

Program of Research

**The Balance of Description and Explanation in Linguistic Theory**

Jacques Lamarche

Department of French,  
The University of Western Ontario

© Jacques Lamarche, May 2005

## Introduction\*

The difference between description and explanation in a theory hinges on the specificity of the statements the theory makes with respect to the facts it seeks to explain. A description of a given fact is a statement that is specific to it, and that provides an account for it alone. An explanation, on the other hand, is an account of a fact through statements that are not specifically about this fact, but that apply elsewhere. The fact is thereby accounted for without recourse to a unique statement in the theory, and is thus explained. In my research, I seek a theoretical understanding of linguistic knowledge by taking the position that the facts of natural language provide a natural division between what must be described, and what should be explained.

The factual knowledge of speakers is their capacity to recognize expressions, that is, to associate strings of sounds (or visual signs, manual or written) with meanings. Any account of this knowledge must recognize the existence of two distinct types of expressions: terms (expressions that cannot be broken down into smaller sound-meaning associations) and constructions (expressions that are made up of smaller expressions). These different realities reflect a natural division between what must be described and what should be explained. Thus, the Saussurean *arbitraire du signe* implies that knowledge about terms can only be described: only specific statements in the theory can provide an account for the fact that *black* in English and *noir* in French are associated with the same notion. In contrast, knowledge about a given construction should be explainable, because speakers can produce and interpret novel complex expressions at will. They can thus recognize an infinite number of sound-meaning relations of their language, and no theory based on specific statements for each expression can provide an adequate account of this knowledge. Any adequate account of linguistic knowledge must therefore recognize the existence of two components in a grammar: a lexical component, a list of basic statements that cannot be explained, but only described; and a combinational component (a morpho-syntactic system), which is responsible for the existence of complex expressions.

Accordingly, all linguistic theories assume the existence of these two components in their account of linguistic competence. However, it is fair to say that no consensus exists in linguistics about what kind of statements are to be made in the lexicon of grammar. Our position with respect to this fundamental question is to assume that the nature of an expression—whether it is simple or complex—is the criterion that determines what should be described and what should be explained in linguistic theory. In other words, all properties of constructions (i.e. complex expressions) are to be explained, and only terms (simple expressions) can be the object of a specific statement in the analysis. More specifically, assuming that specific statements (descriptions) are only made in the lexicon, and no specific statement may be about a construction, it follows that no statement about a term should refer to a construction (a context) in which the term appears. From this follows the terminal criterion, the methodological principle at the center of my research since Lamarche (1998a):

---

\* This document is written as an introduction to my research package for my tenure application at the University of Western Ontario, and should not be cited or distributed without permission.

### **Terminal criterion**

A specific statement, a statement about a term, should be true of that term in all the contexts in which it appears.

The picture of grammar that emerges from the scrupulous application of this criterion is strikingly different from what is generally assumed in contemporary analyses. Under the lens of the terminal criterion, the compositionality of meaning must be altered, because the meaning of a complex expression is always *more* than the meaning of its parts. Many distinctions that are traditionally taken to be properties of terms cannot be considered terminal because they are properties of terms only in context. Under the terminal criterion, these distinctions must be explained, that is, derived by the combinational system.

This leads to the observation that the distinctions relevant for grammatical analysis appear to have an interpretative reality only at the level of combination: under this view, morpho-syntax is a generator of meaning distinctions. In other words, the combinational system not only combines expressions, but also introduces the distinctions that are relevant for grammatical interpretation of expressions. With this hypothesis emerges a clear division in semantics between what belongs to grammar and what is external to it. Semantic distinctions that can be specifically associated with terms—for example substantive concepts and the lexical constants in certain verbs—arguably belong to general cognitive and knowledge of the world. These distinctions have a reality outside of grammar. The distinctions relevant for grammatical semantics—which pertain to the grammatical interpretation of expressions—are a function of the combinational system, and have an interpretative reality at the level of constructions in language.

A truth-value for a sentence can then be seen as a match between a possible state of affairs in general cognition and the constructed distinctions attributed to terms by the combinational system. These distinctions are imposed on the substantive content that is associated with the terms in the sentence: a truth-value will emerge just in case all the distinctions provided by the grammar are compatible with the substantive content.

A strict adherence to the terminal criterion thus reveals a precise program of research that has implications not only for linguistics, but also for other domains of inquiry such as cognitive science and philosophy. This program requires developing a grammatical system where certain meaning distinctions (those pertaining to grammatical semantics) only exist at the level of complex expressions. It should include a component of descriptions that are compatible with the terminal criterion. This component includes two types of specific statements: statements about distinctions that are found outside of grammar, in world knowledge; and statements pertaining to the grammatical markers (the content of determiners, prepositions, inflectional morphology, etc.). The combinational system must be such that when it combines expressions, it introduces only the meaning distinctions that are expressed in complex expressions of the language under study. In principle, the grammar (the combination system and the functional vocabulary that feeds it) can be studied independently of the reality associated with substantive terms, whose reality is only relevant to assess the truth-values of sentences.

However precise the goals of this program of research, its instantiation has proven to be challenging from a formal point of view. For one thing, the terminal criterion is

only capable of excluding information from the descriptive tool of the theory; it does not reveal what the descriptively adequate statements in the theory are. At best, the criterion provides a tool to determine whether the analysis of a specific contrast is adequate or not. Furthermore, because the traditional tools of logic and standard linguistic formalism have been developed within a perspective of grammar that does not respect the terminal criterion, these formal systems cannot be used as the descriptive apparatus of grammatical semantics. Even the formal systems that are designed to account for expressions which are larger than terms (phrases, sentences, discourse) presuppose terminal descriptions that are not terminal, rendering these systems unsuitable for a framework that respects the terminal criterion.<sup>1</sup>

Most of my research, starting with my dissertation (Lamarche 1998a), has been dedicated to the development of a formal system that is compatible with the terminal criterion. My early research focused on determining what is relevant in the interpretation of complex expressions, as well as on making hypotheses about the nature of terminal descriptions that can be fed to a combinational system that would produce relevant distinctions. A combinational system that respects the terminal criterion must be such that the combination of two identical terms can receive distinct values at the output of the combination. To achieve this, I make the assumption that semantic values at the level of constructions are a function of different arrangements of contents in the output: different rules arrange the two inputs differently, and complex expressions made of identical terms can thus receive different values. It became clear early on in my research that this highly formal and integrated system would lend itself to computational modelling, which would be a tremendous tool with which to understand and develop the different components of the theory and test their interaction. Thus, in 2003, I began collaborating with Professor Robert Mercer of the Computer Science Department at UWO, a move that has proven to be extremely fruitful in terms of understanding and formalizing the theory.

In the following pages, I present a synthesis of the more recent results of my research, and provide a comprehensive overview of the research plan that ties together my published material and guides my ongoing research. Much of my early work focused on understanding the fundamental implications of the terminal criterion and argued for its importance from a theoretical point of view. These papers deal with specific problems of verbal semantics. In Lamarche (2002a), which is partly based on my doctoral dissertation, I discuss the polysemy of the French verb *aller* ‘go’ and other related verbs. Lamarche (2002c) discusses the raising and control ambiguity of a verb like *promettre* ‘to promise’ in French. The description of the copula *be* is the topic of Lamarche (1998b, 2002b, 2003), papers which also address the thorny issue of the ambiguity of pseudo-cleft constructions. Copular constructions are a crucial area of study for this approach, because their meaning is based on the most basic element of the grammatical vocabulary, the verb *be*. Given that copular constructions are made up of *be* and two nominal expressions (such as *John is a priest*), to fully understand the construction naturally requires a solid understanding of the nominal system. This question has been at the center of my research

---

<sup>1</sup> It is worth pointing out that the problem does not lie in the relevance of the notions these systems use in their account: rather, the issue lies in the specific associations assumed between these notions and expressions (whether they are terminal or complex). In an account that respects the terminal criterion, the significant notions can only have a reality as the output of the combinational system, and cannot be incorporated as specific statements about expressions.

in the last three years, and Lamarche (2005a, 2005b) constitute preliminary results of this work. My most current research—to be outlined in Lamarche (2005c, 2005d, in preparation)—provides a good measure of the potential of my research program.

Over the next three sections, I illustrate a range of issues that the nominal system poses for an account that respects the terminal criterion, and how my most current research accounts for these empirical problems, before concluding with a brief overview of the computational model that Professor Mercer and I are currently developing, outlining at the same time the direction of my program of research for the future.

### **Description and explanation in the nominal domain**

The distinction between nouns and adjectives is generally considered to be a property of terms. However, the notion of category is clearly based on contextual considerations. Furthermore, the notions these categories represent are not constant properties of many terms, and therefore do not respect the terminal criterion. Let me illustrate the issue with a few examples.

The notions “noun” and “adjective” are associated with certain interpretative properties that characterize terms in specific positions of the sentence. Thus, “adjective” refers to a class of terms that denotes a *property* of some entity, a reading that is correlated with certain positions in a sentence. In English, for example, this interpretation is possible for a term that appears before a noun, as in (1a), after certain intransitive verbs, as in (1b), or after the direct object of certain transitive verbs (1c)

- (1) a. The sad person.
- b. The person is/became sad.
- c. Mary considers/made the person sad.

In contrast, the interpretation is not possible in English for a term that appears with a determiner without a noun, as in (2a), or if the term is inflected for the plural, as in (2b).

- (2) a. \* I saw the sad.
- b. \* sads
- c. \* Sad is a dog.

The position of subject (before the verb) in normal declarative sentence (with no special presupposition or stress pattern) also excludes the adjective as in (2c).

Nouns, on the other hand, are terms that tend to denote entities (rather than properties of entities). The interpretation of “noun” is generally associated with argument positions in the sentence (the subject, the complements of verbs and prepositions). Different classes of nouns can be distinguished depending on their distribution and their capacity to be inflected for plural. For example, what is traditionally called a common noun (or count noun) generally refers to a term that can easily be pluralized (cats, cars, objects), and appears with a determiner in the singular (a cat, the car, this object). In contrast, nouns that appear in the singular without a determiner are either proper nouns or mass nouns (I saw George, I ate beef).

Although the respective semantic values of the grammatical categories “noun” and “adjective” are clearly correlated with the distributional properties of expressions, these categories are generally considered to be properties of terms: *huge* is labelled as an adjective, *Mary* as a proper noun, and so on. An obvious reason for this assumption is the denotation of these terms (that is, the nature of what they designate). Thus, given that the terms *sad*, *beef*, *Mary* and *cat* designate, respectively, a property, a substance, an individual, and a set, these terms are bound to appear in positions of the sentence that are correlated with the semantic interpretation attributed to the categories “adjective”, “mass noun”, “proper noun” and “count noun”. It thus seems perfectly legitimate to associate the categories to the terms themselves.

For many terms of English, however, the terminal criterion excludes a lexical association with such categories. The terminal criterion excludes specific statements about category in the lexical description of any term that can appear in contexts associated with more than one category. The category of the term in such circumstances would not be constant across all contexts, and therefore would not respect the terminal criterion. As is well known, this ability of terms to have different categories in different context is not exceptional, and could arguably be considered the norm. In the following sentences, the terms *red*, *chicken* and *George* illustrate how terms may appear with different categories in different context:

- (3) a. A red car.  
b. A bright red.
- (4) a. I saw a chicken.  
b. I ate chicken.
- (5) a. George is here.  
b. Mary married a George.

*Red* is interpreted as an adjective in (3a), but as a noun in (3b); *chicken* is interpreted as a count noun in (4a), but a mass noun in (4b); and *George* is a proper noun in (5a) (it designates a specific individual), but serves as a count noun in (5b) (the expression means that there is a set of individuals called *George*, and Mary married one of the members of this set).

Under the terminal criterion, any mention of a category at the terminal level is thus excluded for these terms. That is, none of these terms can bear a feature that would account for their capacity to appear in the contexts associated with the interpretation of adjectives, proper nouns, mass nouns and count nouns.<sup>2</sup> Eliminating categories from the description of substantive terms substantially departs from traditional assumptions, but offers an opportunity to reconsider what should constitute the lexical content of substantive terms. In fact, an analysis that respects the terminal criterion must be able to account for the facts that the denotation of a term correlates with its capacity to appear in

---

<sup>2</sup> Of course, an analysis by homonymy (listing two terms with different categories) or rules that alter the category depending on the context—assumed in both formal and functional frameworks, see Partee 1987 and Croft 1991—can provide an account of these facts. Such accounts, however, are not compatible with the terminal criterion, which is based on the hypothesis that what belongs to constructions must be derived. Given that grammatical categories are defined with respect to contextual notions (that is, positions in the sentence), they therefore be explained, not stated in the definition of terms.

certain positions in the context (that is, to appear in position that are associated with categories such as adjectives, proper nouns, and so on). Let me address the question of the content of substantive term, before tackling the issue of the correlation between denotation and grammatical category.

The terminal criterion states that descriptive properties attributed to terms must be constant across all contexts. In Lamarche (2005a), I argue that what is constant with substantive terms is the association between their phonological form and a notion that exists outside of the grammatical domain (call it the cognitive domain for ease of discussion). This is the case in the examples in (3) to (5). Thus, *red* is the same value in the visible light spectrum whether it is interpreted as a noun or as an adjective; the animal that is designated by *a chicken* is the same one that provided the meat we call *chicken*; and a proper name like *George* always designates an individual concept, even when it is used with reference to a set of individuals (each individual in the set of *Georges* has the name *George*). In a sense, the substantive terms are simply labels for notions that exist outside of grammar. With this in mind, I turn to the issue of the correlation between denotation and grammatical categories.

The following analogy provides a useful way of showing how to account for this correlation. Suppose that a grammar is a device that is designed to send information to a hearer in the discourse. The information to be sent is about substantive content, that is, objects in the cognitive domain denoted by terms. Grammar, however, does not send the substantive content directly: rather, it provides *moulds* of different types. For example, one type of mould corresponds to a property; another mould corresponds to a single discontinuous entity, which I call an atom; another mould is the substance mould, design to contain what an atom is made of; and so on.

The functional vocabulary of a language is the material used by the combinational system to construct specific moulds: what is associated with the notion “adjective” is the property mould; “proper noun” the individual atomic mould; “mass noun” the substance mould; and so on. Substantive terms provide labels for the moulds that are sent in the discourse. What they denote is either compatible with a mould or not. Thus, if the “adjective” mould is sent, then the label will need to denote a property (e.g., *huge*); if it is the mould “proper noun” that is sent, then the label will need to denote an individual entity in the discourse (e.g. *Mary*); and so on.

In principle, any label can be assigned to any mould. However, not all denotations associated with labels are compatible with all types of mould. Thus, a term that designates a specific person (a proper name) cannot be used in the “adjective” moulding, because such an expression does not designate a property.<sup>3</sup> However, a proper name can appear in the “count noun” moulding; the name is then imposed on every member of the set that is included in the mould. Terms that designate a reality that can be applied to a set of things (*cat*) will appear in the “count noun” mouldings (constructed with determiners and the plural inflection). Such terms will not appear in the proper name moulding because they designate a set rather than an individual atom. Whether a count

---

<sup>3</sup> In English, the use of a proper name to refer to something that can be considered a property of the individual designated can only be expressed through the use of the combinational system, for example derivational morphology (e.g., the Chomskyan theory of language).

term can appear in the “mass noun” moulding, or vice-versa, depends on the nature of the denotation in question, and other language-specific and cultural factors.<sup>4</sup>

This analysis explains the correlation between denotation and syntactic categorization. Certain denotations will naturally gravitate around certain kinds of moulds. In the lexicon of the grammar, however, the only statement made about the term is its association with a specific concept of the cognitive domain. Nothing else needs to be said: whether *red* denotes a property (or something else) is totally irrelevant to the construction of the mould. Only once a mould is constructed is denotation relevant: it is used to assess whether some aspect of the concept designated by the term can be construed in such a way that is compatible with the value of the mould onto which the term’s label appears. A truth-value emerges when the concept associated with the label fits the value of the mould.

Recently, certain authors (see, for example, Baker 2003, and Borer 2005) have defended the view that lexical features should be totally separate from the features that underlie grammatical interpretation. For example, the position espoused in Borer (2005) is remarkably close to the one I just outlined. There are, however, significant differences between my approach and hers, as will become clear after I present a combinational system that gives rise to grammatical “moulding”.

### **Combinational Semantics**

The terminal criterion is very explicit about what is a possible analysis. In short, it requires a system in which the meaning distinctions associated with the grammatical moulds have a reality only at the level of the combination. Moreover, the distinctions have to be constructed out of specific statements for terms that all apply in all contexts where the terms appear. Before presenting a system that meets these requirements, it is important to be more specific about the relation between labels, concepts and notions in the world.

A label provides a reference to a notion defined in the conceptual domain, and ultimately defines the term.<sup>5</sup> I assume that a label, by definition, only relates to one piece of information at a time. To illustrate what I mean by this, consider the situation depicted in Figure 1.

---

<sup>4</sup> For example, the existence of a specific label associated with the meat of an animal (*beef* for the meat of *cows*) may restrict the use of the label for the animal in a mass reading context.

<sup>5</sup> For example, in true cases of homonymy such as *bank* in English—edge of a river or a financial institution—it is the conceptual domain that provides evidence for the existence of distinct lexical objects.



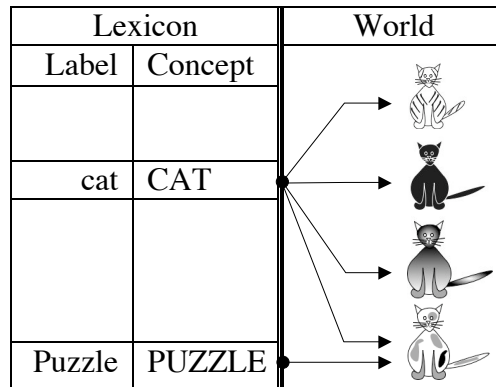


Figure 1

The labels (the form of the term) *cat* and *Puzzle* in the first column respectively relate to the concepts CAT and PUZZLE of the second column, objects of the conceptual domain that exist independently of grammar. How these concepts are characterized is not relevant for the definition of the term: there is likely a web of information associated with each of these concepts, and one of them is clearly embedded in the other (*Puzzle* is a *cat*), but as far as grammar is concerned, this is irrelevant. What grammar cares about is the existence of these conceptual objects, which once associated with a given label provides a value to the substantive term in the lexicon. The concepts themselves also relate to objects in the world. In the case of the concept CAT, a category of things, the objects in the world form a set. In the case of PUZZLE, a concept that applies strictly to an individual, there is only one *cat Puzzle* in the domain of reference in Figure 1.

At the level of the lexicon, the two expressions *cat* and *Puzzle* are treated uniformly: technically, they are individual constants, although not in the usual sense in that their domain of denotation is not the world, but the conceptual domain. The nature of what the labels denote only becomes relevant in the analysis after they are combined with functional content, that is, in context.

Returning to the moulding analogy, any label can be assigned to a generic entity, what I will call an atom. The basic mould of grammar can send a single atom, which can be represented as the dot in (6):

(6) •

When the combinational system constructs the proper noun moulding, it sends one atom in the discourse. For any label to be used in a felicitous manner on this atom, it will have to be the case that this label only relates to one individual in the world. In our little world, this will only be possible if the label used on this atom is *Puzzle*, as in (7a):

(7) a. •<sub>Puzzle</sub>  
 b. •<sub>cat</sub>

In (7a), a single atom labelled *Puzzle* is sent in the discourse: this can therefore be used to talk about the entity *Puzzle* in the world. On the other hand, if the label *cat* is associated with a single atom in the domain of reference in Figure 1, then there will be a problem,

because more than one object in the world bears the label *cat*. A label that can be applied to more than one entity cannot be used in this case<sup>6</sup>.

Because the term *cat* relates to a concept that applies to a set of things in the world, a felicitous use of *cat* requires some manipulation of the label by the combinational system. As I show below, the grammar can create a set of atoms, and the label would have to be duplicated so that it can be applied to this set. This would result in something like (8a), where the term *cat* has been pluralized to *cats* to provide labels for a set of atoms:

- (8) a. •*cat*    •*cat*    •*cat*    ...  
 b. •*Puzzle*    •*Puzzle*    •*Puzzle*    ...

The situation in (8a) can thus be used to talk about the set of things in the world that is such that each member of the set bears the label *cat*. In principle, nothing would prevent pluralizing the term *Puzzle*, and thus sending in the discourse a set of atoms that bear that label, as in (8b). However for this to be felicitous, there would need to be a set of entities in the world that have that name, which is not the case in the world of reference in Figure 1.

So even if the conceptual details of substantive terms ultimately serves to assess whether a sentence is felicitous or not, these details are not relevant at the input of the grammar. All substantive elements have identical values in the lexicon: they are all treated as individual constants with respect to a notion of the conceptual domain. Their value remains the same in all contexts of use: so, even if the grammar duplicates labels to allow reference to a set of things, the conceptual notion the term denotes does not change.

Let me now turn to a formalization of these ideas. Suppose that the mould that provides the means to send one atom in the discourse is represented as the set of brackets in (9), which I will call a *contour* (in earlier work, I used the term *position*).

- (9) ( )

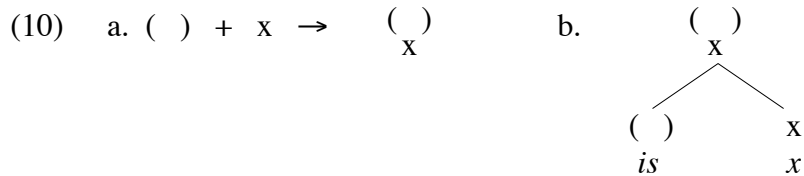
The contour is the basic functional content that occurs in argument-taking elements in grammar, such as verbs, prepositions, and verbal inflection. Verbal inflection is the contour associated with the tense marker and introduces the subject of the proposition. I argue at length in previous work that the content of *be* is only that: a contour associated with the tense marker.

The grammatical value of a term is a function of how its label is combined with a contour. Different rules of combination lead to different arrangements of labels and contours. An element that identifies a contour is a “noun”. By “identify”, I mean that it provides a label for the contour, indicating what label the atom bears as it is sent in the

---

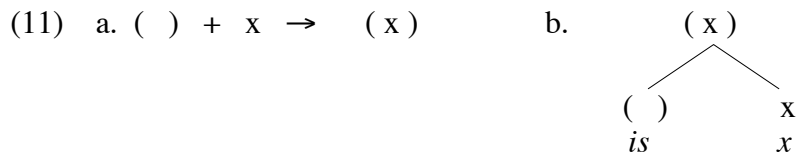
<sup>6</sup> This is not exactly correct. In a vocative context where I address my cat (*Oh cat! you ate the plants again!*) or a teacher addresses a class of pupils (*Class, today we are going to study the foundations of linguistic theory*), then the labels *cat* and *class* are used as designators of individuals. Notice however that such uses are possible only if the labels are applied to the individual *cat* or *class* in the immediate context. For example, these labels cannot be used as a means of conveying information to the hearer (*\*today, cat ate the flowers, \*the teacher taught class the meaning of life*). In this context, they are not associated with any mould, that is, they are not combined functional content in the sentence.

discourse. The basic case of identification relates *one* contour to *one* label *x*. This situation sends one atom that bears the label *x* in the discourse. Identification is formally expressed by stacking elements on top of each other in the representation. (10a) shows the combination of a contour and a label *x* by *the rule of identification* (the order of the stacking in the representation is determined by the linear order of the elements in the sentence; for example, the *x* would sit on top of the contour if *x* was before the contour).



Leaving aside details of tense specification and agreement features, combining the expression *is x* by this rule results in (10b). The contour in (10b) is identified: it bears the label *x*, and *x* is thus a noun. Given that only one atom is sent in the discourse, and that it bears a label *x*, it will have to be the case that only one object in the world bears that label. A label that uniquely applies to an individual thing is a proper noun. In other words, here *x* is a proper noun. The result of the combination in (10) is exemplified in *this is Puzzle*, and corresponds to the case of the proper noun illustrated in (7a).<sup>7</sup>

Whereas a proper noun results from the stacking of a contour and a label an adjectival value for a term arises from inclusion of a label in the contour. The application of the *rule of inclusion* is depicted in (11a), with a derivation for *is x* in (11b):



Unlike in (10), the *x* in (11) does not label the contour at the level of the combination. To be well formed, the contour must be identified, so another label is required somewhere in the sentence to identify the contour (in this case, it will be the subject). The term that labels the contour would then be associated with a contour that includes *x*, restricting its value accordingly. In this configuration, the label is like a filter placed on the contour. The fact that terms interpreted as adjectives require an argument follows from the hypothesis that an adjective is a label that is included in something else (adjectives internal to the NP are discussed at the very end of the section): to be included, they need a recipient, that is, something to be included in. The adjective interpretative cannot be realized without this recipient.

Since grammatical category does not exist at the terminal level, but is a function of how contents are arranged at the level of the combination, different values can be associated with a term without changing its basic terminal content. In other words, the *x*


<sup>7</sup> The fact that a single label can only relate to one atom can be seen as a consequence of the nature of a terminal expression. Such a term is an expression that cannot be broken down into smaller meaningful units: it is a unique formal object in the analysis. The atomic nature of the moulding constructed with a single term is thus not accidental; it has its source in the very nature of what is the term in the analysis.

in *is x* is potentially ambiguous between a proper noun reading and an adjectival reading. A term that can occur in these two configurations is *red* in (12):

- (12) a. Mary's favourite car *is red*.  
 b. Mary's favourite colour *is red*.

Given that *red* in (12a) can only be interpreted as a property of the subject *Mary's favourite car*, the term has to be combined as an adjective in this case (it is combined with *be* by inclusion, just like the *x* in (11)). In contrast, *red* in (12b) identifies the name of *Mary's favourite colour*. It is thus used as a proper noun, and is combined by identification, just as the *x* in (10). Notice that under this analysis of the contrast in (12), both *is* and *red* have the same terminal content in the two contexts, and *red* is labelling the same concept whether it appears as an adjective or as a noun.

The proper noun and adjectival readings are not the only values that can be associated with a singular substantive term in English: a mass noun could also appear as the *x* in *is x*. The mass reading relates to the substance atoms are made of. To arrive at this, I suppose that at the level of the combination, the contour is narrower than the label. The grammar labels the contour in such a way that only a subpart of the information the label denotes is associated with the contour. Formally, I assume that this is achieved by duplicating the label as it is combined, including the contour in the duplicated labels in the result. The rule of *inclusion by duplication* in (13a) shows this:

- (13) a. ( ) + x → x( )x                      b. 

In the result of this rule applied to *is x* in (13b), the region occupied by the label is wider than the contour. This configuration is associated with the mass noun reading of a term, as in *this is chicken*. What is sent in the domain of discourse is not an atomic (individual) chicken, nor is it a set of chickens. A set, as I discuss below, requires separation of the duplicated label (a empty space between two identical labels). In (13), there is a contour between the two labels: what relates with the contour is therefore only a subpart of whatever concept is associated with the label.<sup>8</sup>

Under the definition of noun provided in the discussion of (10), the *x* in (13) should not be a noun, because it is not combined by the rule of identification with the contour. To account for the nominal value of mass interpretation, I assume that a contour included in a duplicated label is labelled indirectly, that is, by inference. The following inference rule, which applies to other cases discussed below, ensures this:

<sup>8</sup> Evidently, as the literature on the subject shows, what 'subpart of' means is not a one-size-fits-all-notion in the conceptual domain. It depends on the denotation of the term combined with the contour. For example, with a homogenous substance like *water*, a subpart is still water. But because the term *chicken* denotes a kind of animal, which is not a homogenous notion, a subpart of a chicken is not another chicken. But as far as grammar is concerned, it would treated all these cases uniformly, by providing a contour for a subpart of whatever the conceptual notion is.

- (14) In a sequence  $\alpha \beta \alpha$  where  $\alpha$  is a label and  $\beta$  is an operator (a functional term), then  $\beta$  is labelled  $\alpha$ .

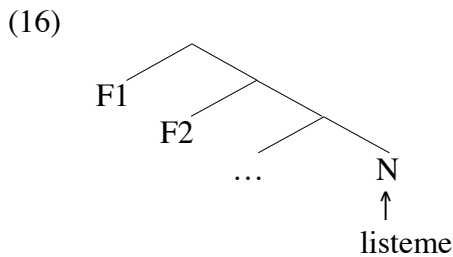
Under (14), the term in (13)  $x$  is a noun: it provides the label associated with the contour.

With these rules, it is possible to assign three values to the  $x$  in expressions of the form *is x* (as in *this is huge*, *this is George* and *this is chicken*). The three semantic representations assigned to this expression, depicted in (15), all arise from unique descriptions of terms, and therefore respects the terminal criterion.

- (15) a.  $\left( \begin{array}{c} \phantom{x} \\ x \end{array} \right)$       b.  $(x)$       c.  $x(\phantom{x})x$

The rules of combination that account for the distinctions are not specific statements about the expression *is x*. They could equally apply to *is y*, *is z*, and so on. Indeed, the proposed rules are sufficiently general to account for other distinctions as well, a point to which I return shortly. First, I want to briefly compare this analysis to other syntactic and semantic approaches to noun interpretation. I discuss the analysis developed in Borer (2005), in which the description attributed to terms is explicitly separated from their semantic interpretation in grammatical context. I then show how the analysis proposed here differs from approach based on semantic types.

In line with the current Chomskyan framework, Borer assumes that a noun is a term (listeme in her terminology) inserted at the bottom of an array of certain functional categories. A simplified structure, where the functional categories are labeled F1, F2, ..., appears in (16).



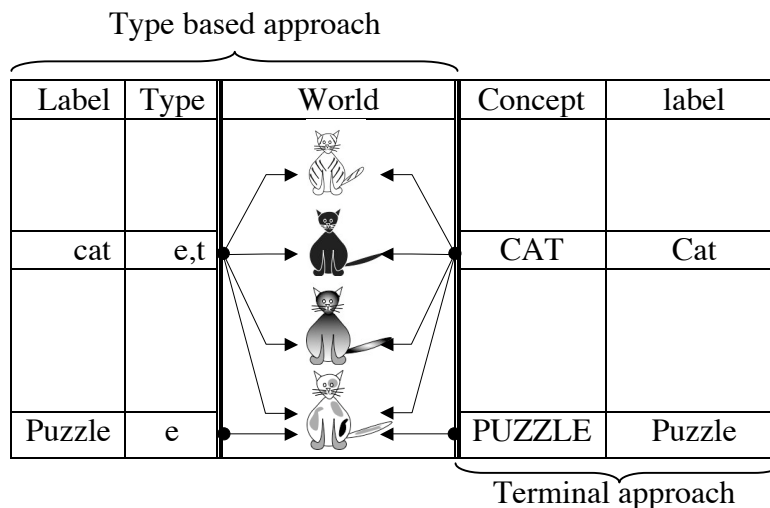
Unlike standard assumptions, however, the term in Borer’s analysis would not bear any formal feature that would be relevant to its insertion in this position. Any term can be inserted in this structure, regardless of its denotation, and thus receive the value N. As in my approach, the denotation of the term and the category “noun” are effectively separate (see also Baker 2003 and Marantz 2000 for similar proposals).

Although in this account, the value noun is determined in the configurational component of grammar—the combinational component, i.e., the syntax—it is nevertheless based on a specific statement at the terminal level of the analysis. To illustrate this point consider Borer’s analysis of proper nouns and mass nouns. Although identical in form (singular bare terms), the interpretations “proper noun” and “mass noun” in her analysis depend on the presence of certain features in the functional projections. The proper noun reading of a term in Borer’s analysis emerges from the presence of a specific abstract feature in the highest terminal node of the structure (in this case F1), usually referred to as the D(eterminer) node. This feature is ‘realized’ by the

determiner *the* when it appears in the structure. In a proper noun configuration, however, no determiners are inserted. For the feature in D to be realized, the term inserted in the lowest node of the structure (that is, in N in (16)) has to move (covertly in English) to the position D. This is why, under this analysis, proper nouns have the definiteness and uniqueness presuppositions that are generally associated with definite NPs. In contrast, mass noun interpretation is assumed to be the basic value associated with the node N itself. Mass noun interpretation arises when there is neither a determiner nor a feature in the structure that would trigger the movement of N. The term is bare and must remain in the base position N, where it gets to have the mass reading that is associated with this position.

Notice that in Borer’s analysis, each reading is accounted for by statements that directly relate to distinct terminal nodes of the structural configuration: the mass noun reading is associated with the base position N, whereas the proper noun interpretation is associated with another terminal node higher in the structure, the node D. In other words, two distinct statements (one associated with the node D, and another, with the node N) are assumed to account for the semantic distinction. The two interpretations are therefore associated with specific statements made at the terminal level of the analysis. Even if no specific term is associated with these statements, they nevertheless need to be made somewhere in the grammar, and associated with terminal nodes of the analysis. As such, they cannot be reduced anything else at this level of the analysis, and are therefore incompatible with the terminal criterion.

Borer’s analysis can be opposed to more standard approaches in formal semantics, where terms relate to semantic types. Semantic types provide the means to relate labels and notions in the world (real or possible): type definition is dependent on the nature of the objects the expression denotes in the world. This means that under this approach, information about the world is used at the terminal level of grammar. For example, the type associated with the terms *cat* and *puzzle* depends on the ‘scope’ of these expressions in the world. This situation is depicted in the left part of Figure 2.



**Figure 2**

A term that relates to the type predicate  $\langle e, t \rangle$ , such as *cat*, in fact relates to a set of individuals in the world, whereas a term like *Puzzle*, that only relates to a specific individual, has a type  $\langle e \rangle$ . When types are used to describe terms, type-shifting rules then becomes necessary each time a term appears in a context with a type that is different than the one it has at the terminal level. Such manipulation does not occur in the terminal analysis, because the input of syntax is blind with respect to the association between the concept and the objects in the world it applies to. Thus, as the right part of Figure 2 illustrates, what feeds into the combinational system is only the relation between the label and the concept. Grammar then manipulates the label (by identification, inclusion, duplication, etc.) to produce moulds that can relate to distinctions that are relevant in the world: this is done without altering the basic value of the term (the concept it denotes). The label can then keep its function as a denotation of a concept, and which part of this concept is targeted in the sentence is dependent on how the label is combined with a contour in the combinational system.

To conclude this section, I show that the rules of combination just presented can be applied to a range of other constructions in the nominal domain. I will discuss the plural, indefinite and definite articles, and adjectival modification in the noun phrase. The plural under my approach amounts to sending more than one atom into the discourse. This requires two things: duplicating the label and separating each instance to allow their association with more than one atom, and providing a sufficient number of contours for duplicated labels. Pluralizing implies the duplication of a label, since a single label can only identify one contour and send one atom in the discourse. The resulting duplicated labels must also be separated from one another, since, as the discussion of mass terms showed, a duplicated label is not necessarily about discontinuous atoms. Only when a label  $x$  is duplicated *and* each instance of the label is separated from the other can the result be used to talk about a set of objects that bear that label.

In English, duplication and separation of the label is the function of the plural inflection on nouns. Plural inflection introduces what I call the *cut* filter. It duplicates a label and separates the duplicated instances by leaving an empty space between the two labels. Assuming that the plural is represented by the straight bar in (17a), the presence of the plural marker *s* on a term  $x$  leads to the representation in (17b):

- (17) a. |                      b. x x

I assume here that the plural operator does not appear in the result in (17b), since its function is to separate identical labels. By leaving a gap between the two  $x$ s, the labels are formally separated. This means the two  $x$ s in (17b) can relate to discontinuous elements. The result in (17b) must relate to a set of things in the world that bear the label  $x$ .<sup>9</sup> The label associated with a plural expression must therefore be associated with a concept that applies to more than one entity in the world. This is typical of count terms, such as *cat*, which naturally appear in the plural—*cats*. It can also be applied elsewhere (for example, plural applied to the name of an individual, as in *Andrews*, identifies the set of individual having that name).

---

<sup>9</sup> Although the representation only shows two  $x$ s, I assume that it relates to a whole set of labels  $x$  by inference. Unless restricted by other means—for example, a quantifier like *two* or *some*—a sequence of two separate labels  $x$  can relate to the whole set of things that bear the label  $x$ .

To send a set of atoms in the discourse requires a sufficient number of contours. Recall that when the contour is identified, as in (10), it sends only one atom (significantly so, given the hypothesis that this configuration is associated with the labelling of an individual). To send a set of atoms, more contours must be provided, a result obtained by duplication of the contour. Assuming that a distribution operation operates on one contour at a time to ensure an even distribution of labels and contours, the identification of duplicated contours with duplicated labels results in (18a):

$$(18) \quad \text{a. } \begin{array}{cc} ( ) & ( ) \\ x & x \end{array} \qquad \text{b. } \quad x ( ) ( ) x$$

This representation expresses that all the members of the set of atoms that bear the label  $x$  are sent in the discourse. Another combinational output that would be well formed under the approach is the inclusion of the duplicated contours in the duplicated and separated labels, as in (18b). With a slight modification to the inference rule in (14) to allow all operators included between identical labels to bear the label, the contours are then labelled  $x$  by inference. In this case, the plural functions like a filter on the output: the scope of the label  $x$  is wider than the set of atoms sent in the discourse. As with the mass reading only a subpart of what is labelled  $x$  is sent in the discourse; they are all atoms, and they all bear the label  $x$  (by inference). The difference between the two results in (18) can be applied to the difference between generic and referential readings of bare plurals: in (18a), the whole set of labels is associated with atoms, which amounts to sending in the discourse the whole set of atoms that bear the label  $x$ . This corresponds to a generic reading. In the case of (18b), the referential reading, only a subset of all labels is associated with atoms, implying the existence of other atoms that are not sent in the discourse.

Duplication of a contour, I should point out, is not morphologically signalled in English. That is, no verbs or prepositions—the terms whose function it is to provide contours in the sentence—show an inflectional marking that is strictly sensitive to the presence of plural inflection on the noun.<sup>10</sup> In English, the width and number of contours are determined by the needs of the context: if the contour is combined into a duplicated label in the mass interpretation, it will then have a narrow width; if it is identified with a unique label, it will have its basic atomic width; and in cases where it is combined with duplicated and separated labels, it will need to be duplicated.

Sending a unique member of a set in the discourse (rather than a set, as just discussed) is also a possibility. Recall that an atomic label  $x$  can only be applied to an individual atom. Thus, only one individual in the discourse that bears a label  $x$  can be identified with an atomic label. This precludes the use of the bare term *cat* in the singular to talk about an individual, given that this label can be applied to a set of objects. To

---

<sup>10</sup> Although the verb *be* changes from *is* to *are* in the plural, *are* is not specific to plural expressions (e.g. *you are a good person*). In contrast, in a language like French, there is an explicit manifestation of contour duplication on verbal endings. Thus, inflectional morphology on the verb signals that the subject is plural in 1<sup>st</sup> and 2<sup>nd</sup> person in all paradigms. Some morphemes might be specific to a verb, to a class of verbs or to tense-modality combinations, but specific markers within each paradigm appear for the 1<sup>st</sup> and 2<sup>nd</sup> person subject in the plural. The only verbal systematic morphology sensitive to contour value in English (the *-s* of 3<sup>rd</sup> person singular in the present) in fact indicates the absence of duplication.



isolate a member of this set, a determiner is necessary. Determiners can be interpreted as filters placed onto a contour to restrict the nature of the atom that is sent in the discourse. The argument of the determiner identifies the restriction, and when combined with a contour, the atom that is let through bears the label associated with this argument.

The indefinite determiner *a* can be analyzed as the set filter. It creates a set of labels *x* and reserves an opening in the set for one contour. Formally, this is captured in the following manner: the indefinite is a filter whose identification requires inclusion in a duplicated and separated label, in a way that is akin to the plural. The filter appears in the representation, reserving a region that is sufficiently wide to include a contour. Taking the curly brackets in (20a) as the formal object associated with *a*, (20b) is the result of combining this filter by inclusion in the duplicated label *x* (as in *an x*):

$$(20) \quad \text{a. } \{ \} \quad \text{b. } x\{ \}x$$

The two *x*s are discontinuous labels, therefore providing a set of labels. This set can be applied to the set of all objects that bear the label. Once a contour is included in the filter, as in (21), it is then identified as a label *x* by inference (rule (14)):

$$(21) \quad x\{ ( ) \}x$$

The configuration that corresponds to *is an x* can then target one atom which bears the label (as in *this is a cat*). Notice how this analysis differs from a standard analysis based on semantic type. Under such an approach, the noun *cat* denotes a set (it is a predicate): determiners select a predicates as their argument, and isolate one member of the set. Under the analysis proposed here, *a* selects a single label (an individual constant) and *creates* a set of labels so that it can be applied to a set of objects. It then delimits a region in which one contour can be included: by inference, the contour is interpreted as having the label, and one atom bearing the label is sent in the discourse.

The definite article *the* can be analyzed as the atomic filter: it seeks out the single individual in the discourse that has the label *x*. Formally, *the* can be represented as an atomic variable, which I represented as  $\alpha$ . The argument of the determiner provides a restriction on the variable. Representing variable identification by a subscript on variables, the expression *the x* results from the combination of the variable by the rule of identification with the term *x* in (22a):

$$(22) \quad \text{a. } \alpha_x \quad \text{b. } ( )_{\alpha_x}$$

When this result is combined by identification with a contour, as in (22b), the variable is treated in a way similar to that of a proper name: it identifies a single atom  $\alpha$ . However, since that atom is a variable, it needs an antecedent in the context, an antecedent that happens to be restricted by the label *x*. For a felicitous use, then, there will need to be one entity that bears the label *x* in the discourse (as in *this is the cat*).

I conclude this section with a brief mention of a last case where these operations apply in the nominal domain, the case of adjectival modification inside the noun phrase. Adjectives in verbal contexts, as I briefly mentioned in the discussion of (11), are also filters; they can be interpreted as narrow filters. They restrict the value of an atom by

inclusion in a contour. This operation strips them of their function as labels, making their interpretation dependent on another label in the context. In a construction like *the cat is grey*, the adjective is the essence of the informational content of the proposition, since it applies to the contour that is identified by the subject of the sentence: the proposition informs the hearer about a property of the cat. In a nominal phrase, however, adjectives restrict the nature of the label itself, creating a complex label that is applied to the atoms sent in the discourse. For a sequence of terms  $x y$  such *grey cat*, the adjectival value obtains by duplicating and including the  $x$  inside the duplicated label  $y$  (i.e.,  $yxy$ ). The representations in (23a) and (23b) are the outcome of combining this result with the indefinite and definite determiners respectively.

$$(23) \quad \text{a. } \lambda x y x y \{ \} y x y \quad \text{b. } \alpha_{x y x}$$

Representation (23a) is a filter that can identify a member of the set of atom  $x$ s that have the propriety  $y$  (*a grey cat*), whereas (23b) is the filter that looks for the unique individual  $x$  with the propriety  $y$  (*the grey cat*) in the discourse.<sup>11</sup>

This concludes this brief overview of an account of some basic semantic and formal properties of nominal expressions in the grammar of English, which at the same time summarizes my most recent research. My goal here has been to highlight the explanatory power of the approach, rather than seeking to test predictions that it makes. As the discussion has made clear, the theory has tremendous explanatory potential: very few descriptions (that is, terminal statements) are required to account for many semantic distinctions; furthermore, the rules and conditions that account for these distinctions have a broad range of application. Although these results are based on a very small fragment of facts, I believe they nevertheless provide a good measure of the potential of the theory. I shall now turn to the last component of my research program, the computational model of the theory.

## A Meaning Engine

Although the previous presentation has been informal, the overall framework presented clearly lends itself to a rigorous formalization, and over the last three years or so, I have been collaborating with Robert Mercer to develop a computational model of the theory. I will now briefly present the current state of this model, before discussing what I intend to use it for. This provides an overview of the future direction of my research program.

---

<sup>11</sup> In Lamarche (in preparation), I incorporate the insights of Bouchard's (2002) analysis of adjectival modification, arguing that what he calls the *subset* interpretation of certain prenominal adjectives (which would yield the reading *old as a friend* for *old friend*) is actually a case of identification (stacking) of the term (the adjective) with a noun. This is not a true adjectival "configuration" (the result of the inclusion of a denotation). Rather, the adjectival interpretation arises from the denotation of the term (that of a property). The less-defined type of modification these adjectives allow (they have a wide range of interpretations, as Bouchard eloquently demonstrates) follows from the less-specific relation the noun establishes with this "pseudo"-adjective: when a term is included, it no longer has a labelling function, being completely dependent on another label. Prenominal terms that are not included still act as labels, and how they are interpreted with respect to the noun is dependent on the interaction of their respective denotations.

In July of 2003, Professor Mercer and I were awarded a small grant from the Academic Development Fund (ADF) of The University of Western Ontario to start the development of a computational model of this approach to grammatical analysis. A part of this grant was used to develop the templates that would support the construction of the semantic representations. The template developed is a two-dimensional grid such as the one in (24), which expands horizontally towards the sides from a central (zero) value, and vertically toward the bottom (see White, 2003).

(24)

	...	3	2	1	0	1	2	3	...
a									
b									
⋮									

The values associated with the different symbols for each term dictate how the information is to be placed on the grid. As the information is fed to the system, representations such as the ones in (25) can be constructed out of simple rules that determine the possible embedding and stacking of information.

(25)

	...	3	2	1	0	1	2	3	...
a					x				
b		y	{	(	)	}	y		
⋮									

Once the grid-template was established, we were free to focus our efforts on determining the nature of the objects that would appear on it. Before developing the computational model proper, we felt it was important to establish what kind of objects would be manipulated by the system and how they would interact. In many respects, the overall design of the nominal system presented in the previous section is the natural outcome of the numerous discussions Professor Mercer and I have had over the last two years.

The theory is now described in terms that can be instantiated in an actual program to be outlined in Lamarche and Mercer (in preparation). The remainder of our ADF grant will pay for the writing of code for a program that will instantiate the theory of contours labels, and filters. The first version of the model will be written in the programming language PROLOG, which is sufficiently flexible to test a program as it is developed. Once we have a version that satisfies the needs of the theory, more definitive versions will be written in a more efficient language, such as C or any of its modern variants. Professor Mercer and I intend to apply for a SSHRC grant in the fall of 2005, which would in part support this project.

Rather than expanding on the technical aspects of the computational implementation, I want to show how such a tool can be used to develop and test the theory. The computational model should be such that it is capable of manipulating parameters and testing the consequences of adding new specific and general statements to

the analysis. To illustrate what I have in mind, suppose the statements in (26) are added to the hypotheses described previously (that is, the values and widths assigned to the terms discussed, the combinational rules, the inference rules and the conditions on the well-formedness of contours and filters):

- (26) a. Directionality parameter  
Morphological operators (plural and tense) appear after their argument;  
Syntactic operators (contours and determiners) precede their argument.
- b. Precedence condition  
In a sequence of adjacent operators  $\alpha \beta$ ,  $\beta$  can only be combined with  $\alpha$  once  $\beta$  has an argument.
- c. Constraint on expansion  
All labelling information presented after the contour that bears the tense marker expands the representation outside of the limits of this contour.

The directionality parameter, set for English in (26a), governs the order of the relation between functional terms and their arguments. The condition in (26b) is stated to respect the constituent structure of V+NP constructions, ensuring that in the expression *is a x*, the filter *a* is combined with *x* before the result is combined with *is*. The constraint in (26c) is necessary to accurately describe the relations between the different arguments and filters in copular constructions.<sup>12</sup>

The theory thus constrained derives the set of representations in (27), which can be associated with the copular expressions that results from the combination of *x*, *y*, *a* and *is*:

(27)	Input		Output	Operations	Labeled expressions
a.	$x \text{ is } y$ $x \text{ ( ) } y$	i.	$x$ $(y)$	1. $x \text{ ID ( )}$ 2. $y \text{ IN in R1}$	Mary is tall
		ii.	$x$ $( )$ $y$	1. $x \text{ ID ( )}$ 2. $y \text{ ID R1}$	Mary is Wonderwoman
		iii.	$x$ $y( )y$	1. $x \text{ ID ( )}$ 2. $R1 \text{ IN-DU } y$	Garfield is trouble
		iv.	$x(y)x$	1. $( ) \text{ IN-DU } x$ 2. $y \text{ IN in R1}$	Beer is good
		v.	$yx( )xy$	1. $( ) \text{ IN-DU } x$ 2. $R1 \text{ IN-DU } y$	Beer is alcohol
b.	$x \text{ is a } y$ $x \text{ ( ) } \{ \} y$	i.	$x$ $y\{ ( ) \}y$	1. $x \text{ ID ( )}$ 2. $\{ \} \text{ IN-DU } y$ 3. $R1 \text{ IN R2}$	Puzzle is a cat.

<sup>12</sup> Notice that the constraint is stated on *labelling* information, that is, nominal elements: it would not apply to the case of adjective, which can be included inside of a contour even if they appear after the inflectional node. I believe that what underlies this constraint is likely to be highly significant for this theory; however, I shall not address this issue here, and have only stated the constraint for the purpose of the current discussion. See also note 15 for some further discussion.

		ii.	$y\{x( )x\}y$	1. ( ) IN-DU x 2. { } IN-DU y 3. R1 IN R2	Gold is a metal.
c.	a x is y { } x ( ) y	i.	$x\{(y)\}x$	1. { } IN-DU x 2. ( ) IN-DU R1 3. y IN R2	A cat is cuddly.
		ii.	$yx\{( )\}xy$	1. { } IN-DU x 2. ( ) IN R1 3. R2 IN-DU y	A cat is fur. <sup>13</sup>
d.	a x is a y { } x ( ) { } y		$y\{x\{( )\}x\}y$	1. { } IN-DU x 2. ( ) IN in R1 3. { } IN-DU y 4. R2 IN R3	A cat is an animal.

The input, in the first column, is the description assigned to each term of the expression. Each input in (27a-d) has one or more output, which appears in the second column: (27a) has 5 (i, ii, iii, etc.), (27b) has 2, and so on. These outputs are the representations generated by the application of the operations summarized in the third column, while the last column adds actual labels to the expressions, that is, the atomic terms  $x$  and  $y$  are replaced by words of English.<sup>14</sup> The denotation of the labels is such that the expressions can receive truth-values.<sup>15</sup>

The reader can establish that these representations do follow from the premises. The important point here is that a combinational component that introduces semantic distinctions is capable of generating a range of representations out of very few descriptive statements. So, out of the four descriptions attributed to  $x$ ,  $y$ ,  $a$ , and  $is$ , the combinational component generates ten distinct copular constructions. Furthermore, the system generates representations that provide correct interpretations for sentences of English. With few specific descriptions and general statements, the analysis is thus capable of

<sup>13</sup> The fact that a sentence like *a cat is fur* is possible (albeit, not your run-of-the-mill sentence) indicates that although contours are sensitive to the width of filters, other filters are not. This sentence states that an individual that belongs to the set of *cats* is a subpart of the substance *fur*. To have a mass value in this context, the inclusion of *a cat is* in the duplicated labeled *fur* cannot have the effect of separating the label *fur* (otherwise, it would not be a mass term).

<sup>14</sup> The numbered lines in column 3 read the following way: 1.  $x$  ID ( ) reads **operation 1 is:** *x is combined by the rule of identification with ( )*; 2.  $y$  IN in R1 reads **operation 2 is:** *y is included (by the rule of inclusion) in the result of the combination on line 1*; R1 IN-DU  $y$  reads **operation 2 is:** *result of line 1 is included in duplicated y*.

<sup>15</sup> The effect of the constraint in (26c) is illustrated in the examples in (27a-5) and (27d). The reader will observe that the subject must be combined with the verb before the complement if the labeled sentences are to have truth-values. Thus, *beer is alcohol* and *a cat is an animal* receive truth-values, but *alcohol is beer* and *an animal is a cat* do not (at least, not in the normal world: for example, if alcohol was only found in beer, and the only animals left on earth were cats, then these sentences would receive a truth-value). The fact that the order of the derivation is the reverse of what is generally assumed in standard analysis of subject-verb-complement structures (where the complement is combined with the verb before, not after, the subject) is a consequence of having *be* as the only verb in the expression. Under a terminal description, *be* is only the inflectional node of the sentence, and has no independent root meaning. With any other verb of English, the root would support the construction of a representation that is independent of inflection. The information introduced by the complement structure of the sentence can then relate to the root meaning independently of the subject-inflection complex, accounting for the standard embedding of VP inside of IP.

making predictions about quite a few expressions of English. Suppose now that we start to add more functional distinctions to the grammar. What would happen, for example, if we assumed that the tense marker is a kind of contour? How would it relate to the rest of the content of the proposition? What happens if the direction of operator–argument relations is reversed for nominal filters? With a computer model, such hypotheses could be implemented easily, and the interaction between the different statements could be tested quickly. This is a rather exciting perspective for the development of this theory, and I strongly believe that it would provide an invaluable research tool for the study of language typology.

Let me briefly discuss some of the empirical and theoretical questions I intend to address with this computational model. The most obvious and immediate goal of the research is to develop the formal apparatus that is to be associated with functional content, so that a larger vocabulary can be described and fed to the combinational system. Since it takes very little content to generate quite a few distinctions, I believe that a fairly limited set of notions will emerge from this research. This task will evidently need to be pursued by studying more than one language, since not all languages make the same distinctions, nor do they use the same formal means to make distinctions (English, French, and Spanish are to be my first targets). In this area, I believe the essential question to be addressed is how and where labels and contours meet in a language. To illustrate what I mean by this, consider the fact that in French, the mass noun reading requires a partitive determiner (*du, de la*, etc.) in argument position. A simple account of this follows from the assumption that mass reading requires that a label and a contour meet at the *point of inflection*, that is, where the grammatical number is marked in the sentence. In a sense, this amounts to saying that inclusion of the contour inside a duplicated label is only possible in a local relation. Adopting the hypothesis in Bouchard (2002) (see also, Lamarche 1991) that grammatical number appears on the noun in English, but on the determiner in French, the presence of the partitive determiner in French follows. The mass reading is possible with just the noun in English, because the label sits directly on the contour. In French however, the point of inflection is not on the label, but on whatever determiner precedes it. In other words, a label such as *vin* (wine) cannot relate directly with a contour in French, because it is the determiner that sits on the point of inflection. The partitive provides a bridge so that the two elements can meet (the details are discussed in Lamarche, in preparation).

I believe that fundamental issues of cross-linguistic variation will involve determining where the points of inflection are in a language, and how the language signals these points.<sup>16</sup> A significant question that should be addressed, once a range of data in various languages is described, is how the forms and functional objects are associated to one another. I am inclined to think that there is some fairly principled manner in which these associations are established, which is dependent on the number of

---

<sup>16</sup> I have been toying for many years now with the idea that there is a principle in grammar (call it *The Exception that Confirms the Rule Principle*), which holds that important things are signaled in language by unexpected formal manifestation in an otherwise general paradigm. The ECRP naturally applies to the identification of the points of inflection in the sentence. For example, the suppletive nature of the copula paradigm in English (*am, is, are, was, were*) happens exactly at the point where morphological marking occurs on verbs (the –s of 3<sup>rd</sup> person singular in the present, and the –ed of the past): under the ECRP, this suggests that the copula is sitting on something important—the verbal point of inflection in English.

formal distinctions available in a language. The more formal distinctions available in a given language, the more semantic distinctions it can express. Considering that the different distinctions expressed on verbal and nominal morphology are made to fit each other, I expect that fairly general statements should handle the association of semantic and morphological objects.

Finally, given the function of the grammar as a means to transmit information, the system will need to be set in a model of the context where it operates. Although contours, labels and filters can operate autonomously, their function is to carry information about the world. This part of the research amounts to making hypotheses about the nature of the carriers of the grammar (that is, the speaker and hearer), how they internally organize the information their grammar conveys, and the nature of the world they live in. I believe that a computational model of the grammar will prove an invaluable tool to test hypotheses about the organization of speaker's knowledge of the world. Thus, under the null hypothesis that the organization of the different pieces of information in semantic representations is isomorphic with the organization of this information in the brain, then the semantic representations of sentences with truth-values are a direct reflection of the organization of cognitive information in humans.

## **Conclusion**

I have been pursuing for some time the idea that the content of terms should not make reference to contextual information, with the conviction that the fundamental questions of linguistic theory would not be solved by looking at more and more complex constructions of more and more different languages, but rather by an adequate understanding of the basic atoms of language. I believe that my research is at a point where I can foresee the correctness of this conviction. Perhaps the most compelling reason to make this claim lies in the observation that the foundations of the two main schools of thoughts in contemporary linguistics converge in the terminal analysis, which provides evidence that they are both correct in their basic assumptions.

Under the influence of Chomsky, who pioneered the school of formal linguistics, many linguists have assumed that the study of language can be restricted to its formal manifestations. In other words, it is possible, in fact some would say necessary, to study linguistic expressions without making reference to other external realities (e.g. general cognition, the function of language, etc.). Many other linguists have never accepted this position, and have in fact argued for the opposite. For the school of functional linguistics, it is the function of language that determines its formal manifestation. This position makes the formal study of language meaningless.

I would suggest that these schools each begin with a sound premise, but go wrong on the nature of the basic descriptions they assume in their account of natural language. As the field stands today, no grammatical framework I am aware of, formal or functional, respects the terminal criterion. However, with analyses that respect the terminal criterion such as the ones proposed here, I see arguments that support both formal and functional linguistics. Thus, there is definitely a formal reality to language, which I have shown can be studied without making reference to factors outside of grammar (the denotation of terms). At the same time, the fact that language is made to talk about the world is significant in its design. The notion of contour is a formal means of sending information

to a hearer. Furthermore, the packaging that grammar provides to fulfil its function is specifically designed to convey distinctions whose ultimate reality lies outside of grammatical knowledge. There is thus convergence of both schools of linguistics in the analysis, providing what I would argue is the strongest theoretical and conceptual support of its explanatory adequacy.

## References

- Baker, Mark. 2003. *Lexical categories: Verbs, nouns and adjectives*. Cambridge: Cambridge University Press.
- Borer, Hagit. 2005. *Structuring sense, volume 1: In name only*. Oxford: Oxford University Press.
- Bouchard, Denis. 2002. *Adjectives, number, and interfaces: why languages vary*. Amsterdam: Elsevier, 2002.
- Croft, William. 1991. *Syntactic categories and grammatical relations: The cognitive organization of information*. Chicago: University of Chicago Press.
- Lamarque, Jacques. In preparation. Contours, Labels and Filters: the Logic of Natural Language Semantics. Working title.
- Lamarque, Jacques. 2005a. Terminal Description and Explanation in Linguistics. To appear in *Linguistica Atlantica*, 25, Spring 2005.
- Lamarque, Jacques. 2005b. Construction du sens en grammaire. To appear in In Evrard, I., D. Bouchard, and E. Vocaj (eds.), *Représentations du sens linguistique II*, Brussels: Duculot.
- Lamarque, Jacques. 2005c. "Filtres grammaticaux et substantifs nus en français et en anglais," paper to be presented at a *Workshop on Bare Nouns*, to be held at the University of Western Ontario, London, Ontario, May, 2005.
- Lamarque, Jacques. 2005d. "Atome grammatical," paper to be presented at the 35<sup>th</sup> *Annual Meeting of the Canadian Linguistics Association*, University of Western Ontario, London, Ontario, June 2005.
- Lamarque, Jacques. 2003. Be the One. In M. Weisgerber (ed.) *Proceedings of the Conference "sub7 -- Sinn und Bedeutung"*, 7th Annual Meeting of the *Gesellschaft für Semantik*, Konstanz, Germany: Arbeitspapiere des Fachbereichs Sprachwissenschaft 114, 184-198.
- Lamarque, Jacques. 2002a. Invariance et contingence. *Travaux de Linguistique*, 45(2): 15-44.
- Lamarque, Jacques. 2002b. Sémantique linguistique en syntaxe. In Lagorgette, D. and P. Larrivée (eds.), *Représentations du sens linguistique*, Munich: Lincom Europa, 233-253.
- Lamarque, Jacques. 2002c. Raising and Controlling Ambiguities. Proceedings from the 32<sup>nd</sup> *Annual Meeting of the Canadian Linguistics Association*, Université du Québec à Montréal, 193-207.



- Lamarche, Jacques. 1998a. *Terminalité en grammaire*. Doctoral Dissertation, Montreal, Université du Québec à Montréal.
- Lamarche, Jacques. 1998b. Etre ou ne pas être une entité. Proceedings from the 28<sup>th</sup> *Annual Meeting of the Canadian Linguistics Association*, Cahiers Linguistiques d'Ottawa, 227-238.
- Lamarche, Jacques. 1991. Sur la morphosyntaxe du déterminant en français. *Proceedings from the 21<sup>st</sup> Annual Meeting of the Canadian Linguistics Association*, Toronto Working Papers in Linguistics, 145-158.
- Lamarche, Jacques and Robert Mercer. In preparation. A Computation Model of a Semantic Combinational System.
- Marantz, Alec. 2000. Reconstructing the lexical domain with a single generative engine. Unpublished ms., MIT, Cambridge, MA.
- Partee, Barbara. 1987. Noun phrase interpretation and type-shifting principles. In *Studies in discourse representation theory and the theory of generalized quantifiers*, eds. Jeroen Groenendijk, Dick de Jongh and Martin Stokhof. 115-143. Dordrecht: Foris Publications.
- White, Barbara. 2003. *Une grille sémantique: Modèle computationnel de la construction du sens en anglais*, M.A. thesis. London, Ontario, The University of Western Ontario.