a block of wood), and some entities that are perceived as non-solid substances are described by count nouns (e.g., *a nice wine* to describe the contents of a glass). There is always

---

[29]Examples are easy to find. For instance, the utterance *John has a wug* justifies the inference that there is at least one wug, and this is an inference about sentence meaning, one based entirely on the presence of *a* (which is the cue in Gordon's "syntactic condition"). Or consider the shift in interpretation in *Mary ate a cake* vs. *Mary ate cake.*

some risk in using perception to infer count/mass semantics. But attending to the semantic features of determiners will never fail; every noun that follows a [+INDIVIDUAL] determiner must be a count noun. In light of this, the children's responses are quite appropriate.

This leads to the issue of how children were construing these novel words taught in the *Conflict* condition: mass nouns describing objects and count nouns describing substances. It is quite possible that they found it impossible to give the word any coherent interpretation and simply chose to focus only on a single cue – most of the time linguistic. The large proportion of children (about one-third) who either always pluralized or never pluralized suggests that children found it such a difficult task that many just gave up on trying to construct a suitable interpretation (only two children were excluded due to response bias in the *Accord* condition). But if children were able to integrate the cues they might do so as follows: For the mass object condition the child could construe the referent as the stuff that the objects are made of. For the count substance condition the child could be construing the referent as the test-tube with the liquid, not the liquid itself, just as someone would call a glass of water *a drink*.

Gordon's first experiment, then, actually supports the SCH, since he shows that children are sensitive to both linguistic cues and perceptual cues, two critical predictions of this hypothesis. The second experiment, however, apparently refutes the semantic theory in all of its versions. Here Gordon tested children's sensitivity to semantic and syntactic cues in isolation. (A third condition, testing the effect of quantity, is irrelevant for our concerns here). There were 39 children tested (mean age: 4;5). In the *Semantics Alone* condition, children were taught a novel word in neutral syntax (*the garn*) as describing either a novel object or a novel non-solid substance. In the *Syntax Alone* condition, children were taught a word with either count syntax (*a garn*) to refer a bean, or with mass syntax (*some garn*) to refer to more than one bean; the idea was that as a food term of a certain size, beans would be perceptually ambiguous as whether or not to be semantically classified as count

or mass (Wierzbicka, 1985; see Chapter 4 for discussion). The test was the same as in the first experiment, whether or not children would pluralize the novel word in an elicitation task.

Gordon found that while all children were sensitive to the syntactic cues, only the older children (4;6 to 5;11) were able to categorize the new word solely on the basis of the semantic cue; the younger children (3;0 to 4;5) could not. Each child had four trials, and their responses were coded as to how many of the four were semantically biased (for the *Semantics Alone* condition) and how many were syntactically biased (for the *Syntax Alone* condition). The relevant results here are those of the children on the *Semantics Alone* condition (from Gordon's Figure 3), shown below:

Table 5.2 -- Number of subjects by age and response type for Semantics Alone condition

	Younger	Older
SEM:4	3	4
SEM:3	2	4
SEM:2	2	0
SEM:1	1	0
SEM:0	1	0
All plurals:	3	6
No plurals:	7	6

Gordon concludes

> The quite bad performance of the younger children with semantic cues is quite surprising, given the enhancing effect of the accord condition of the previous experiment. It would appear that the semantic cues are only truly effective in a parasitic sense of relying on a syntactic base (p.228).

This is an important finding, because Gordon has apparently discovered the following: If young children hear a novel entity named with neutral syntax, they are incapable of using perceptual cues to categorize the new word. If so, this would solidly refute the semantic theory, in all its versions. But it should be noted that other studies *have* found sensitivity to the semantic basis of count/mass syntax in young children (e.g., Brown, 1957; Soja, 1990). And in the next chapter, I report experimental evidence that 3- and 4-year-olds are sensitive to the semantic basis of count/mass syntax even for words denoting non-material entities, as well as evidence from spontaneous speech that children sometimes miscategorize mass nouns that describe objects as count nouns, presumably because of perceptual information. Thus Gordon's failure to find competence on the part of young children is interesting, but cannot be taken to show that that they never attend to non-linguistic cues.

In any case, there are some difficulties with Gordon's interpretation of the data. Gordon claims that the distribution of the younger children's responses on the semantic condition is "quite random." He is correct that the over-all performance is not above chance, but there are two important caveats. First, if children have any response bias -- if they have an overall tendency to pluralize or to not pluralize, for instance, or if they tend to favor either count or mass interpretations of novel words -- this would lead to fewer SEM:3 and SEM:4 responses. The relevant contrast here is not between the children's performance and "perfect performance", it is between the children's responses (i.e., their frequency of pluralization) with the object stimuli vs. their responses with the substance stimuli -- and Gordon does not provide these analyses. As it stands, we can draw no firm conclusions from the

children's poor performance; it *could* be due to insensitivity to semantics, but it could also be due to some sort of general performance difficulty. And second, there is some indication that the children were *not* random. If children were responding at chance, then only 1 out of 16 should give the SEM:4 response -- yet 3 out of 9 children gave this response. A binomial probability test shows this to be significant, $p = 0.016$, suggesting some sort of semantic competence.

Finally, there is Gordon's third experiment. Here Gordon prompted 2-, 3-, and 4-year-old children to use existing English words in a context like the following: Children were told that a Paddington Bear had to go to a food store to buy some food and that on the way he passed by various other stores. They would then be told something like: ... *next he came to a toy store. Do you know what you get in a toy store?*. If the children categorize *toy* as a count noun they should say *toys*; if they categorize it as a mass noun, they should say *toy*. Children were tested on mass superordinates (e.g., *furniture*), count superordinates (e.g., *toy*), mass food terms (e.g., *lettuce*), and count food terms (e.g., *carrots*). In a pre-test, children were shown pictures and asked to point to the picture of the relevant noun; this was done to ensure that these words were familiar to them.

The prediction is that if children categorize words on the basis of perceptual cues they should tend to miscategorize mass nouns describing objects and count nouns describing substances. Although 2-year-olds found this a very difficult task -- over half the data are lost to failures on the pre-test or because of irrelevant responses or silence -- there were very few errors, and there was no tendency for the errors that did occur to be more frequent for such "non-canonical" count nouns or mass nouns.

Gordon concludes from this study that "count/mass categories are not represented in terms of an object/substance distinction", which goes directly against the main prediction of the DST. As noted above, however, the SCH predicts that as soon as children acquire the semantic properties of determiners, they should be able to categorize words like *furniture* correctly. It is therefore not inconsistent with the SCH that children should correctly categorize these words. However since

children are sensitive to perceptual cues and since superordinates are often presented in neutral syntax (e.g., *look at the furniture*) it seems likely that children should occasionally miscategorize such words. And in the following chapter, I present evidence from studies of children's spontaneous speech suggesting that children actually do sometimes miscategorize mass superordinate object names as count nouns. But such errors are rare and it is not surprising that Gordon did not find them.

It is possible, moreover, that Gordon's design underestimates the extent of errors in the child's categorization. Gordon assumes that if children fail to pluralize a noun when describing a plural referent, then they have categorized the noun as mass. But Oatway (1990) has pointed out that children may start off categorizing words like *furniture* and *money* as count nouns with null plural forms. While it is true that all count nouns can follow numbers, determiners like *a* and *another*, and so on, they do not all have overt plural morphemes. Consider words like *fish* and *sheep*. If I point to a goldfish and say *Look at the fish* and then show you several goldfish and ask *What are these?*, a correct answer is *fish* (though *fishes* is also acceptable in some dialects). Nevertheless *fish* is a count noun. Thus failure to pluralize is not a good test for mass-nounhood; Gordon's children could have categorized the mass superordinates as mass nouns, but they could have also categorized them as count nouns taking irregular plurals. This isn't a mere skeptical point; the conditions where children learn the plural form of nouns like *fish* are very similar to the conditions under which children learn superordinates, i.e., they are used in contexts like *Look at the ...* to describe groups. In this regard, Gordon's conclusions should be treated cautiously.

5.4.2. Gordon (1989)

Gordon tested the same hypothesis -- that children would miscategorize non-canonical nouns -- in a study of children's spontaneous speech. He studied the speech of two children, one of them 1;9 to 3;6, the other 2;3 to 3;5. He searched for four types of nouns, as shown below; the examples and subcategories are Gordon's.

Table 5.3 -- Nouns used in search of spontaneous speech

-- Count canonical
(chair, truck, arm)

-- Count non-canonical
abstract (way, ache, story, somersault, bath)
food names (carrot, lambchop)
substance names (drink)

-- Mass canonical
(milk, juice, playdo)

-- Mass non-canonical
abstract (company, room, tennis)
food names (celery, meat)
object names (furniture, silverware)

The prediction, clearly enough, is that if children initially start off with an object/substance distinction, they should miscategorize non-canonical count nouns as mass nouns and non-canonical mass nouns as count nouns. Gordon did not find this; in fact, for count nouns he found the opposite effect -- it was the non-canonical nouns that appeared more frequently with count-selective determiners. Gordon suggests that the non-canonical count nouns are standardly used to refer to "non-particularized" entities, e.g., one says *I want a bath,* but rarely *I want the bath.* Thus these nouns are more likely to appear in count-selective contexts, such as with *a,* rat.. .r than neutral contexts such as with *the.* But in any case, children are clearly not miscategorizing the non-canonical nouns, contrary to the DST. In a second

analysis, Gordon compared the child's knowledge of the obligatory determiner rule -- which is that count nouns cannot appear without a determiner (see Gordon, 1988, and Chapter 6) -- across the two types of count nouns, canonical and non-canonical. Once again, there was no sign that children suffered from a deficit with the non-canonical nouns, providing further evidence against the DST.

There are three points to stress here. First, the force of this study only works against the DST. While it would be evidence *for* the SCH if children have problems with non-canonical nouns (since this would suggest that children are using perception to categorize novel nouns), it is not evidence against the Semantic Competence Hypothesis if they do not have problems. After all, the SCH predicts that children should be able to attend to determiner semantics and thus should be able to correctly categorize nouns like *beans* and *nap* correctly.

Second, the non-canonical nouns are an odd bunch. The abstract nouns denote neither objects nor substances, and thus there is no perceptual bias either way. One might imagine, according to the SCH, that the child would simply attend to linguistic cues in these cases and be error-free. Some of the other nouns are object mass nouns and substance count nouns and it is these cases where one might expect occasional errors according to the SCH, because here cognition-semantics mappings can lead the child astray. In other words, Gordon has collapsed real non-canonical nouns (those which run opposite to the object/substance contrast), with nouns that are neither canonical nor non-canonical -- those for which the object/substance distinction simply does not apply.

Finally, some of the non-canonical nouns do not really seem to be that non-canonical; Gordon gives *carrot* and *lambchop* as examples of non-canonical count nouns, despite the fact that carrots and lambchops are perfectly distinct objects and he gives *meat* as an example of a non-canonical mass noun even though meat is homogeneous and unbounded. These are puzzling decisions and suggest the need for further analyses; these are reported in the following chapter.

5.4.3. Gathercole, (1985)

Gathercole studied children's use of *much* and *many*, as a way to compare (her versions of) a semantic theory and a distributional theory. The design was as follows: A puppet produced sentences where *much* and *many* were followed by nouns belonging to different classes. These nouns appeared in both the singular and the plural form (e.g., *We have much water, We have much waters, We have many water, We have many waters*). 88 children, aged 3;6 to 9;0, were tested. The children gave grammaticality judgments on the sentences, and corrected the sentences that they found ungrammatical. Children were also given a Piagetian conservation task, which did not relate to their performance and will not be discussed further. The word classes are as follows:

Table 5.4 -- Gathercole's word classes

A. Prototypical mass nouns (e.g., *water*)

B. Prototypical count nouns (e.g., *boy*)

C. Non-prototypical mass nouns (e.g., *money*)

D. Non-prototypical count nouns (e.g., *rock*)

E. Flexible nouns (e.g., *cake*)

F. Opaque nouns (e.g., *teeth*)

G. Nominal measure words (e.g., *pile of fish*)

H. Nonsense words (e.g., *smay*)

Gathercole made the following predictions based on her renditions of semantic and distributional theories. The main contrast was between prototypical and non-prototypical nouns; according to Gathercole, the semantic view predicts that children should do more poorly on C and D than on A and B. Also, the semantic view predicts difficulty on flexible nouns (so-called because they can appear as both

count and mass), because they "may cause some confusion (does *cake* refer to the substance or the whole?)". Presumably these should be easy according to a distributional theory, since the nouns could be categorized as both count and mass merely by hearing them used in the appropriate contexts (e.g., by hearing both *cake* and *a cake*). Opaque nouns are those whose phonological shape suggests that they are plural when they are singular and vice-versa. The distributional theory, but not the semantic theory, predicts problems with these words. Nominal measure words (actually word combinations) should be difficult under the semantic theory, because the child might not be able to understand what they refer to, and nonsense words should be a real problem according to the semantic theory, because the child cannot know what they mean. But the distributional theory predicts that children should do fine on both nominal measure words and nonsense words, because each can be classified without appeal to meaning (i.e., the child could attend to whether or not they have a plural morpheme). In sum, according to Gathercole, the semantic theory predicts that children should do better on A and B than on C and D, and should do well on F, and poorly on E, G, and H, while the distributional theory predicts that children should do worse on F than on all the others.

The major finding was that all children did very poorly in all conditions. The youngest children got an average of 43.8% correct, and the oldest group got 73% correct. This was calculated by averaging over all sentences, and thus chance is 50%. Children performed equally well on types A - E, but more poorly on F - H, and there is never a stage where children restrict *much* to non-solid substance names (A) or *many* to object names (B). Finally, the children's earliest accomplishment is realizing that *many* corresponds with plural usage. Gathercole suggests that these findings show that children start off acquiring the count/mass distinction through distributional cues. She concludes that

> Children do not approach the co-occurrence conditions of *much* and *many* with various nouns from a semantic point of view, but rather from a morphosyntactic, or surface distributional one (p.414).

There are several difficulties with Gathercole's conclusions. First, though this might be getting redundant, it is just mistaken to equate "a semantic point of view" with the DST. The finding that children do not encode the count/mass distinction as objects vs. non-solid substances, then, does not force one to a distributional view. Similarly, the early connection between *many* and plurals is not surprising from the semantic theory proposed here, as the plural marker expresses the semantic feature [-SINGULAR], which is selected for by the quantifier *many*.

It is also unclear what to make of children's performance. Even children as old as eight-and-a-half were only rejecting phrases like *much fingers* only 53% of the time, which gives the impression that there is still a considerable way to go before the *much/many* distinction is acquired. Nevertheless, this comparison with adults is misleading; the relevant contrast is not between the child's performance and perfect performance, since such an analysis confounds children's knowledge with their response biases (this is the same point that was made about the second experiment of Gordon (1985)). For instance, if a child hates correcting a puppet, then this will lower her mean number of correct responses, even though she could have perfect understanding that *much* only appears with mass nouns. The relevant contrast, then, is between (e.g,) how frequently children reject phrases like *much fingers* (plural count) vs. how frequently children reject phrases like *much water* (mass), because any response bias would be equal across these two conditions. Here the difference is quite striking: 53% vs. 0%. This suggests that the older children *do* understand the selectional properties of *much*. Since Gathercole presents none of the relevant statistical analyses, it is difficult to evaluate the performance of the younger children, where the differences in percent correct are less striking, and it is impossible to draw clear overall conclusions about their competence.

There were also methodological problems that could depress the children's performance. For one thing, half the sentences with *much* and *many* were presented as declaratives (e.g., *Big Bird saw much smoke*), and these sentences are at best

marginally acceptable. This could make difficult grammaticality judgments even harder. And it could not have helped the children to be asked to make judgments about the grammaticality of sentences with nonsense words, such as *Oscar took many koodges*. In sum, while Gathercole's findings are interesting, they provide no evidence either for or against semantic theories (see Gordon, 1989 for further discussion).

5.5. Semantics without cognition?

Gordon holds a different view from Gathercole and Levy, both who support distributional theories, in that he argues *against* distributional theories and presents instead a version of the SCH. His most recent discussion of this theory is as follows:

> In impugning the validity of the referentially based account embodied in the SCH [Semantic Categories Hypothesis; what is described here as the DST], I am not thereby claiming that categories represent purely formal, non-semantic, distinctions. On the contrary, the count/mass distinction is a semantic distinction ... The count/mass distinction does not serve to pick out names of objects vs. substances. What the distinction does is distinguish between the way in which nouns are quantified. That is, count nouns are quantified as discrete individuals, whereas mass nouns are not individuated when quantified. This difference in the mode of quantification, in turn, affects how we construe the nature of what is being referred to (Gordon, 1989: 10).

With regard to the adult state, Gordon's claim seems exactly right. He argues that the adult distinction is not arbitrary, while at the same time it is not defined in terms of what words objectively refer to. It is a distinction between modes of quantification, between words that are quantified as individuals vs. those that are not. He also notes that the syntactic categorization of the word affects how we conceptualize its referent, which is precisely the argument presented in Chapter 4. In all of these regards, his view is identical to the Semantic Competence Hypothesis.

But Gordon is making another claim, one that motivated the series of experiments reviewed above. He draws a sharp distinction between referential theories and semantic theories. So when he stresses the importance of individuation in the distinction, he takes this as an argument *against* the view that children are sensitive to the object/substance distinction when categorizing new words. He claims, for instance, that

> ... whether or not a noun is individuated depends not on what it denotes, but on whether the language allows it. Therefore, individuation appears essentially to be a matter of linguistic convention rather than semantic entailment (Gordon, 1985: 212).

Like Gathercole, Gordon holds that the referential correlates of the count/mass distinction come in later; they are acquired. So only the older children are sensitive to the fact that objects tend to be described by count nouns and non-solid substances tend to be described by mass nouns. But nevertheless the count-mass distinction is initially semantic, under his view.

This theory suffers from a rather serious problem. According to Gordon, young child go through a stage where they know some words refer to individuals when quantified while others do not. To put it another way, they know some nouns have the feature [+INDIVIDUAL] and others have the feature [-INDIVIDUAL]. Yet they cannot tell whether a noun is count or mass by attending to what it refers to; there is no link between cognition and semantics. So the child could categorize *water* as a count noun and *dog* as a mass noun, just as easily as she could categorize *dog* as count and *water* as mass. Is this consistent?

It doesn't seem so. If anything can be a count noun and anything can be a mass noun, then *what makes the distinction semantic?* To put it another way, what sense does it make to say that children represent one class of words as semantically differently from another if how they think about the meaning of the words has nothing to do with how they are categorized? The only way to hold onto a

semantic theory of count-mass is to assume that the grammatical distinctions interact with non-linguistic cognitive mechanisms.

This drives us towards a strong claim. If the count/mass distinction is an innate semantic distinction then there must be some innate link between the semantic contrast [+/-INDIVIDUAL] and some aspect of non-linguistic cognition. The nature of the semantics-cognition link, and the role it plays in language acquisition, will be discussed in the next chapter.

Chapter 6

Acquisition of the language-specific aspects of the count/mass distinction

6.1. Introduction

Under the theory presented here, the count/mass distinction does not have to be acquired; children are born with it. The semantic feature [+/- INDIVIDUAL] and its links to cognition and to grammar are innate. With regard to cognition, one could tentatively propose that children have a mapping from the semantic feature [+INDIVIDUAL] to the cognitive notion of "discrete bounded entity", where "entity" can be defined over either space or time. With regard to grammar, although there is some variation across languages, children are constrained to map this semantic feature onto aspects of NP syntax (nouns and determiners) in certain prescribed ways, as discussed below.[30]

Although much of the cognitive architecture is unlearned, children still face a considerable task. Assuming that they have already segregated the adult input into determiners and nouns, they must still learn the following:

---

[30]Interestingly, the same feature [+/- INDIVIDUAL] may also show up within the VP system, to express different modes of quantifying over events (i.e., aspectual distinctions, see Bach, 1986; Jackendoff, 1990b). Thus verbs like *slapped* refer to sets of individual events, while verbs like *sleep* do not. While the discussion in this thesis has solely concerned the semantics of nominals and determiners, an important test of the theory presented here would be to see if it could be extended to account for phenomena in the acquisition and representation of verbal aspect.

- The selectional properties of particular English determiners (e.g., that *much* co-occurs with mass nouns and *many* co-occurs with plural count nouns).
- The count-mass status of particular English nouns (e.g., that *dog* is a count noun and *water* is a mass noun).

These tasks are clearly necessary for the child, as the semantic properties of specific English determiners and nouns are not encoded in the human genome. Section 6.2 extends the representational theory presented in Chapter 4 to explain how children might acquire the semantic properties of these words. In Sections 6.3 to 6.5 some specific predictions about acquisition are tested. It is suggested that young children are capable of using perceptual information to categorize new nouns and that they are sensitive to the cognitive implications of semantic features. These claims are defended in three studies. The first shows that young children occasionally miscategorize English mass nouns that describe objects as count nouns, presumably because the children are (incorrectly) using the cognition-semantics mapping to determine the status of these nouns. The second and third studies show that children are sensitive to the cognitive implications of count/mass semantics for both material and non-material entities.

It might appear that learning the properties of determiners and nouns is all that is necessary for the child, but in fact there may be further development that takes place. Cognitive changes may enhance and alter the relationship between semantics and cognition. This is the topic of the final section.

6.2. Assigning semantic features

This is not a theory of how children acquire and represent determiners in general. Given the great complexity of the determiner system (as reviewed in Quirk et al., 1987, for instance), this would have to be the topic for another, much longer, work. Nevertheless, it is impossible to make claims about count-mass acquisition without considering the broader question of how children acquire all semantic contrasts within the nominal system, and this will lead me to make claims about other distinctions, such as number and definiteness. At the same time, there is an important dimension of the acquisition problem that is absent from the discussion below: the precise nature of the algorithm that children use to map semantic features onto nouns and determiners in real-time. Below I will only be concerned with the question of which sources of information children are *capable* of exploiting during language development; for a detailed theory of the actual acquisition process, see Pinker (1984).

6.2.1. The framework

Based on the theory in Chapter 4, one can make the following suggestions regarding the semantic representation of determiners and nouns.

1. All grammatical contrasts that convey cognitive distinctions are expressed in terms of semantic features.
2. These features are either connected with some aspects of non-linguistic cognition or they are defined in terms of other features that are connected with some aspects of non-linguistic cognition.
3. These features constitute the semantic representation of open-class words such as nouns and verbs. But open-class words also have a cognitive representation associated with them; it is this that is commonly thought of as the word's "meaning", "construal", or "conceptual representation".

4. These features constitute the semantic representation of closed-class words such as determiners, prepositions, and modals. These words have no associated cognitive representation above and beyond the links between these semantic features and cognition.
5. Semantic features can be added to a word through innate cognitive-semantics mappings.
6. Semantic features can be added to a word through constrained inferences based on the semantic features of another word.

The first claim is that all grammatical distinctions that express thought must be defined through semantic features. For instance, because one's thoughts about the numerosity of an entity can be conveyed through the grammatical contrast of singular/plural (e.g., *dog/dogs*), there must be a semantic feature corresponding to this numerical contrast. There must be an intermediate level of semantics that serves as a link between thoughts about numerosity and the grammatical contrast, and thus the distinction between singular and plural must be expressible at the semantic level.

This is not to deny that there are semantically arbitrary contrasts within language, such as gender marking in French, but these are solely autonomous properties of grammar, and do not express thought. Also, this does not imply that all semantic features must appear in the grammars of all languages. Languages differ as to which semantic features they exploit; some languages have a grammatical count/mass distinction, some do not; some have singular/plural marking on nouns, others have singular/dual/plural, or singular/dual/trial/plural, or no marking at all; some encode shape-contrasts, others do not, and so on. Like many biological systems, language has a considerable degree of redundancy; a community of speakers does not need to employ all innate semantic (or phonological, or grammatical) distinctions in order to effectively communicate thought.

The second claim is the same made in the critique of Gordon at the end of the

previous chapter; if a feature is not related to thought (either directly or by being defined through other features that are related to thought) then it is not a semantic feature and cannot be used to explain cognitively-relevant systematicities within grammar.

The third claim is that open-class categories have cognitive representations that are independent from their semantic representations. For example, the nouns *dog* and *country* are identical at the semantic level (they each have the semantic features of [+INDIVIDUAL] and [+SINGULAR]), but they differ in other ways, ways that are invisible to grammar but obviously relevant to thought. This was discussed earlier as evidence for the claim that semantic structure and non-linguistic cognition are distinct levels of mental representation.

In contrast, and this is the fourth claim, determiners only contrast on semantic features; there is no other cognitive difference between them. In other words, the only cognitively relevant contrasts between determiners will exist because of differences at the semantic level. There is no deep theoretical reason for claiming this; it just seems to be true. In fact, one could make a more general claim about all functional categories -- determiners, prepositions, inflections, modals, complementizers, and maybe others -- perhaps members of all of these categories can only be cognitively distinguished from one another through differences at the semantic level. So while there are nouns that have identical semantic features but different construals (*dog* vs. *country*), all cognitively-relevant contrasts between functional categories (*many* vs. *much*, *in* vs. *on*, *will* vs. *must*, *that* vs. *which*, etc.) must correspond to distinctions at the semantic level (see Pinker, 1989 for discussion).

What about classifiers; should they be classified with closed-class categories like prepositions and determiners or with open-class categories like nouns and adjectives? In an extensive review, Allan (1977) notes that classifiers encode the same narrow range of semantic notions that show up in other aspects of grammar: MATERIAL, SHAPE, CONSISTENCY, SIZE, LOCATION, ARRANGEMENT, AND NUMBER (but

not color, symmetry, time of day, temperature, emotional state of the speaker, or countless other possible cognitive distinctions). This would suggest that classifiers are defined through semantic features. In contrast, Lakoff (1987) has argued that some classifier systems have a cognitive family-resemblance structure. He gives the example of one Dyibal classifier that co-occurs with nouns depicting women, fire, and dangerous things, and explains this pattern in terms of the social and religious beliefs of Dyibal speakers. If Lakoff is right, then there is non-semantic cognitive content for at least some classifiers, and they are more similar to adjectives than to determiners.

It is unclear, however, whether most Dyibal speakers actually understand (even at an unconscious level) the metaphoric basis of their classifier system. The metaphors may only be active when choosing the appropriate classifier to go with a word that is entering the language, while children might simply acquire the proper classifier by rote (Maratsos, 1989). One possibility is that the only "psychologically real" aspects of classifier systems are those that conform to the semantic primitives found by Allan (1977), Talmy (1985), and others. All the aspects of classifier systems that are not expressible at the semantic level (such as those based on social beliefs) might play no role at all in the acquisition process.

The fifth claim is that at least some links between the semantic level and the cognitive level are innate and that children can exploit these links in the acquisition process. The strongest version of this is that *all* semantic contrasts within the nominal system have innate mappings onto cognition. This is a strong claim, but it is not implausible. Consider the semantic dimensions sub-dividing the determiner system. These would include INDIVIDUATION (*many* vs. *much*), NUMBER (*a* vs. *both* vs. *many*), DEFINITENESS (*a* vs. *the*), EXISTENCE (*no* vs. *a*), PROXIMITY (*this* vs. *that*), and MAGNITUDE (*many* vs. *few*). Some languages draw a contrast between male and female; this is called SEX to distinguish it from arbitrary gender contrasts. Less obviously, consider the contrast between referring to part of a group vs. an entire group: as in *both* vs. *two*. If there are three apples on the table you can say *give me*

two, but not *give me both*; if there are two, you can say *give me both*, but *give me two* is marginal. Similarly, the contrast between *all* and *every* is between referring to the group (*all men*) and referring to all the members of the group (*every man*). (For lack of better terms, call these contrasts COMPLETENESS and GROUPING respectively). Finally, there is POLARITY (Klima, 1967); some determiners have negative polarity and cannot appear in non-negated declarative sentences; consider *There weren't any apples* vs. **There were any apples.*

Do all of these semantic contrasts have innately specified counter-parts in children's cognition? Obviously, nobody knows -- but the evidence is suggestive. Infants are sensitive to individual objects (Spelke, 1988) and sounds (Starkey et al., 1982), both perhaps linked to INDIVIDUATION. They can distinguish different numbers of entities (Strauss and Curtiss, 1981) (NUMBER) and are sensitive to the contrast between male and female faces, (Fagan and Singer, 1979) (SEX). There is evidence that 2-year-olds grasp the concept of "exhausting a set", e.g., pointing to each member once and only once) (Potter and Levy, 1968) (COMPLETENESS, GROUPING). There is a growing base of data suggesting that young children have a sophisticated understanding of belief-desire psychology (Wellman, 1990) and DEFINITENESS may be rooted in the understanding of what others are and are not aware of. Further, it is suggestive that the youngest children studied (3-year-olds) are essentially adult-like in their use of the English determiners *a* and *the* (Maratsos, 1972). As for EXISTENCE, PROXIMITY, and MAGNITUDE, it would be surprising to find any complex organism that was insensitive to the cognitive correlates of these semantic properties. The one case where it is very hard to find a cognitive correlate is POLARITY, and here it has been argued (see Bach, 1989) that POLARITY is not itself a semantic primitive but is instead defined in terms of other aspects of determiner semantics.

All of these claims about infants and young children are highly controversial. Furthermore, there is a large gap between showing that the cognitive correlate to a semantic feature is innate and showing that there is also an innate link between the

cognitive correlate and the semantic feature. Thus even if we could be sure that children innately distinguish numbers, it does not necessarily follow that this cognitive capacity is innately linked to the semantic feature [+/-SINGULAR]. Nevertheless, it is reasonable to take the studies reviewed above as providing tentative support for the view that cognitive correlates of the semantic features encoded in the determiner system are unlearned. In fact, there is one specific case where the notion of innate cognition-semantics links makes some surprisingly accurate predictions about the relation between semantics, cognition, and grammar. This is the domain of number.

Number is encoded in grammar, but only to a limited extent. In English, there is the presence or absence of plural marking on nouns, which can be described as marking the contrast between one and more than one. Also some determiners, like *a*, select for a singular referent while others, like *many*, select for a plural referent. Although English is sometimes assumed to only have a singular/plural distinction, there are certain aspects of English grammar that mark twoness. The determiner *both* selects for nouns describing two entities. And there is a class of nouns -- including *scissors, pants, tweezers,* and *trousers* -- which is defined in terms of having two identical parts conspiring to fulfill a single function (McCawley, 1985; Wierzbicka, 1985). While English has no grammatical means for expressing threeness, other languages do, both in the noun system (Greenberg, 1978) and the pronoun system (Ingram, 1978). No language grammatically encodes fourness. In sum, the notions of one, two, three, and more than three correspond to primitives at the semantic level and it is these notions that are capable of being expressed in grammar.

It was suggested above that (i) distinctions within the determiner system are expressed as differences in semantic features, and that these features have innate mappings onto cognitive primitives (claims 1, 2 and 5), and (ii) some of the semantic features that define determiners are also encoded in the semantic representation of nouns (claims 3 and 4).. This leads to two predictions. First, there

should be innate mappings from the semantic notions of one, two, three, and more than three onto cognitive primitives. But there is no reason to expect mappings for specific higher numbers, as these are not innately encoded at the semantic level. And second, the words *one, two,* and *three* should have a special status relative to the number words *four* and above. Number words smaller than *four* are defined in terms of innate semantic primitives and are presumably acquired the same way as determiners like *a* and *the,* while all the rest are acquired through some other mode of mental representation, perhaps as the result of inductive processes at the cognitive level (see Wynn, 1990b for discussion).

Both predictions hold. Infants can discriminate small numerosities, up to the distinction between three and bigger than three, but no more (e.g., Antell and Keating, 1983; Starkey and Cooper, 1980; Strauss and Curtis, 1981). And 3-year-old children can subitize (precisely quantify without conscious counting) numerosities up to two or three but no more (Chi and Klahr, 1975; Schaeffer, Eggleston and Scott, 1974; Silverman and Rose, 1980). Finally, in a recent study, Wynn (1990a) discovered the following pattern in lexical acquisition: children first acquire the meaning of *one,* then *two,* then *three.* Once they have acquired *three,* the next step is acquiring all the rest of the numbers within the counting sequence -- but all at once, not in sequence. This suggests a strong dissociation between the words *one, two, three,* and all the rest. Both effects follow from positing links between semantics and cognition and between grammatical features and lexical representation.

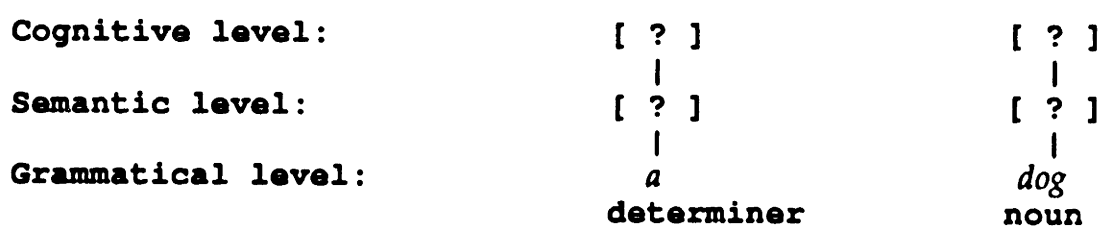
The final claim is that children should be able to infer that a word has certain semantic features because of the presence of these features on another word (see also the discussion in Chapter 3). For example, assume that *a* has the semantic features [+INDIVIDUAL] and [+SINGULAR]. The child who knows this should be able to infer from *a dog* that *dog* also has the features [+INDIVIDUAL] and [+SINGULAR]. And vice-versa; the child could infer the properties of the determiner from the properties of the noun. But there is an asymmetry here; every feature on a noun can transfer to a determiner, but not every feature on a determiner can transfer to a

noun. This is because determiners can also express semantic notions that are not inherent in the representation of the noun, such as definiteness. For example, *the* has the property of forming an NP that picks out a definite referent but there is no such thing as a definite noun. In general, determiners tend to inherit features from nouns, and not the other way around (see Greenberg, 1966).

6.2.2. Methods of assignment

For all cases below, assume that the child has somehow determined that *a* is a determiner and *dog* is a noun, but knows nothing of the semantic features of these words or their cognitive representation. We can express this as in (1).

(1)

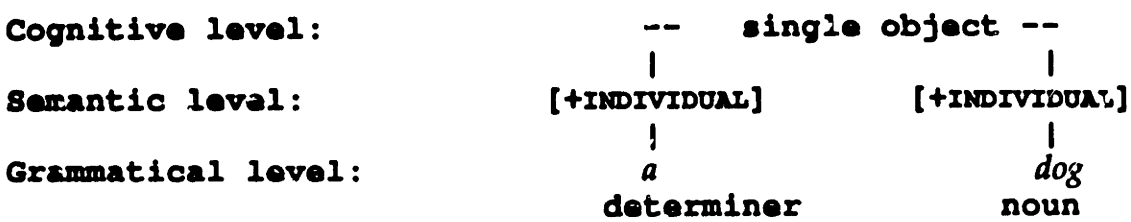


There are four possible ways that semantic features can be assigned to these words.

1. Assignment from cognition to noun and determiner

The simplest case is through cognition-semantics links. Thus the child could learn that *dog* and *a* each have the feature [+INDIVIDUAL] by seeing a dog, construing it as a single individual, and associating it with the NP *a dog*. This is shown in (2):

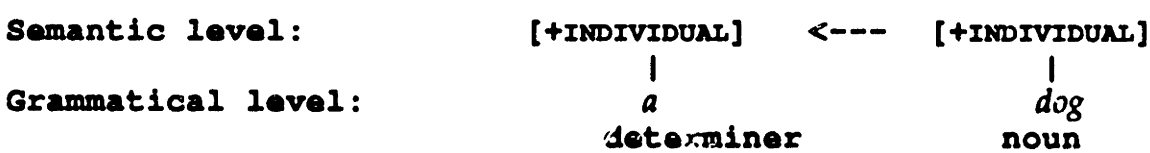
(2)



2. Assignment from noun semantics to determiner semantics

The second case is when the child already knows the semantic feature of the noun, perhaps because she has heard the noun used with another determiner. But there is no referent available to determine the semantic features of the determiner through cognition-semantics mappings. Thus the child could learn the semantic representation of *a* through hearing *a dog*, as in (3).

(3)



This puts us in a position to solve Gordon's puzzle (1988), which is how children learn -- without negative evidence -- that bare singular count nouns cannot appear as NPs (e.g., **dog barked*). Gordon's own solution was that children observe that determinerless nouns have generic reference and that mass nouns and plural count nouns fulfill this role for individuated and non-individuated entities. Thus a bare singular count noun like *dog* would have to form an NP with the identical meaning of *dogs* (as there's nothing else that it could mean) and this overlap in meaning is prohibited by some version of a uniqueness principle. One problem with this, as noted in Chapter 4, is that it forces a dissociation between the absence of singular

count nouns in bare NPs and the absence of singular count nouns following determiners like *more* and *a lot of*. If there is some connection between the two phenomena, Gordon's theory cannot capture it. Another complaint – which is a general worry with explanations based on pre-emption of meaning – is that it doesn't explain why children do not hypothesize that bare singular count noun NPs differ from bare plural NPs on some *different* semantic dimension, such as magnitude or polarity.

Here's an alternative: Suppose that it is a condition at the semantic level that all nouns have a determiner associated with them; the determiner maps the noun meaning onto the NP. This assumption was proposed in Chapter 3 as part of a general theory of nominal semantics, and is held within other semantic frameworks as well (e.g., Barwise and Cooper, 1981; Higginbotham, 1983). There are two immediate advantages of this proposal. First, it allows us to simplify the semantics considerably: by assuming that all nouns need a determiner to become NPs, one need not posit a basic division among noun classes where some need determiners to establish reference and others do not. And second, it avoids positing a strong semantic discontinuity with other languages, such as French, where mass nouns and plural count nouns must have an overt determiner. Under this analysis, all languages form NPs by combining determiners and nouns; the only cross-linguistic difference is relatively superficial – in some languages a given determiner is phonologically overt; in others, it isn't.

Thus, when children hear *dogs bark* they will mark the null determiner as [+INDIVIDUAL, -SINGULAR] and when they hear *water is wet* they will mark the null determiner as [-INDIVIDUAL, +SINGULAR], because all determiners, whether or not they are phonologically overt, inherit the semantic features of the nouns they appear with. As a result, the null determiner subcategorizes for plural count nouns and mass nouns, but not for singular count nouns. This is an example of conservative learning; children use the null determiner preceding only nouns with the features [+INDIVIDUAL, -SINGULAR] and [-INDIVIDUAL, +SINGULAR] because

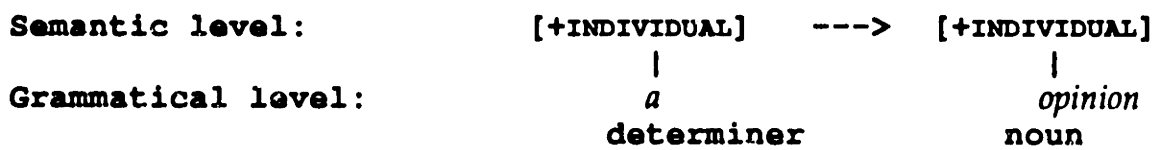
these are the only contexts where that they have heard adults use it.[31]

Gordon quite reasonably rejects a different form of conservatism, which is that children assume "that if a fairly simple construction has never been heard after some period of time, then it is not generated by the grammar". Besides the fact that it is vague, Gordon notes that children could produce an indeterminate number of hypotheses of this sort, mostly bizarre -- "do children test the hypothesis .. that mass nouns do not occur with relative clauses? It seems doubtful." (1988: 123). But the form of conservatism presented above is limited to a well-defined and highly restricted domain, and avoids this problem of vagueness and bizarre hypotheses.

3. Assignment from determiner semantics to noun semantics

The third case is where the child could assign a semantic feature to the noun through a feature on the determiner. This is shown in (4).

(4)



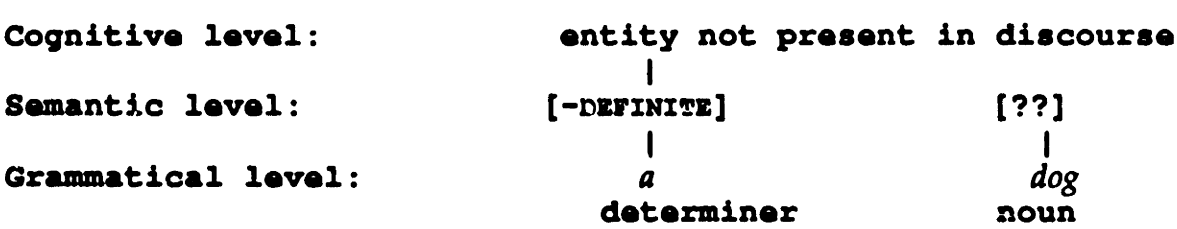
4. Assignment of semantic features to a determiner through cognition

Finally, some of the semantic features on determiners cannot be inferred from nouns, since nouns do not encode them. They must be inferred from properties of the situation. (Pinker, 1984: 198-201 describes these features as "context-specific" to distinguish them from "stem-specific" features such as INDIVIDUATION). This is shown in (5).

---

[31]Alternatively, as suggested in Chapter 4, mass nouns and plural count nouns might share the feature [+EXTENT], and children would thus learn that the null determiner subcategorizes for [+EXTENT].

(5)



6.2.3. Two predictions

The section above makes certain predictions about children's competence. One is that children should be able to assign semantic features to nouns on the basis of determiner semantics, and vice-versa (mappings 2 and 3). Further, they should be able to infer semantic features on the basis of cognition-semantics links (mappings 1 and 4). While it is not in doubt that children can use determiner-type to infer count/mass categorization (Gordon, 1985, 1988), it is frequently claimed that they cannot determine the count-mass categorization of a word through their construal of the word's meaning (e.g., Gathercole, 1986; Gordon, 1985; Levy, 1988a). Study 1 presents evidence from children's spontaneous speech that suggests otherwise.

Another prediction is that children should be able to draw cognitive implications on the basis of the semantic features of the determiner used. This follows from the claims that determiners are defined through semantic features and that there exist innate links between semantics and cognition. Studies 2 and 3 address this issue, considering the role of determiner semantics in effecting inferences about material and non-material entities.

6.3. Study 1 -- Children's count-mass errors

What would we expect if children can use cognition to determine the semantic categorization of novel nouns? One prediction concerns how children categorize mass nouns referring to objects, such as superordinates like *furniture* and *money* and food terms like *bread* and *lettuce*. Recall that these are the sorts of words that vary cross-linguistically: in some languages they are count, in others they are mass. Children often hear these words in neutral syntax, in contexts like *look at the furniture*, where *furniture* could be either count or mass, since *the* selects for both [+INDIVIDUAL] nouns and [-INDIVIDUAL] nouns. If children are capable of using perceptual cues, some might mistakenly categorize such words as count nouns. That is, they can use the mapping from cognition to semantics to assign to these words the feature [+INDIVIDUAL]. After all, children have to categorize these words one way or another (either [+INDIVIDUAL] or [-INDIVIDUAL]) and the bounded object interpretation is cognitively accessible as a construal of their meanings.[32]

Thus the prediction is that children should sometimes miscategorize words like *furniture* and *bread* as count nouns. I studied the spontaneous speech transcripts recorned on the CHILDES computer database system (MacWhinney and Snow, 1985) to see if such errors spontaneously occurred. As a control, I looked at the children's use of the words *juice, milk, mud, sand,* and *water*. Since the referents of these words cannot readily be construed as bounded individuals, the theory above predicts that children should not have problems with them, because even without the aid of the appropriate determiners, the cognition-semantics mapping would lead these words to be categorized (correctly) as [-INDIVIDUAL].

---

[32]Similarly, we might expect children to occasionally miscategorize certain count nouns as mass nouns. These cases would be very hard to find in spontaneous speech, however. Children frequently omit determiners in their spontaneous speech (e.g., *I want cookie*), and it would be difficult to distinguish cases where children have actually categorized a word as a mass noun and were using a null determiner, from cases where they were merely omitting the appropriate determiner as the result of performance limitations in language production (see Bloom, 1990b for discussion). Because of this, children's miscategorizations of count nouns were not searched for.

6.3.1. Subjects

The speech of five children were studied. Their ages ranged from 1;6 to 5;2 (see Table 6.1).

Table 6.1 - Transcripts analyzed

Child	Age Range	Source
Abe	2;4 - 5;0	(Kuczaj, 1976)
Adam	2;3 - 5;2	(Brown, 1973)
Eve	1;6 - 2;3	(Brown, 1973)
Sarah	2;3 - 5;1	(Brown, 1973)
Peter	1;9 - 3;1	(Bloom, 1970)

6.3.2. Materials and Procedure

I searched for three classes of words: mass nouns that describe objects, mass nouns that describe substances, and the word *people*. The mass object list contains 14 words, all which denote entities that can be construed as discrete objects: *bacon, bread, cheese, celery, clothing, furniture, fruit, jewelry, lettuce, mail, money, paper, spaghetti* and *toast*. The mass substance words searched for were *juice, milk, mud, sand,* and *water,* all which describe unbounded substances or substance-like collections of particles.

People was searched for because although it is a count noun (*many people* vs. **much people*), it cannot be used to describe an individual; *a people* and *another people* are unacceptable. Since the word is relatively frequent in child language, it is interesting to see whether children will notice that it refers to groups of individuals and thus encode it, incorrectly, as being able to act as a singular count noun and refer to a single individual. This would suggest that children are not merely accumulating semantic features but that they also apply semantic rules, one of which can convert a noun which is [+INDIVIDUAL,-SINGULAR] to [+INDIVIDUAL,+SINGULAR]. If they apply such a rule to *people,* this would lead them to make errors like *a people.*

For this analysis, I looked through all utterances with these words (extracted by computer search) and listed all the cases where they were used in a manner acceptable for count nouns and unacceptable for mass nouns. Such usages would occur when the nouns would co-occur with the determiners *a, another, both, one,* or with numerals. *People,* being plural, is also unacceptable with singular verb agreement.

One potential problem with this analysis is that children make many sounds similar to *a*. Thus a child who says *piece a cake* is probably not encoding *cake* as a count noun, but rather means to say *piece of cake*. Similarly, a child who says *I fell a water* probably means to say *I fell on/in water*. This was particularly frequent in the earlier transcripts of Eve. Because of this, I omitted any example where the determiner *a* could be replaced with *on, of,* or *in* to form a meaningful utterance.

6.3.3. Results and Discussion

There were 38 cases where the object mass nouns were used as count nouns (see Table 6.2), 12 cases where the substance mass nouns were used as count nouns (see Table 6.3), and 20 cases where *people* was used with singular count syntax (see Table 6.4).

Table 6.2 -- Errors with object mass nouns

Abe
I got a paper (3;1)
What is that? Just a paper (3;5)
I'm gonta have a lettuce (3;6)

Adam
dat a mail (2;7)
want a paper (2;7)
I don't like a paper (3;2)
this is a paper (3;3)
look at me pack ten money (3;4)
I could use both of my money (4;0)
I have a new money (4;0)
where's a paper (4;5)
I want a paper (4;7)

Eve
a paper (1;7)
I drop a celery (1;7)
a cheese (1;7)
a nice cheese (1;7)
Mommy fix a paper (1;9)
that a new paper (1;11)
eating a bacon (2;2)
a bread with some butter on it (2;3)

Sarah
a money (2;3)
I see write on a paper (3;0)
I get a paper (3;3)
I write another paper (4;3)
they have a spooky furniture (4;6)
give me another paper (4;8)

Peter
I find a paper (2;2)
It's a paper (2;3)
there's a paper (2;3)
this pencil on this a paper (2;3)
I'm gonta get a paper (2;3)

that's a paper (2;4)
that's a money (2;5)
a money (2;5)
there's a money in here (2;5)
you get another paper (2;5)
a bacon (2;7)
I want a paper (2;9)

Table 6.3 -- Errors with substance mass nouns

Abe
this not a water this water (2;5)
this is not water this a cave! this is a water! (2;5)
this a water that's a water and that's a cave (2;5)
which water? not a water (3;2)

Adam
like another water (2;10)

Eve
drink a juice (1;7)

Sarah
a water (2;11)
I want drink a water (3;2)
that's how a water is (4;3)

Peter
have a juice (2;2)
where's a juice (2;2)

Table 6.4 -- Errors with *people*

Abe
you be a people (2;11)
you're a people (3;1)
he could bite a people (3;1)
is that a people? (3;8)

Adam
and another people on that seat (3;4)
he blind people (3;7)

Sarah
you come some another people (4;5)
a people (4;11)

Peter
is he a people? (2;8)
this people [=to one individual] (2;8)
there's one people in there (2;9)
where a people (2;9)
that one's small people, that's one's a bigger people (2;10)
one people (2;10)
he's a people (2;10)
one people (2;10)
look a people (2;10)
one little people (3;1)

The first thing to note is that children clearly make mistakes with words like *furniture* and *people*. Utterances like *they have a spooky furniture, I have a new money* and *one little people* are strikingly bad. Further, the errors appear to be randomly scattered across the sample, rather than appearing only with very young children. While the majority of errors involved *a*, other errors occurred: 8 of the object mass noun errors (21%) and 5 of the *people* errors (25%) were due to other count-selective properties of grammar.

A large proportion of the object mass noun errors (55%) were with *paper*. While

this may be due to the children miscategorizing the word *paper* as the result of cognition-semantics links, there is an alternative: For adults, *paper* can also mean "newspaper", and children might have heard this usage of the word with count noun syntax and over-extended it to the "piece of paper" meaning. This would be a case of miscategorization through linguistic information, not cognitive information.

There are two reasons to doubt this alternative, however. First, a newspaper is perceptually different from the pieces of paper that children write on (it is much bigger, for one thing) and it is not obvious that children would infer that the "newspaper" meaning of paper could be extended to the "piece of paper" meaning. A second consideration is that while a newspaper is often called *the paper* this usage would not cue the child that *paper* was a count noun. An exclusively count noun context such as *a paper* or *another paper* would be necessary. While these usages are not non-existent in adult speech, my intuition is that they are relatively rare -- and all of the five children produced utterances like *a paper*. Nevertheless, the status of these errors is an open question.

What can we make of the fact that children also made errors with substance mass nouns? There were only three incorrect phrases: *a water*, *a juice*, and *another water* (which is the only usage that is not due to over-extension of *a*: (8%)). The utterances *a water* and *a juice* are acceptable under the rule of construal that switches a mass noun to a count noun if the mode of individuating over the substance is contextually available (i.e., "restaurant talk"; Langacker, 1987, see Chapter 4). Five of the count usages were clearly of this form. It is also worth noting that three of Abe's utterances were from a single session where Abe, quite excitedly, was using *water* as both a count noun and a mass noun. In sum, the evidence from Table 6.3 does not provide evidence that these children had actually categorized any of these mass nouns as count nouns, as opposed to either making speech errors (Abe) or using productive rules of category transfer.

Another difference between the two classes of words involved frequency of errors. While both sorts are rare, children tend to miscategorize object mass nouns

more frequently than substance mass nouns. The proportions for each child are shown in Table 6.5.

Table 6.5 -- Proportion of errors

Child	Object words	Substance words
Abe	0.8%	2.9%
Adam	1.8%	0.2%
Eve	2.6%	0.3%
Sarah	3.8%	1.1%
Peter	4.5%	0.7%

Abe made more errors on substance words than object words, though the difference is not significant ($\chi^2(1) = 0.34$, $p > .5$). The four other children made significantly more errors for the object words than for the substance words (Adam: $\chi^2(1) = 5.56$ Eve: $\chi^2(1) = 5.18$, Sarah: $\chi^2(1) = 3.94$, Peter: $\chi^2(1) = 6.82$, all p's < 0.05, two-tailed). The most reasonable way to explain the difference between these two classes of words is in terms of differences in meaning. In sum, the kinds of errors found suggest that children do categorize some words as [+INDIVIDUAL] through cognition-semantics mappings, as predicted by the theory here.

Finally, *people* was frequently used as a count noun, suggesting that children do over-apply a rule transforming a plural count noun (which is how adults use *people*) to a singular count noun (which is how all of the children except for Eve used it). The *people* errors occur significantly later than the object mass noun errors (mean age for *people* errors: 3;2, mean age for object mass noun errors: 2;9, $F(1,53) = 4.64$, $p = 0.03$, two-tailed). This suggests that the productive rule may be coined only by slightly older children.

6.4. Study 2 -- 4-year-olds' sensitivity to count-mass syntax in material and non-material domains.

The goal of this study was to see if children are sensitive to the relation between the semantic encoding of a determiner and its cognitive construal. In particular, it tested whether children would make different inferences on the basis of whether a word was presented with [+INDIVIDUAL] determiners vs. [-INDIVIDUAL] determiners. This was done by teaching children a new word as either a count or mass noun and then seeing if they would respond differently when this word is used as part of a request -- either to give something to a puppet, or to perform an action.

This experiment was motivated by a classic study by Brown (1957), which was done to test a rather similar hypothesis, that: "the part-of-speech membership of [a] new word could operate as a filter selecting for attention probably relevant features of the nonlinguistic world."[33] Brown showed 3- to 5-year-old children pictures that depicted (among other things) a substance and an object. He told them either *In this picture you can see a sib. Now show me another picture of a sib* (count condition) or *In this picture you can see sib. Now show me sib* (mass condition). When the word was a count noun, 11 out of 16 children pointed to the object, but when the word was a mass noun, 12 out of 16 children pointed to the substance. Brown concluded that "young English-speaking children take the part-of-speech membership of a new word as a clue to the meaning of a new word.", suggesting an early sensitivity to the relationship between grammar and thought.

In her discussion of this, Gathercole (1986) makes the astute point that Brown did not actually show that children were sensitive to the semantic properties of the noun, since it is possible that they did not attend to the noun at all. What he did

---

[33]The main difference between my view and Brown's is that he posits that linguistic categories are merely correlated with different non-linguistic notions and "the native speaker detects this tendency while he is the process of learning the language". My position is that there is an intermediate level of semantics between grammatical categories and non-linguistic thought, and that the semantics-cognition links are innate and non-probabilistic.

show was that the children understood the semantic import of the *determiners* -- that *a* refers to an individual and the null determiner does not. Gathercole's point is a good one, and the experiment that follows should not be taken as testing the hypothesis that children can construe the meaning of nouns differently as a function of determiner use. Instead, it should be taken as testing the hypothesis that they understand the cognitive import of determiner semantics.

There are two major differences between Brown's study and the one below. First, Brown studied whether children could focus on different entities (an object vs. a substance) on the basis of linguistic information. This study tests a stronger claim; that determiner semantics can affect how children construe a single perceptually ambiguous entity. And second, it studies the effects of count/mass semantics not only for the construal of a material entity (particles of food), but also for the construal of a non-material entity (sounds). If children are sensitive to this, it would suggest that the cognitive correlate of the count/mass distinction is not merely that of object vs. non-solid substance, but something considerably more abstract.

6.4.1. Subjects

16 4-year-olds were tested, aged 3;10 to 5;1 (mean age: 4;6). All were from day care centers in the Boston area.

6.4.2. Materials

For the *food*-condition the stimuli were either lentils or blue pieces of broken raw spaghetti, which are the sorts of entities that could be named with either count nouns or with mass nouns. For the *sound* condition, the stimuli were strings of sounds -- a bell was struck with a baton at a rate of two strikes per second; this was recorded on a tape-recorder. There were two bells used, of different pitches, each recorded on a separate tape. This is the sort of sound that could be construed as a set of individuals or as an undifferentiated stream.

The bells themselves were shown to the children -- there was a circle of multi-colored bells, with the two relevant bells (a red one and a green one) on opposite ends. The children were also shown the baton.

6.4.3. Procedure

Pre-test

Children were introduced to a hand-puppet, which looked like Queen Elizabeth and was holding a cup. The children were told:

> This is the Queen Puppet, and in this game she is going to ask us to do certain things for her. They aren't going to be that hard to do, but you have to listen very carefully. The Queen Puppet is very picky, so you have to do *exactly* what she says.

Then they were shown a pile of five pennies, and the puppet asked the children *Please give me one of the pennies. Put one of the pennies in my cup, please.* Then the children were told *Please give me all the pennies. Put all the pennies in my cup, please.* The purpose of this was to get the children used to putting different numbers of entities in the puppet's cup. None of the children had any problem with this task.

The same procedure was done to familiarize the children with the process of making different numbers of sounds. So children were handed the baton and shown the bells. Then the Queen Puppet said:

> Do you know you can use this stick to make sounds? You can make sounds by hitting it against this thing. Do you want to play with it?

The children would almost always bang away on the bell. They were then told by the Queen Puppet: *Please make just one sound. Make just one sound please.* and then *Please make a lot of sounds. Make a lot of sounds, please.* Once again, the children found this is a fairly easy task, though some of the children tended to bang away repeatedly at the bell, and had to be discouraged from doing this.

Experiment

There were two main conditions: *food* and *sound*. Within each condition, there were two types of stimuli; for food, it was the lentils and the spaghetti, for sound, it was two different bells of different pitches. The nonsense words *sib* and *gav* were assigned to the object conditions, and the words *moop* and *fep* were assigned to the sound conditions. The critical variable was word syntax: whether the word was taught as a count noun or a mass noun.

Each child was given four trials: *Food-Count, Food-Mass, Sound-Count,* and *Sound-Mass*. There were four different orders, with the caveat that the two Food conditions were taught together and the two Sound conditions were taught together. The assignment of nonsense words and stimulus type were counter-balanced. The stories were as follows:

Food-Count

EXPERIMENTER: Do you know what these are called? [pointing to a plate of lentils/spaghetti] These are gavs. I first put one gav here and then another gav and then another gav until I had a lot of gavs here. There really are a lot of gavs here. What are there here? (if child is silent, he or she is prompted to say *gavs*). These are gavs ... Now listen to the Queen Puppet.

QUEEN PUPPET: Please give me a gav. Put a gav in my cup please.

Food-Mass

EXPERIMENTER: Do you know what this is called? [pointing to a plate of lentils/spaghetti] This is gav. I first put some gav here and then some more gav and then some more gav until I had a lot of gav here. There really is a lot of gav here. What is there here? (if child is silent, he or she is prompted to say *gav*). This is gav ... Now listen to the Queen Puppet.

QUEEN PUPPET: Please give me gav. Put gav in my cup please.

Sound-Count

EXPERIMENTER: If you hit this over and over [pointing to the bell], you can make a lot of moops. I made a tape recording of a lot of moops [turning on the recorder]. First, I made one moop and then another moop and then another moop until I had a lot of moops. There really are a lot of moops coming out of the machine. What's coming out of the machine? (if child is silent, he or she is prompted to say *moops*). These are moops ... Now listen to the Queen Puppet.

QUEEN PUPPET: Please make a moop for me. Use this stick to make a moop please.

Sound-Mass

EXPERIMENTER: If you hit this over and over [pointing to the bell], you can make a lot of moop. I made a tape recording of a lot of moop [turning on the recorder]. First, I made some moop and then some more moop and then some more moop until I had a lot of moop. There really is a lot of moop coming out of the machine. What's coming out of the machine? (if child is silent, he or she is prompted to say *moop*). This is moop ... Now listen to the Queen Puppet.

QUEEN PUPPET: Please make moop for me. Use this stick to make moop please.

Thus the count noun stories had used the novel word following *one, a another,* with plural morphology, and with plural agreement. The mass noun stories had the novel word following *some* and *a lot of* (without pluralization) and as a bare NP. The difference in the test questions was whether the word appeared with *a* or with no overt determiner at all.

6.4.4. Results

The prediction is that children should respond differently to the count noun stories than in the mass noun stories -- they should give one item or make one sound when the word is presented with count syntax and give many items or make many sounds when the word is presented with mass syntax. (Strictly speaking, giving one item or making one sound when exposed to the mass noun is not incorrect; if children do this, however, it would diminish the role of word syntax

and would thus go against the predicted direction of the effect). Separate ANOVAs were performed for the food and sound conditions, with stimuli assignment, question order, nonsense word assignment, and word syntax as the independent variables and the child's response ("0" if she gave one item or made one sound, "1" if she gave more than one item or made more than one sound) as the dependent variable.

There were no effects of stimuli assignment, question order, or nonsense word and assignment for either the food or sound analyses, nor were any of the interactions significant. There was a significant effect for word syntax for the object condition ($F(1,8) = 18.000$, $p < 0.005$, one-tailed) and a marginally significant effect of word syntax for the sound condition ($F(1,8) = 3.000$, $p = 0.06$, one-tailed). This suggests that word syntax affected the children's response in both conditions (see Table 6.6 for the breakdown of responses).

Table 6.6 -- Children's responses on all questions as a function of syntax

Object

	Count	**Mass**
one	13	9
more than one	3	7

Sound

	Count	**Mass**
one	6	2
more than one	10	14

There was a significant effect of order for the object condition ($F(3,8) = 17.333$, $p < 0.005$), though not for the sound condition. Observations during testing suggested that children found the first trial of a given condition (either sound or food) the easiest, and would sometimes act impulsively or confused on the second trial. In

light of this, a second analysis was done, looking only at children's first trial on a given condition. The results are shown in Table 6.7. When only the first trial was studied, the effect is somewhat stronger [Objects: $p = 0.00008$, Sounds: $p = 0.04$, both one-tailed Fisher Exact Tests].

Table 6.7 -- Children's responses on initial question as a function of syntax

Object

	Count	**Mass**
one	8	0
more than one	0	8

Sound

	Count	**Mass**
one	4	0
more than one	4	8

The children's difficulty with the sound condition is clear from both the data and from observations of the children. There is a tendency for children to repeatedly hit the bell; it was obvious during the testing that children enjoyed doing this a great deal and this might have led some of them to ignore the syntax. Nevertheless, for both conditions, the 4-year-olds responded to the word syntax to a significant degree. That is, they tended to give one item or make one sound when exposed to count syntax and to give more than one item or make more than one sound when exposed to mass syntax. This suggests some sensitivity to the semantic basis of the count/mass distinction.

6.5. Study 3 -- 3-year-olds' sensitivity to count-mass syntax in material and non-material domains.

Study 3 looks for the same knowledge in 3-year-olds. Some changes in the procedure were done to facilitate performance and to reduce the child's tendency to bang on the bell in the sound condition.

6.5.1. Subjects

16 3-year-olds were tested, aged 3;1 to 4;1 (mean age: 3;7). All were from day care centers in the Boston area.

6.5.2. Materials and Procedure

The stimuli were the same as in Study 2 and so was the procedure, with the following differences, which were designed to reduce the children's tendency to bang on the bell, and to also reduce the effect of order. One difference is that each child was tested on separate trials, at least two days apart. On each trial, the child was given one object story and one sound story. A second difference was that instead of giving both the object practice and the sound practice at the beginning of the study -- before the child was actually tested -- the children were given the object practice immediately before each of the two object stories and the sound practice immediately before each of the two sound stories. Finally, the sound practice was expanded as follows:

> Do you know that you can use this stick to make sounds? You can make sounds by hitting it against this thing [pointing to the bell]. Do you want to play with it?
>
> Now each of these bells makes a different sound. And I'm going to teach you names for some of these sounds and then the Queen Puppet will ask you to hit the bell and make some sounds. But you have to be very careful not to hit it more times than the Queen Puppet asks you to, because she hates a lot of noise when she doesn't ask for it [at this point, the puppet clutches her ears and moans]. Ok?

6.5.3. Results

As in Study 2, separate ANOVAs were performed for the food and sound condition, with stimuli assignment, question order, nonsense word assignment, and word syntax as the independent variables and with the child's response ("0" if she gave one item or made one sound, "1" if she gave more than one item or made more than one sound) as the dependent variable. There were no effects of stimuli assignment, question order, or nonsense word assignment for either the food or sound analyses, nor were any of the interactions significant. Both of the word syntax manipulations were highly significant, however. For the object condition, $F(1,8) = 7.00$, $p = 0.007$, one-tailed, for the sound condition, $F(1,8) = 8.00$, $p = 0.005$, one-tailed. The results are shown in Table 6.8.

Table 6.8 -- Children response's on all questions as a function of syntax

Object

	Count	**Mass**
one	12	5
more than one	4	11

Sound

	Count	**Mass**
one	9	4
more than one	7	12

A second set of analyses were performed, this time focusing only the children's first response for the each condition. These results are shown in 6.9. The effect is significant for both (Objects: $p = 0.07$, Sounds: $p = 0.01$, both Fisher Exact Tests, one-tailed.) All of this suggests that even 3-year-olds are sensitive to the semantic basis of the count-mass distinction for both material and non-material entities.

Table 6.9 -- Children's responses on initial questions as a function of syntax

Object

	Count	Mass
one	6	2
more than one	2	6

Sound

	Count	Mass
one	5	0
more than one	3	8

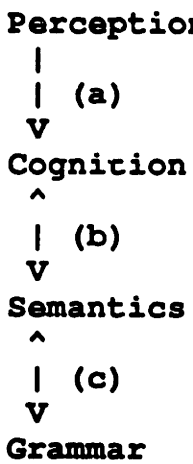
6.6. What develops?

Conceptual development happens -- adults construe the world very differently from children. For instance, 4-year-olds think very differently from older children and adults within the domain of biology; differences have been found in beliefs and inferences about animals, plants, humans, gender, reproduction, illness, and death (Carey, 1985, 1989; Keil, 1989). It has been hypothesized that these phenomena may be due to theory change; children have different theories of the world than adults do, and the developmental shift may be in some way akin to a "paradigm shift" (Kuhn, 1962) in the thought of adult scientists (Carey, 1985, 1989). A second general sort of change is the shift from novice-to-expert that has been intensively studied in adults (see Chi, Glaser, and Rees, 1982). Adult experts in chess, for instance, think very differently from novices with regard to how they categorize chess problems, memorize the position of pieces on the board, and plan their moves. It is not merely that the experts know more strategy than the novices or that they have better memories or can plan faster -- in some important sense, experts and novices

think *differently*. And for virtually all human capacities where there are individual differences in ability that are due to practice, learning, and experience, young children are going to do much worse than adults. Relatively speaking, adults are experts in most everything and children are "universal novices" (Brown and De Loache, 1978).

Does any of this affect the cognitive basis of the count-mass distinction? Consider again the framework developed in Chapters 3 and 4, slightly abbreviated so as to focus on issues relevant to this discussion:

(6)



First, what *doesn't* change? Following Fodor (1983), I will assume that link (a) is modular, informationally encapsulated from other aspects of cognition, and thus applies independently of any cognitive development. However the assumption is not necessary. If perception was non-modular, then the boxology would change somewhat, and link (b) would relate semantics and cognition/perception. Nothing critical turns on this difference, because regardless of whether modularity is correct, the link between objective reality and semantic structure is still mediated by the central cognitive system, which is the critical issue here.

The discussion in Section 6.2 concerned how children learn the language-specific aspects of the count-mass distinction, some of which involve link (c). Children

must learn that *a* has the feature [+INDIVIDUAL] and that *water* has the feature [-INDIVIDUAL]. Older children can learn this by attending to semantic inference; they can categorize determiners by attending to the semantic properties of adjacent nouns and vice-versa. But the link between cognition and semantics, link (b), is essential when learning the semantic properties of at least some determiners and nouns, i.e., at the very beginning of language acquisition.

One speculation about the initial link between semantics and cognition for the count-mass distinction is that the semantic feature [+INDIVIDUAL] maps onto the cognitive notions of discrete bounded object. In addition, this semantic primitive is linked with other semantic primitives, like [+SINGULAR] and [+DEFINITE], according to a set of autonomous semantic principles (see Gil, 1987 for discussion). And the cognitive notions are linked to other cognitive notions, having to do with movement and stability and duration and continuity, and they play a role in inference and memory. They also link up with perception; in this way, language interacts with the world.

It follows that conceptual development can change the nature of the semantics-cognition relation. This can happen in two ways: (i) the extension of the cognitive correlates to count/mass to new cognitive domains, and (ii) the creative capacity to construe bounded objects as [-INDIVIDUAL] and unbounded substances as [+INDIVIDUAL]. I will discuss these in turn.

6.6.1. New Vistas

As children develop cognitively, they come to know about different domains, and they must understand concepts within these domains. For instance, computer software is a domain where all 2-year-olds are utter novices and some older children and adults are experts. Some concepts within this domain are described by nouns (e.g., *symbol, gate, flowchart, byte, program,* etc.), none of which can be semantically categorized through innate cognition-semantics links. Only people who understand the domain will be capable of productively assigning count-mass

semantics to such terms, while novices are limited to making inferences from determiner semantics. Thus an expert knows that programs have properties that would lead them to be readily thought of as individuals and could categorize *program* as a count noun, even without hearing the word used with a determiner. But a novice could not infer this from meaning alone and would have to be exposed to linguistic input such as *a program*.

Another such domain is higher mathematics -- statistics, for instance. Why is *kurtosis* a mass noun while *central location* is a count noun? It is possibly because *kurtosis* is "the flatness or peakedness of one distribution with respect to another", and is thus a continuous notion while *central location* is "a value of the variable near the center of the frequency distribution", and thus is an individual point (definitions from Ferguson, 1981: 31). Moving onto more exciting pursuits, if you have never played poker, you are unlikely to be able to figure out why *ante* is a count noun while *betting* is a mass noun. And someone who is sexually inexperienced is unlikely to figure out, in the absence of grammatical cues, why *foreplay* is a mass noun while *orgasm* is a count noun.

It should be clear that the problems here are different from, say, not knowing the count-mass status of *aardvark* because you have no idea what the word means. It is not the banal point that you cannot determine a word's semantic properties through cognition if you do not know anything about the word. Rather it may be that there is considerable restructuring necessary to understand how the initial cognitive correlates of the semantic count/mass distinction can apply in unusual domains. In domains such as computer programming, abstract mathematics, card games, and sex, you have to have some understanding of the domain before you can be sensitive to how the relevant semantic distinctions apply within that domain. An interesting question is whether there exist cognitive domains that are familiar to both children and adults, and where all adults can readily extend the cognitive correlates of count/mass semantics, and children cannot.

Importantly, it is not the semantic count-mass distinction itself that is changing,

nor the link between cognition and semantics. It is cognition that changes, and, as a result of the interaction between thought and semantics, indirectly affects the course of the acquisition of grammar and meaning.

6.6.2. Cognitive Flexibility

The above section concerned the acquisition of count nouns and mass nouns which describe concepts in unfamiliar domains, where the semantic correlates are not wired in by evolution. But there is a flip side to this problem. What about cases where the innate link exists, but where the child must learn semantic categorizations that go *against* this link, where mass nouns describe bounded objects and count nouns describe unbounded individuals? The first case comes up for solid substance words, like *wood* and *steel*, which can be used to describe what appear to the child to be bounded objects (e.g., pieces of wood or steel). The second case concerns count nouns that have a "kind-of" interpretation, as when describing the contents of a glass as *a liquid* or *a dry wine*.

As for the first case, children definitely have a problem with solid substance nouns. They are acquired late and with some difficulty (Au and Markman, 1987; Dickinson, 1986, 1988; Markman and Wachtel, 1988; Soja, 1987). Interestingly, the problem with these nouns appears to apply only to mapping cognition to semantics, in construing an object as [-INDIVIDUAL]. Young children have no problem with *non-linguistic* inductions over solid substances (Massey and Gelman, 1988; Soja, 1987).

In a recent study, Prasada (1990) showed that children as young as 2-and-a-half can learn solid substance names when they are taught in the proper context. This is an interesting finding, but two caveats are necessary. First, Prasada did not show that children encoded these words as mass nouns; they could just as well have been encoded as adjectives (e.g., as in *wooden*) which would not run counter to any cognition-semantics link. And second, Prasada's experiment was designed to exclude the object word interpretation. This was done by using homogeneous

items, by exposing children to contrastive stimuli to focus on the substance interpretation, and, most important, by using items which already had object names so that the principle of contrast (Clark, 1987) would drive the children away from construing the new word as also being an object name. Thus Prasada's experiment does not run against the finding that children find it very difficult to learn mass nouns that describe entities that are perceived as bounded objects.

The other phenomena, of count nouns like *liquid*, has never been studied in children. My own search of children's speech (the five children examined in Study 1) found such count nouns (or such exceptional usages of mass nouns) to be non-existent, which is hardly a surprise.[34] There is no evidence one way or another for these words.

Of course, adults are capable of learning words like *liquid* and words like *wood*. These are tame examples; at our peak, adults can use creative re-construals to map non-linguistic ideas onto grammar in ways that utterly defy our initial biases. So consider the languages of the fictional planet Tion, none of which have nouns. In the languages of the southern hemisphere

> there are impersonal verbs, modified by nonsyllabic suffixes (or prefixes) with an adverbial value. For example: there is no word corresponding to the word "moon", but there is a verb which in English would be ... "to moonate".

while in the languages of the northern hemisphere

[34]The word *drink* does sometimes appear as a count noun in children's speech, as in *I want a drink*. But this usage does not run against the relation between [+INDIVIDUAL] and bounded objects, because *drink* is used here to describe a bounded object, e.g., a glass of milk. A more relevant case would be if a child said something like *Milk is a great drink*, where *drink* has the "kind-of" interpretation. No such usages appeared in the transcripts that I analyzed.

> the prime unit is not the verb, but the monosyllabic adjective. The noun is formed by an accumulation of adjectives. They do not say "moon" but rather "round airy-light on dark" or "pale-orange-of-the-sky", or any other such combination. In the example above the mass of adjectives refers to a real object, but this is purely fortuitous (Borges, 1964: 8-9).

There are two morals to draw from Borges' fantasy. The first is that adults have great creative power; we can construe an individual object as an intransitive state or as a set of properties. And second, outside of fiction, it is not our nature to do so. Our natural bias is to encode an object in the world as a count noun, not because of any special device for language learning, but due to the nature of language itself.

References

Abney, S. (1987). The English Noun Phrase in its sentential aspect. Unpublished doctoral dissertation, MIT.

Allan, K. (1977). Classifiers. *Language*, **53**, 285-311.

Allan, K. (1980). Nouns and countability. *Language*, **56**, 541-567.

Anglin, J. (1977). *Word, object, and conceptual development*. New York: Norton.

Antell, S. and Keating, D.P. (1983). Perception of numerical invariance in neonates. *Child Development*, **54**, 695-701.

Aronoff, M. (1980). Contextuals. *Language*, **56**, 744-758.

Aronoff, M. (1982). Potential words, actual words, productivity, and frequency. In S. Hattori and K. Inoue (Eds.), *Proceedings of the XIIth International Congress of Linguists*. Tokyo: CIPL.

Ashcraft, M.H. (1976). Property norms for typical and atypical items from 17 categories: A description and discussion. *Memory and Cognition*, **6**, 227-232.

Au, T.K. and Markman, E.M. (1987). Acquiring word meanings via linguistic contrast. *Cognitive Development*, **2**, 217-236.

Bach, E. (1986). The algerbra of events. *Linguistics and Philosophy*, **9**, 5-16.

Bach, E. (1989). *Informal lectures on formal semantics*. New York: SUNY.

Barwise, J. and Cooper, R. (1981). Generalized quantifiers and natural language. *Linguistics and Philosophy*, **4**, 159-219.

Bates, E. and MacWhinney, B. (1987). Competition, variation, and language learning. In B. MacWhinney (Ed.), *Mechanisms of language acquisition*. Hillsdale, NJ: Erlbaum.

Bates, E. and MacWhinney, B. (1989). Functionalism and the competition model. In B. MacWhinney and E. Bates (Eds.), *The crosslinguistic study of sentence processing*. New York: Cambridge University Press.

Battig, W.F. and Montague, W.E. (1969). Category norms for verbal items in 56 categories: A replication and extension of the Connecticut Category Norms. *Journal of Experimental Psychology Monographs*, **80**.

Benedict, H. (1979). Early lexical development: Comprehension and production. *Journal of Child Language*, **6**, 183-200.

Berwick, R.C. (1986). *The acquisition of syntactic knowledge.* Cambridge, MA: MIT Press.

Bierwisch, M. (1981). Basic issues in the development of word meaning. In W. Deutsch (Ed.), *The child's construction of language.* NY: Academic.

Blewitt, P. (1983). Dog versus collie: Vocabulary in the speech of young children. *Developmental Psychology*, **19**, 602-609.

Bloom, L. (1970). *Language development: form and function in emerging grammars.* Cambridge, MA: MIT Press.

Bloom, L. (1973). *One word at a time: the use of single word utterances before syntax.* The Hague: Mouton.

Bloom, L. and Lahey, M. (1978). *Language development and language disorders.* New York: Wiley.

Bloom, L., Lightbown, P. and Hood, L. (1975). Structure and variation in child language. *Monographs of the Society for Research in Child Development*, **40**.

Bloom, P. (1990a). Syntactic distinctions in child language. *Journal of Child Language.* **17**, 343-355.

Bloom, P. (1990b). Subjectless sentences in child language. *Linguistic Inquiry.* **21**, 491-504.

Bloom, P. (1990c). Functions of the mind. *Mind and Language*, in press.

Bloomfield, L. (1933). *Language.* NY: Holt.

Bolinger, D. (1967). Adjectives in English: attribution and predication. *Lingua*, **18**, 1-34.

Borer, H. and Wexler, K. (1987). The maturation of syntax. In T. Roeper and E. Williams (Eds.), *Parameter-setting and language acquisition.* Dordrecht: Reidel.

Borges, J.L. (1962). Tion, Uqbar, Orbis Tertius. In J.L. Borges (Ed.), *Labyrinths.* New York: New Directions.

Bowerman, M. (1973). *Early syntactic development: a cross-linguistic study, with special reference to Finnish.* New York: Cambridge University Press.

Bowerman, M. (1982). Reorganizational processes in lexical and syntactic development. In E. Wanner and L.R. Gleitman (Eds.), *Language acquisition: the state of the art*. New York: Cambridge University Press.

Braine, M.D.S. (1976). Children's first word combinations. *Monographs of the Society for Research in Child Development*, **41**.

Braine, M.D.S. (1987). What in learned in acquiring word classes – A step towards an acquisition theory. In B. MacWhinney (Ed.), *Mechanisms of language acquisition*. Hillsdale, NJ: Erlbaum.

Bresnan, J. (Ed.) (1982). *The mental representation of grammatical relations*. Cambridge, MA: MIT Press.

Bresnan, J. and Grimshaw, J. (1978). The syntax of free relatives in English. *Linguistic Inquiry*, **9**, 331-391.

Brown, A.L. and De Loache, J.S. (1978). Skills, plans, and self-regulation. In R. Siegler (Ed.), *Children's thinking: What develops?* Hillsdale, NJ: Erlbaum.

Brown, R. (1957). Linguistic determinism and the part of speech. *Journal of Abnormal and Social Psychology*, **55**, 1-5.

Brown, R. (1958). How shall a thing be called? *Psychological Review*, **65**, 14-21.

Brown, R. (1973). *A first language: The early stages*. Cambridge, MA: Harvard University Press.

Bunt, H.C. (1979). Ensembles and formal semantic properties of mass terms. In F. Pelletier (Ed.) *Mass terms: Some philosophical problems*. Dordrecht: Reidel.

Burge, T. (1979). Individualism and the mental. In P.A. French, T.E. Ueling Jr., and H.K. Wettstein (Eds.), *Midwest Studies in Philosophy, Vol. IV*. Minneapolis: Minneapolis University Press.

Bybee, J. (1985). *Morphology: A study of the relation between meaning and form*. Philadelphia: Benjamins.

Callanan, M.A. (1985). How parents label objects for young children: The role of input in the acquisition of category hierarchies. *Child Development*, **56**, 508-523.

Carey, S. (1978). The child as word learner. In M. Halle, J. Bresnan, and G. Miller (Eds.), *Linguistic theory and psychological reality*. Cambridge, MA: MIT Press.

Carey, S. (1982). *Semantic development: the state of the art.* In E. Wanner and L.R. Gleitman (Eds.) *Language acquisition: The state of the art.* New York: Cambridge University Press.

Carey, S. (1986). *Conceptual change in childhood.* Cambridge: MIT Press.

Carey, S. (1989). Conceptual differences between children and adults. *Mind and Language,* **3**, 167-181.

Carlson, G. (1977). Reference to kinds in English. Unpublished doctoral dissertation. Department of Linguistics. University of Massachusetts, Amherst.

Cartwright, H. (1979). Some remarks about mass terms and plurality. *Synthese,* **31**, 395-410.

Chi, M., Glaser, R., and Rees, E. (1982). Expertise in problem solving. In R. Sternberg (Ed.), *Advances in the psychology of human intelligence, Vol 1.* Hillsdale, NJ: Erlbaum.

Chi, M. and Klahr, D. (1975). Span and rate of apprehension in children and adults. *Journal of Experimental Child Psychology,* **19**, 434-439.

Chomsky, N. (1957). *Syntactic structures.* Mouton: The Hague.

Chomsky, N. (1965). *Aspects of the theory of syntax.* Cambridge, MA: MIT Press.

Chomsky, N. (1977). *Language and responsibility.* New York: Random House.

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

Chomsky, N. (1981). *Lectures on government and binding.* Dordrecht: Foris.

Chomsky, N. (1986). *Knowledge of language: Its nature, origin, and use.* New York: Prager.

Clark, E.V. (1973). What's in a word? On the child's acquisition of semantics in his first language. In T.E. Moore (Ed.), *Cognitive development and the acquisition of language.* New York: Academic.

Clark, E.V. (1983). Meaning and concepts. In J.H. Flavell and E.M. Markman (Eds.), P.H. Mussen (Series Ed.) *Handbook of Child Development, Vol 3. Cognitive Development.* New York: Wiley.

Clark, E.V. (1987). The principle of contrast: A constraint on language acquisition. In B. MacWhinney, (Ed.), *Mechanisms of language acquisition*. Hollsdale, NJ: Erlbaum.

Cooper, W.E. and Paccia-Cooper, J. (1980). *Syntax and speech*. Cambridge, MA: Harvard University Press.

Creswell, M. (1978). Semantic competence. In M. Guenther-Reutter and F. Guenther (Eds.), *Meaning and translation*. London: Duckworth.

Cutler, A. (1982). Idioms: The colder the older. *Linguistic Inquiry*, **13**, 317-320.

di Sciullo, A.M., and Williams, E. (1987). *On the definition of word*. Cambridge, MA: MIT Press.

Dickinson, D.K. (1986). Learning about materials: A study of conceptual development. Paper presented at the meeting of the Jean Piaget Society, Philadelphia, PA.

Dickinson, D.K. (1988). Learning names for materials: Factors constraining and limiting hypotheses about word meaning. *Cognitive Development*, **3**, 15-35.

Dik, S.C. (1987). A typology of entities. In *Ins and outs of the predication*. Dordrecht: Foris.

Dowty, D.R. (1989). Type raising, function composition, and cross-constituent conjunction. In D. Oehrle, E. Bach, and Wheeler (Eds.), *Categorical grammar and natural language structures*. Kluwer.

Emmonds, J.E. (1987). The invisible category principle. *Linguistic Inquiry*, **18**, 613-632.

Enc, M. (1988). The syntax-semantics interface. In F.J. Newmeyer (Ed.), *Linguistics: The Cambridge Survey, Vol 1: Linguistic Theory: Foundations*. New York: Cambridge University Press.

Fagan, J.F. and Singer (1979). The role of simple feature differences in infant's recognition of faces. *Infant Behavior and Development*, **2**, 29-45.

Ferguson, G.A. (1981). *Statistical analysis in psychology and education*. New York: McGraw-Hill.

Fodor, J. (1966). How to learn to talk: some simple ways. In F. Smith and G. Miller (Eds.), *The genesis of language*. Cambridge, MA: MIT Press.

Fodor, J. (1975). *The language of thought*. New York: Crowell.

Fodor, J. (1981). The present status of the innateness controversy. In J.A. Fodor (Eds.), *Representations*, Cambridge: MIT Press.

Fodor, J. (1983). *The modularity of mind*. Cambridge, MA: MIT Press.

Fodor, J. (1987). *Psychosemantics*. Cambridge, MA: MIT Press.

Fukui, N. (1987). A theory of category projection and its implications. Unpublished doctoral dissertation, MIT.

Gathercole, V. (1985). "He has too many hard questions": The acquisition of the linguistic mass-count distinction in *much* and *many*. *Journal of Child Language*, **12**, 395-415.

Gathercole, V. (1986). Evaluating competing linguistic theories with child language data: The case of the mass-count distinction. *Linguistics and Philosophy*, **9**, 151-190.

Gazdar, G., Klein, E., Pullum, G.K. and Sag, I.A. (1985). *Generalized phrase structure grammar*. Cambridge, MA: Harvard University Press.

Gelman, R. and Baillergeon, R. (1983). A review of some Piagetian concepts. In J.H. Flavell and E.M. Markman (Eds.), P.H. Mussen (Series Ed.) *Handbook of child development, Vol 3: Cognitive development*. New York: Wiley.

Gelman, S.A. (1984). Children's inductive inferences from natural kind and artifact categories. Unpublished doctoral dissertation, Stanford University.

Gelman, S.A., and Markman, E.M. (1985). Implicit contrast in adjectives vs. nouns: Implications for word-learning in preschoolers. *Journal of Child Language*, **12**, 125-143.

Gelman, S.A., and Taylor, M. (1984). How two-year-old children interpret proper and common names for unfamiliar objects. *Child Development*, **55**, 1535-1540.

Gentner, D. (1982). Why nouns are learned before verbs: Linguistic relativity versus natural partitioning. In S. A. Kuczaj II (Ed.), *Language development, Vol II: Language, thought, and culture*. Hillsdale, NJ: Erlbaum.

Gibbs, R. and Nayak, N. (1989). Psycholinguistic studies on the syntactic behavior of idioms. *Cognitive Psychology*, **21**, 100-138.

Gil, D. (1987). Definiteness, Noun Phrase Configurationality, and the Count-mass distinction. In E.J. Reuland and A. ter Meulen (Eds.), *The representation of (in)definiteness*. Cambridge, MA: MIT Press.

Gleason, H.A. (1969). *An introduction to descriptive linguistics.* London, Holt: Rinehart and Winston.

Gleitman, L.R. (1981). Maturational determinants of language growth. *Cognition*, **10**, 103-114.

Gleitman, L.R. (1990). The structural sources of word meaning. *Language Acquisition*, **1**, 3-55.

Gleitman, L.R., Gleitman, H., Landau, B. and Wanner, E. (1987). Where learning begins: Initial representations for language learning. In F.J. Newmeyer (Ed.), *Linguistics: The Cambridge survey: Volume III, Language: Psychological and biological aspects*. New York: Cambridge University Press.

Gordon, P. (1982). The acquisition of syntactic categories: The case of the count/mass distinction. Unpublished doctoral dissertation, MIT.

Gordon, P. (1985). Evaluating the semantic categories hypothesis: The case of the count/mass distinction. *Cognition*, **20**, 209-242.

Gordon, P. (1987). Determiner and adjective categories in children's grammar. Paper presented at the Society for Research in Child Development, 1987, Baltimore, MD.

Gordon, P. (1988). Count/mass category acquisition: Distributional distinctions in children's speech. *Journal of Child Language*, **15**, 109-128.

Gordon, P. (1989). Object, substance and individuation: Canonical vs. non-canonical count/mass nouns in children's speech. Unpublished manuscript, Department of Psychology, University of Pittsburgh.

Grandy, R. (1973). Reply to Moravcsik. In J. Hintikka et al. (Eds.), *Approaches to natural language*, Dordrecht: Reidel.

Green, G.M. (1974). *Semantics and syntactic regularity.* Bloomington: Indiana University Press.

Greenberg, J.H. (1966). Some universals of grammar with particular reference to the order of meaningful elements. In J. H. Greenberg (Ed.), *Universals of language.* Cambridge, MA: MIT Press.

Greenberg, J.H. (1978). Generalizations about numerical systems. In J.H. Greenberg (Ed.), *Universals of human language, Volume 3: Word structure.* Stanford: Stanford University Press.

Grimshaw, J. (1979). Complement selection and the lexicon. *Linguistic Inquiry,* **10,** 279-326.

Grimshaw, J. (1981). Form, function, and the language acquisition device. In C.L. Baker and J. McCarthy, (Eds.), *The logical problem of language acquisition.* Cambridge, MA: MIT Press.

Grimshaw, J. (1990). *Argument structure.* Cambridge, MA: MIT Press.

Gropen, J. (1989). Learning locative verbs: How universal linking rules constrain productivity. Unpublished doctoral dissertation, MIT.

Gruber, J. (1965). Studies in lexical relations. Doctoral dissertation, MIT. Reprinted, 1976, as *Lexical structures in syntax and semantics.* Amsterdam: North-Holland.

Gupta, A.K. (1980). *The logic of common nouns.* New Haven: Yale University Press.

Hale, K. and Keyser, S.J. (1987). A view from the middle. *Lexicon Project Working Papers #10.* Cambridge, MA: MIT Center for Cognitive Science.

Heim, I. (1982). The semantics of definite and indefinite Noun Phrases. Unpublished doctoral dissertation, University of Massachusetts, Amherst.

Higginbotham, J. (1983). Logical form, binding, and nominals. *Linguistic Inquiry,* **14,** 395-420.

Higginson, R.P. (1985). Fixing: A demonstration of the child's active role in language acquisition. Unpublished doctoral dissertation, Washington State University.

Hirsh-Pasek, K., Kemler-Nelson, D.G., Jusczyk, P.W., Wright, K., Cassidy, K., Druss,

B., and Kennedy, B. (1987). Clauses are perceptual units for young infants. *Cognition*, **26**, 269-286.

Hopper, P.J. and Thompson, S.A. (1984). The discourse basis for lexical categories in universal grammar. *Language*, **60**, 703-751.

Horton, M.S. and Markman, E.S. (1980). Developmental differences in the acquisition of basic and superordinate categories. *Child Development*, **51**, 708-719.

Huttenlocher, J. and Smiley, P. (1987). Early word meanings: The case of object names. *Cognitive Psychology*, **19**, 63-89.

Hyams, N. (1986). *Language acquisition and the theory of parameters.* Dordrecht: Reidel.

Ingram, D. (1978). Typology and universals of personal pronouns. In J.H. Greenberg (Ed.), *Universals of human language, Volume 3: Word structure.* Stanford: Stanford University Press.

Jackendoff, R. (1977). *X-bar syntax: a study of phrase structure.* Cambridge, MA: MIT Press.

Jackendoff, R. (1983). *Semantics and cognition.* Cambridge, MA: MIT Press.

Jackendoff, R. (1987). The status of thematic relations in linguistic theory. *Linguistic Inquiry*, **18**, 369-411.

Jackendoff, R. (1990a). *Semantic structures.* Cambridge, MA: MIT Press.

Jackendoff, R. (1990b). Parts and boundaries. *Cognition*, in press.

Joshi, A.K. and Levy, L.S. (1982). Phrase structure trees can bear more fruit than you would have thought. *American Journal of Computational Linguistics*, **8**, 1-12.

Karmiloff-Smith, A. (1979). *A functional approach to language acquisition.* New York: Cambridge University Press.

Katz, J. and Fodor, J. (1963). The structure of a semantic theory. *Language*, **39**, 170-210.

Katz, N., Baker, E., and Macnamara, J. (1974). What's in a name? A study of how children learn common and proper names. *Child Development*, **45**, 469-473.

Keil, F.C. (1979). *Semantic and conceptual development; An ontological perspective.* Cambridge, MA: MIT Press.

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

Kim, [illegible]-J. (1987). The acquisition of relative clauses in English and Korean: Development in spontaneous production. Unpublished doctoral dissertation, Harvard University.

Klima, E.S. (1967). Negation in English. In J. Fodor and J. Katz (Eds.), *The structure of language.* Englewood Cliffs, NJ: Prentice-Hall.

Kucera, N. and Frances, W.N. (1986). *Computational analysis of present-day American English.* Providence, RI: Brown University Press.

Kuhn, T.S. (1962). *The structure of scientific revolutions.* Chicago: University of Chicago Press.

Lakoff, G. (1970). Global rules. *Language,* **46**, 627-639.

Lakoff, G. (1987). *Women, fire, and dangerous things: what categories reveal about the mind.* Chicago: Chicago University Press.

Landau, E. and Gleitman, L.R. (1985). *Language and experience.* Cambridge, MA: Harvard University Press.

Langacker, R.W. (1987). Nouns and verbs. *Language,* **63**, 53-94.

Langacker, R.W. (1990). *Foundations of cognitive grammar, Volume II: Descriptive application,* Unpublished manuscript, UCSD.

Larson, R. (1985). Bare NP adverbs. *Linguistic Inquiry,* **16**, 595-621.

Levin, B. and Rappaport, M. (1986). The formation of adjectival passives. *Linguistic Inquiry,* **17**, 623-661.

Levy, Y. (1988a). On the early learning of formal grammatical systems: Evidence from studies of the acquisition of gender and countability, *Journal of Child Language,* **15**, 179-186.

Levy, Y. (1988b). The nature of early language: Evidence from the development of Hebrew morphology. In Y. Levy, I.M. Schlesinger, and M.D.S. Braine (Eds.), *Categories and processes in language acquisition.* Hillsdale, NJ: Erlbaum.

Levy, Y., Schlesinger, I.M. and Braine, M.D.S. (Eds.) (1988). *Categories and processes in language acquisition.* Hillsdale, NJ: Erlbaum.

Lewis, D. (1972). General semantics. In D. Davidson and G. Harman (Eds)., *Semantics of natural language*. Dordrecht: Reidel.

Macnamara, J. (1982). *Names for things: A study of human learning*. Cambridge, MA: MIT Press.

Macnamara, J. (1986). *A border dispute: The place of logic in psychology*. Cambridge, MA: MIT Press.

Macnamara, J. and Reyes, G. (1990). The learning of proper names and count nouns: Foundational and empirical issues. Unpublished manuscript, McGill University and Universite de Montreal.

MacWhinney, B., and Snow, C. (1985). The Child Language Data Exchange System. *Journal of Child Language*, **12**, 271-296.

Maratsos, M. (1976). *The use of definite and indefinite reference in young children*. New York: Cambridge University Press.

Maratsos, M. (1982). The child's construction of grammatical categories. In L.R. Gleitman and H. Wanner (Eds.), *Language acquisition: the state of the art*. New York: Cambridge University Press.

Maratsos, M. (1983). Some current issues in the study of the acquisition of grammar. In P.H. Mussen (Ed.), *Handbook of child psychology* (4th ed., Vol 3) *Cognitive development*. New York: Wiley.

Maratsos, M. (1988). Vote yes on nouns for semantic bootstrapping. Paper presented at the Thirteenth Annual Boston University Language Conference, Boston, MA.

Maratsos, M. (1989). Metaphors of language: Metaphors of the mind? (Review of G. Lakoff's "Women, fire, and dangerous things"), *Contemporary Psychology*, **34**, 5-7.

Maratsos, M., and Chalkley, M.A. (1981). The internal language of children's syntax: The ontogenesis and representation of syntactic categories. In K. Nelson (Ed.), *Children's language, Vol. 2*. New York: Gardner Press.

Markman, E.M. (1985). Why superordinate category names can be mass nouns. *Cognition*, **19** 31-53.

Markman, E.M. (1989). *Categorization and naming in children.* Cambridge, MA: MIT Press.

Markman, E.M., Horton, M.S., and McLanahan, A.G. (1980). Classes and collections: Principles of Organization in the learning of hierarchical relations. *Cognition,* **8,** 227-241.

Markman, E.M. and Hutchinson, J.E. (1984). Children's sensitivity to constraints in word meaning: Taxonomic versus thematic relations. *Cognitive Psychology,* **16,** 1-27.

Markman, E.M. and Wachtel, G.F. (1988). Children's use of mutual exclusivity to constrain the meaning of words. *Cognitive Psychology,* **20,** 121-157.

Massey, C.M. and Gelman, R. (1988). Preschoolers ability to decide whether a photographed unfamiliar object can move itself. *Developmental Psychology,* **24,** 307-317.

Mayr, E. (1982). *The growth of biological thought.* Cambridge, MA: Harvard University Press.

McBain, E. (1985). *Eight black horses.* New York: Avon.

McCawley, J. (1975). Lexicography and the count-mass distinction. *Proceedings of the first annual conference of the Berkeley Linguistic Society.*

Montague, R. (1974). *Formal philosophy.* (Edited by R.H. Thomason). New Haven, Conn.: Yale University Press.

Moravcsik, J. (1979). Mass terms in English. In F. Pelletier (Ed.) *Mass terms: Some philosophical problems.* Dordrecht: Reidel.

Mufwene, S. (1984). The count/mass distinction and the English lexicon. *1984 CLS Parasession on Lexical Semantics.*

Naigles, L. (1990). Children use syntax to learn verb meanings. *Journal of Child Language,* 357-374.

Nelson, K. (1976). Some attributes of adjectives used by young children. *Cognition,* **4,** 1-31.

Newmeyer, F. (1983). *Grammatical Theory: Its limits and possibilities.* Chicago: University of Chicago Press.

Newmeyer, F. (1990). Functional explanation in linguistics and the origin of language. *Language and Communication*, in press.

Nishigauchi, T. and Roeper, T. (1987). Deductive parameters and the growth of empty categories. In T. Roeper and E. Williams (Eds.) *Parameter setting and language-acquisition*. D. Riedel: Dordrecht.

Oatway, R. (1990). Semantic and syntactic cues in the acquisition of novel nouns and noun syntax. Unpublished Bachelor Degree Thesis, MIT.

O'Grady, W. (1987). *Principles of grammar and learning*. Chicago: University of Chicago Press.

Palmer, F.R. (1990). The semantics of grammar. (Review of A. Wierzbicka's, "The semantics of grammar"). *Journal of Linguistics*, **26**, 223-233.

Partee, B.H. (1979). Semantics--mathematics or psychology? In R. Bauerle, U. Egli, and A. von Stechow (Eds.) *Semantics from different points of view*. Berlin: de Gruyter.

Partee, B.H. and Rooth, M. (1983). Generalized quantifiers and type ambiguity. In R. Bauerle, C. Schwarze, and A. von Stechow (Eds.), *Meaning, use, and the interpretation of language*. Berlin: de Gruyter.

Partee, B.H. (1987). Noun phrase interpretation and type-shifting principles. In J. Groenendijk, D. de Jongh, and M. Stokhof (Eds.) *Studies in discourse representation theory and the theory of generalized quantifiers*. Dordrecht: Foris.

Pelletier, F. (1979). Non-Singular reference: Some preliminaries. In F. Pelletier (Ed.), *Mass terms: Some philosophical problems*. Dordrecht: Reidel.

Pesetsky, D. (1983). Paths and categories. Unpublished doctoral dissertation, Department of Linguistics and Philosophy, MIT.

Piaget, J. (1982). The psychogenesis of knowledge and its epistemological significance. In M. Piattelli-Palmarini (Ed.), *Language and learning: The debate between Jean Piaget and Noam Chomsky*. Cambridge, MA: Harvard University Press.

Pinker, S. (1979). Formal models of language learning. *Cognition*, **1**, 217-283.

Pinker, S. (1984). *Language learnability and language development.* Cambridge, MA: Harvard University Press.

Pinker, S. (1987). The bootstrapping problem in language acquisition. In B. MacWhinney (Ed.), *Mechanisms of language acquisition.* Hillsdale, NJ: Erlbaum.

Pinker, S. (1989). *Learnability and cognition.* Cambridge, MA: Harvard University Press.

Pinker, S. (1990). Of mice and men. Unpublished manuscript, MIT.

Pinker, S., Lebeaux, D.S. and Frost, L.A. (1987). Productivity and constraints in the acquisition of the passive. *Cognition*, **26**, 195-267.

Postal, P.M. and Pullum, G.K. (1988). Expletive Noun Phrases in subcategorized positions. *Linguistic Inquiry*, **19**, 635-670.

Potter, M.C. and Levy, E.I. (1968). Spatial ennumeration without counting. *Child Development*, **39**, 265-272.

Prasada, S. (1990). Young children's linguistic and nonlinguistic knowledge of solid substances. Unpublished manuscript, MIT.

Pustejovsky, J. (1989). Type coercion and selection. Paper presented at West Coast Conference in Formal Linguistics, Vancover, BC.

Pye, C. (1990). The ergative parameter. *Journal of Linguistics*, in press.

Quine, W.V. (1960). *Word and object.* Cambridge, MA: MIT Press.

Quirk, R., Greenbaum, S., Leech, G., and Svartvik, J. (1985). *A comprehensive grammar of the English language.* New York: Longman.

Radford, A. (1988). *Transformational grammar.* New York: Cambridge University Press.

Richards, D.D. and Goldfarb, J. (1986). The episodic memory model of conceptual development: An integrative viewpoint. *Cognitive Development*, **1**, 183-219.

Rosch, E. (1975). Cognitive representations of semantic categories. *Journal of Experimental Psychology*, **104**, 192-223.

Rosch, E. (1978). Principles of categorization. In E. Rosch and B. Lloyd (Eds.), *Cognition and categorization.* Hillsdale, NJ: Erlbaum.

Rosch, E., Mervis, C.B., Gray, W.D., Johnson, D.M., and Boyes-Braem, P. (1976). Basic objects in natural categories. *Cognitive Psychology*, **8**, 382-439.

Rosenthal, R. and Rosnow, R.L. (1984). *Essentials of Behavioral Research*. New York: Mc-Graw Hill.

Rothstein, S. (1983). *The syntactic forms of predication*. Unpublished doctoral dissertation, MIT.

Sachs, J. (1983). Talking about the there and then: The emergence of displaced reference in parent-child discourse. In K.E. Nelson (Ed.), *Children's language, Vol 4*. Hillsdale, NJ: Erlbaum.

Sag, I.A., Gazdar, G., Wasow, T, and Weisler, S. (1985). Coordination and how to distinguish categories. *Natural Language and Linguistic Theory*, **3**, 117-171.

Sapir, E. (1921). *Language*. New York: Harcourt, Brace.

Schacter, P. (1985). Parts-of-speech systems. In T. Shopen (Ed.) *Language typology and syntactic description, Vol 1: Clause Structure*. New York: Cambridge University Press.

Schaeffer, B., Eggleston, V. H., and Scott, J. L. (1974). Number development in young children. *Cognitive Psychology*, **6**, 357-379.

Schieffelin, B. (1985). The acquisition of Kaluli. In D.I. Slobin (Ed.), *The crosslinguistic study of language acquisition, Vol. 1: The data*. Hillsdale, NJ: Erlbaum.

Schlesinger, I.M. (1981). Semantic assimilation in the acquisition of relational categories. In W. Deutsch (Ed.), *The child's construction of language*. New York: Academic Press.

Schlesinger, I.M. (1988). The origin of relational categories. In Y. Levy, I.M. Schlesinger, and M.D.S. Braine (Eds.), *Categories and processes in language acquisition*. Hillsdale, NJ: Erlbaum.

Shipley, E.F., Kuhn, I.F., and Madden, E.C. (1983). Mother's use of superordinate category terms. *Journal of Child Language*, **10**, 571-588.

Silverman, I. W., and Rose, A. P. (1980). Subitizing and counting skills in 3-year-olds. *Developmental Psychology*, **16**, 539-540.

Skinner, B.F. (1957). *Verbal Behavior.* New York: Appleton-Century-Crofts.

Slobin, D.I. (1985). Crosslinguistic evidence for the language-making capacity. In D.I. Slobin (Ed.), *The crosslinguistic study of language acquisition, Vol. 2: Theoretical Issues.* Hillsdale, NJ: Erlbaum.

Soja, N. (1987). Ontological constraints on 2-year-olds' induction of word meanings. Unpublished doctoral dissertation, MIT.

Soja, N. (1990). Semantic inferences: The role of count/mass syntax. *Papers and Reports in Child Language Development,* **29**.

Soja, N., Carey S., and Spelke, E.S. (1990). Ontological categories guide young children's inductions of word meaning: Object terms and substance terms. *Cognition,* in press.

Spelke, E.S. (1988). Where perception ends and thinking begins: The apprehension of objects in infancy. In A. Yonas (Ed.), *Minnesota Symposia on Child Psychology,* Hillsdale, NJ: Erlbaum.

Starkey, P., and Cooper, R.G. Jr. (1980). Perception of numbers by human infants. *Science,* **210**, 1033-1035.

Strauss, M.S. and Curtis, L.E. (1984). Development of numerical concepts in infancy. In C. Sophian (Ed.), *Origins of cognitive skills.* Nillsdale, NJ: Erlbaum.

Talmy, L. (1975). Semantics and syntax of motion. In J. Kimball (Ed.), *Syntax and Semantics, Vol. 4.* New York: Academic.

Talmy, L. (1983). How language structures space. In H. Pick and L. Acredolo (Eds.) *Spatial orientation: Theory, research, and application.* New York: Plenum.

Talmy, L. (1985). Lexicalization patterns: semantic structure in lexical forms. In T. Shopen (Ed.), *Language typology and syntactic description, Vol 3: Grammatical categories and the lexicon.* New York: Cambridge University Press.

Valian, V (1986). Syntactic categories in the speech of young children. *Developmental Psychology,* **22**, 562-579.

Wales, R., Colman, M., and Pattison, P. (1983). How a thing is called: A study of mothers' and childrens' naming. *Journal of Experimental Child Psychology,* **36**, 1-17.

Ware, R. (1979). Some bits and pieces. In F. Pelletier (Ed.), *Mass terms: Some philosophical problems*. Dordrecht: Reidel.

Waxman, S. (1990). Linguistic biases and the establishment of conceptual hierarchies: Evidence from preschool children. *Cognitive Development*. **5**, 123-150.

Waxman, S. and Gelman, R. (1986). Preschoolers' use of superordinate relations in classifications and language, *Cognitive Development*, **1**, 139-156.

Weinreich, U. (1966). Explorations in semantic theory. In T. Sebeok (Ed.), *Current trends in linguistics, Vol. 3*. Mouton: The Hague.

Wellman, H. (1990). *Children's theory of mind*. Cambridge, MA: MIT Press.

Wexler, K. and Cullicover, P. (1980). *Formal principles of language acquisition*. Cambridge, MA: MIT Press.

Wierzbicka, A (1985). Oats and wheat: the fallacy of arbitrariness. In J. Haiman (Ed.), *Iconicity in Syntax*. Amsterdam: Benjamins.

White, T.G. (1982). Naming practices, typicality, and underextension in child language. *Journal of Experimental Child Psychology*, **33**, 324-346.

Whorf, B. (1956). *Language, thought, and reality*. Cambridge, MA: MIT Press.

Williams, E. (1980). Predication. *Linguistic Inquiry*, **11**, 203-238.

Williams, E. (1983). Semantic vs. syntactic categories. *Linguistics and Philosophy*, **6**, 423-446.

Wynn, K. (1990a). Children's understanding of counting, *Cognition*, in press.

Wynn, K. (1990b). Children's acquisition of the meaning of number words, unpublished manuscript, MIT.