DISCOURSE INFORMATION GRAMMAR AND EXPLANATORY ADEQUACY

Alexandre Sévigny
McMaster University

Abstract

This paper examines the concept of explanatory adequacy and addresses some of the problems triggered by rapid theoretical paradigm shifts that occurred in linguistics in the twentieth century. Concurrently, it also examines and comments on some of the difficulties which have inevitably arisen due to a noisy terminological landscape that makes it difficult to formulate, let alone evaluate, theories and hypotheses for their explanatory force. A concept of languaging is introduced which encompasses the information-based approach to linguistics that underlies Discourse Information Grammar, a cognitive pragmatically-based and usage-based approach to the scientific study of language and cognition.

1. INTRODUCTION

Chomsky's work has been remarkable and revolutionary for many different reasons, the most important of which is the cognitive revolution that his work has triggered in the social scientific, scientific and humanistic disciplines focused on language. The introduction of a cognitive approach to the study of language has had many implications, among which are the opening up of empirical studies using formal and semi-formal models. Moreover, given that semi-formal and formal models are based upon increasingly precise characterizations, coupled with rigorous demands placed on definitions and analyses of results and hypotheses, the model-based approach is forcing researchers to re-examine many concepts which have traditionally been either assumed or simply glossed over. Of course, even empirical methods can differ or indeed clash, but providing all efforts are sincere and consistent, such differences are welcome since they enable us to engage in verifiable, meaningful debates concerning the foundations of the theory, its first principles and its explanatory power. Advances in non-invasive brain imaging technology have made it possible to study language cognition in a more direct fashion, permitting researchers to better link linguistic behaviours to their biological (neural) substrate.

Hornstein (1998: 1) summarized Chomsky's work on the problem of adequacy in grammar when he wrote:

©LINGUISTICA *atlantica* 25 (2004) 143–164

> Chomsky has fundamentally restructured grammatical research. Due to his work the central object of study in linguistics is "the language faculty", a postulated mental organ that is dedicated to acquiring linguistic knowledge[1] and is involved in various aspects of language-use, including the production and understanding of utterances. The aim of linguistic theory is to describe the initial state of this faculty and how it changes with exposure to linguistic data.

As Anderson and Lightfoot (2002) point out, the shift which Chomsky triggered gave linguistics and linguistic theory a new definition, a new focus, new goals and new components by re-orienting fundamental linguistic activity from the structuralist preoccupation with external and socially-derived E-language and its various concrete forms to a focus on internal, personalized I-language with a focus on biological and mentalist phenomena. In essence, Chomsky has claimed linguistics as part of biology and has taken the initiative to extract it from the humanities.[2] This conception fits the agenda of formal and/or computational approaches. Should this reorientation prove workable and/or correct, it will have contributed considerably to providing a working definition for the discipline of Linguistics,[3] a much needed definition that continues to be problematic. Predictably, not everyone is pleased with the new approach and goals, and yet, currently, much of the effort expended by linguistic theorists is focused on describing and defining the initial state of the language faculty or organ. Specifically, much effort is being put forth to begin to describe how the "initial state" of the language "organ" changes with exposure to *actual linguistic data*.[4] This effort has not been without its own difficulties. For instance, in 1981 Chomsky characterized the initial state of the language faculty in terms of a set of principles and parameters. Language acquisition was viewed as a process during which the possibility of setting these open parameter values on the basis of (incomplete) linguistic data was universally available to a child. The pre-parametrized initial state of the system was viewed as a Universal Grammar (UG): a super-recipe for concocting language-specific grammars. These grammars, in turn, constitute the actualized and concrete knowledge of particular languages that result as parametric values become fixed. In 1995, the Minimalist Program simplified the approach to the problem of language acquisition and the development of specific,

[1]One of the reasons for writing this paper is specifically to clarify what is referred to by this term. In sections 2 and 4, there are additional comments on the lack of a clear referent implied by this and other terms.

[2]Some would argue that he has spurred a redefinition of the humanities. This is to be seen in the large and growing literature of cognitive scientists working in fields that have traditionally been the purview of the humanities and social sciences, such as cognitive anthropology, literary criticism, sociology, communication studies, music and art cognition, etc.

[3]I use the big-L version to speak of the field as a coherent entity and movement here.

[4]See sections 2 and 4 for comments on the lack of a clear referent implied by this term.

feature-based, lexicalist grammars, but the overall goal of describing and accounting for the language organ remained unaltered.

Given these goals and views, linguistic theory has a double mission. First, it aims to "adequately" characterize the grammars (and hence the mental states) attained by native speakers. Theories are "descriptively adequate" if they attain this characterization with sufficient accuracy as to exclude utterances which fall outside the range of acceptable productions. In addition, linguistic theory also needs to address the problem of explaining how such grammatical competence is attained. An immediate by-product of explanatory adequacy is to generate a principled way of choosing between two grammars generated by the same data set.[5] To cite Chomsky (2000: 7):

> A genuine theory of language has to satisfy two conditions: "descriptive adequacy" and "explanatory adequacy". The grammar of a particular language satisfies the condition of descriptive adequacy insomuch as it gives a full and accurate account of the properties of the language, of what the speaker of the language knows. To satisfy the condition of explanatory adequacy, a theory of language must show how each particular language can be derived from a uniform initial state under the "boundary conditions" set by experience.

This approach to explanatory adequacy rests on an articulated theory of UG, and, in particular, a detailed theory of the general principles and open parameters that characterize the initial state of the language faculty, that is, "the set of mental habits and structures or patterns each (normal) person is presumed to receive genetically" (Hornstein 1998:1).

As mentioned above, Chomsky's research has further led to the postulate that each human possesses an I-language where "I" stands for "internal", "individual" and "intensional" (Chomsky 2002). As an indication of the current reorientation of linguistic theory, it is interesting to note that such traditionally important individual language communities as "Latin", "French" or "English" have been demoted, in the new linguistics, to mere collections of idiosyncrasies which Chomsky treats as epiphenomenal, at best objects which the I-language generically specifies. So far, we have introduced two questions which Chomsky's program explores: "What are the properties of language?" and "How might they be described?" Chomsky (2001) has raised a third question: "Why are language properties the way they are?" The answer to this last question entails going "beyond explanatory adequacy", as discussed in section 5.

Finally, I note that in his approach to language, Chomsky advocates using a rationalist approach and a conception of mind that can be modelled using mathematical constructs. One of his aims is to provide falsifiable models of language

[5] In this case, the data set in question is "primary linguistic data" which is the data children are exposed to and use in attaining their native grammars.

structure and operations that presumably will also reflect cognitive reality by capturing analogous processes and structures in the brain. This aspect of Chomsky's program is tantamount to requiring that linguistic theory seek to provide three general models (1):

(1) a. a model of language structure, including initial state phenomena

b. a model of language acquisition, including initial state interfaces with given data and initial state developments until linguistic developmental maturity occurs, evidently at or near puberty

c. a (partial) model of the brain's linguistic activities

The Chomskian program is exciting and ambitious notwithstanding many difficult problems still to be overcome. Among these, it is important to address the current reality that there are many instances where key terms are defined either insufficiently and/or inconsistently or defined in ways that do not appear to be convincingly falsifiable. This leads to methodological and theoretical inconsistencies that require clarification given the fundamental role that modelling has been assigned in the biolinguistic program. A few of these problematic terms will be dealt with briefly in the next section.

2. PROBLEMS WITH BASIC TERMS OF REFERENCE

First, we need to be more specific about the nature and representation of "actual linguistic data", "linguistic knowledge" and "primary linguistic data". It is also critical that we represent and define processes underlying such expressions as "production and understanding of utterances", including "initial state changes". It would be much easier if these concepts were self-evident, but such is not the case. If it were so, then we would know conclusively what the process of "languaging"[6] involves. However, given the range and variety of theoretical work currently underway, there does not seem to be much of a consensus on this matter. For instance, there is little consensus on the essential nature and use of language (2).

(2) Does language proper include all of the following:

a. phonetics

b. morphology

c. syntax

d. semantics

e. pragmatics

f. knowledge and information

g. neural and/or other cognitive processes

[6] I am using the term "languaging" as a convenient short form for the clumsy expressions "the process or processes which result in the product we refer to as language" and "the use of language and its result(s)".

Or does language include only a subset of this list, or possibly even a different list? Moreover, we may ask whether the interfaces between language per se and the cognitive/pragmatic contexts are to be included as part of language or just as an interlayer glue of some sort? These are difficult questions and the fact that there are, at present, no universal solutions to them indicates that linguists, just like other scientists, must deal with describing elusive phenomena. This is nothing new. For instance, physicists do not agree on what exactly constitutes or defines electricity (a discrete phenomenon? a wave? both?). Yet, science has managed to be able to predict a great deal of the behaviour and side-effects of electrical phenomena. To do so, it has had to formulate various (reasonably) consistent hypotheses and develop units of measurement, as well as adequate conceptual building blocks. It is interesting and relevant to this paper that these building blocks and units of measurement (volt, ampere, watt, etc.) are not electricity itself, but mere tools developed to help us understand the use of electricity and predict its behaviour. In linguistics, we are, at best, at the same level. Thus, such terms as "linguistic knowledge", "actual linguistic data", "linguistic data" and "primary linguistic data", "understanding and representation of utterances" need to be clearly formulated in terms of what they are intended to be: either bits of language itself or mere concepts used to help us understand the nature, functioning and limits of language. When we say that a child develops his/her grammar from "primary linguistic data", it becomes essential that everyone understand what is being referred to (3).

(3) Is the reference to:
   a. genetically hardwired cognitive processes?
   b. intuitive understanding of how language is stored and accessed?
   c. overlearned patterns stored in memory?
   d. innate cognitive abilities to store, retrieve, represent and transmit concepts? etc.

There have been many theories about the nature of language, each with a different viewpoint, each with the objective of rendering comprehensible what language is all about. Thus far, we have made few solid gains over tradition.[7] However, the last several decades have seen progress, especially when formal modelling techniques

[7]When we say tradition it is important to note that it is not only grammarians and linguists who have spoken intelligently about language. There has been input to the field of "languaging" from almost every human, social scientific and scientific discipline that is interested in unlocking the secrets to humankind's nature. It would be important and interesting to linguistics to draw a map of these various influences and contributions from across the disciplines and see what everyone has been saying, compare it, and then winnow out conflicting arguments. Who is to say that there are not ideas about language put forth in cognate fields such as philosophy, literature, communication and drama, to name a few, that could contribute to our modelling efforts. An example of how this disciplinary cross-pollination can work is the influence of the architect, Christopher Alexander's (1977) work in computer science and design. I am convinced that cognitive science, in its many manifestations, is a transdisciplinary means to achieve this end.

and approaches have been used because such models lend themselves to making falsifiable claims which can then be pushed systematically until they either break or survive. Among the various gains that have been made is the possibility of looking at language from different angles.

For instance, linguistic data could and have been conceived as patterns and (hierarchical) organization of structures that can be represented either as tree-structures or as labelled brackets that can then be mapped onto utterances. However, as Saussure pointed out, language also consists of content and contexts. This leads to the need to effect mappings between structures and content. Again, what is the content of language? What relationships obtain among form, content, expression and context? Is it possible to start not from form, but say, from content, or even from a (postulated) essential purpose? If so, what is the essential purpose of language (4)?

(4) a. to communicate with others
b. to store concepts and expressions
c. to talk to oneself
d. to express one's inner thoughts
e. none of the above: rather, language exists to enable any of the above, all of which are side-effects of "languaging"
f. none of the above

It is also possible to ask whether there are significant differences between the essential purpose of language as a cognitive capacity (Saussurean *langue* or possibly Chomskian "I-language") and those many concrete examples of individual languages used by people all over the world (Saussurean *parole* or possibly Chomskian "E-language"). If such differences exist, what exactly are they? What units or processes do we need to develop to arrive at a satisfactory description and explanation of these differences?

In an attempt to deal with some of these difficulties, we might try to identify the components of languaging. An immediate result was that languaging appears to consist of eleven generic activities (5).

(5) a. storing
b. representing
c. accessing
d. transmitting
e. receiving
f. monitoring
g. editing on-the-fly or updating
h. contextualizing
i. referencing
j. networking
k. interpreting

Viewed from this perspective, languaging is seen as primarily dealing with the construction of mental representations based on information. It is also possible to interpret the five commonly given purposes of language listed above in terms of information processing. For example, "talking to oneself" simply involves the sending and receiving of information with sender and receiver coinciding. This has a number of implications for linguistic theory, the most important of which deals with the fact that the theory of syntax cannot be separated from the temporal frame of language production. Another implication is that the primacy of syntax in the cognitive system of language may have to be questioned. At least the conception of syntax as a system of declarative, unordered rules. Another important implication is that the link between language use and the theory of universal grammar may have to be reexamined. It could be that both Chomsky's theory of universal grammar and the time-linear processing models such as Dynamic Syntax or Discourse Information Grammar may both be correct. Perhaps the theory of the initial state can be considered separately from the theory of human langauge production and understanding necessary to understand language as a natural object existing in time and having a material instantiation in the brain. The next two sections will present a sketch of the approach used by Discourse Information Grammar (henceforth DIG) and indicate how such an information-based approach might contribute to some of the problems raised in this section.

3. Brief Overview of Discourse Information Grammar (DIG)

The practice of symbolic communication is a defining trait of *homo sapiens*. We do this by receiving signals which are processed sequentially, in real-time. The products of these processing activities are accumulations of information that cover the entire range of knowledge and meaning significant to us. When we considered these accumulations of information, four questions immediately suggested themselves (6).

(6) a. How do we effect these transmissions?
b. What is the nature of these transmissions?
c. What is the nature of the information being transmitted?
d. How do we learn to effect these transmissions?

Broadly speaking, there are three generic approaches that can be developed for any or all of these questions (7).

(7) a. the analysis of the prerequisites necessary for such transmissions (ante-transmission)
b. the analysis of the transmissions themselves (intra-transmission)
c. the analysis of the products qua products (post-transmission)

Finally, the approach adopted for any or all of these analyses can be formal, semi-formal (meaning that we adopt a system of representation that is not fully mathematically specified) or informal. The research results described in section 4 use a

semi-formal approach to analyze an example of the natural language understanding process (intra-transmission). Focus is placed on how the processes could be modelled and on what sorts of information are accumulated when the grammar is approached from an information-based time–linear perspective.

Recently, there has been a growing interest in the concept of procedural grammars that model "knowledge of language" from a left-to-right, functionalist, usage-based perspective: Dynamic Syntax (Kempson, Meyer-Viol and Gabbay 2001; Cann, Kempson and Marten 2005), LeftAssociative Grammar (Hausser 1999), Markov Grammar (Tugwell 1998), Axiomatic Grammar (Milward 1994), Linearized Phrase Structure Grammar (Shin 1987) and Discourse Information Grammar (Sévigny 2002a, 2002b, 2003). What distinguishes all of these approaches from the phrase structural (PSG) tradition used in most varieties of generative grammar is the underlying and guiding metaphor. PSGs are based on the metaphor that natural languages are formal languages and that there exists an autonomous syntactic module (largely) independent of semantics. In contrast, the time-linear approaches see a grammar of a language as a series of procedures permitting humans to construct partial representations as a sentence is being processed and understood or (re)-constructed. Thus, knowledge of language is knowledge of the processes and information necessary to understand and use the language. In the words of Tomasello (1998: xi):

> Many linguists and psychologists believe that there is a biological basis for language, just not in the form of an autonomous Generative Grammar. Just as plausible for these linguists is the hypothesis that language rests on more general biological predispositions, such as the abilities to create and learn symbols, to form concepts and categories, to process information rapidly, and to interact and communicate with other persons intersubjectively.

DIG attempts to create a framework within which cognitive scientists can frame and articulate linguistic theories that are empirically-based and psychologically plausible. It does not necessitate the *a priori* assumption of an independent language module in the mind, but does encourage hypotheses about the interplay of different types of information in the languaging process.

DIG accumulates and assembles information on the fly, in real-time. In order to do this, it relies on a specialized lexicon tailored to meet the needs of linear information accumulation. There are specialized lexical templates for nominals, verbs, linkers, etc. — even templates for structures of various types as well as functional roles. Lexical entries are specified in terms of attribute-feature pairs which may vary in degree of specification: *full specification*, for such "hard-wired" items as the French definite article *la*, which must be marked [+singular], [+feminine]; *partial specification*, needed to accumulate information on the fly; *underspecified specification*, which is the usual default specification. At times, underspecified becomes altered to *non-specified* in the case of an item which is simply left unclear when the discourse/utterance closes. Information accumulation proceeds word by

word as structures are built up;[8] completed structures are then functionalized by being assigned functional roles via a small set of operators. Eventually, a minimum discourse unit is built up, triggering the appearance of a local context with primitive roles such as agent, doer, verbal relationship, complements and various potential or realized links. As information accumulates, it is networked and semantic fields, topic chains, logical structure and types are marked as initiated, on-going or completed as the case may be. (Distant) anaphora are filled as soon as possible. Eventually, closures occur and discourse units are accumulated. These bring in additional parameters such as discourse type.

Throughout the process, there are constant checks for feature-value compatibility, link relationship updates, accumulated information updates, etc. The net result is that the information contained in fragments and sentences is built up and networked. At any stage, it is possible to obtain a "snapshot" of the information state. This represents a current informational state of affairs. In reality, this models our ability to be able to summarize information accumulated so far and be aware that it is not yet complete. There may be incomplete or even incorrect assumptions and links but these may be corrected as soon as relevant information becomes available. Again, this represents our ability to correct ourselves on the fly. Also important, especially given the severe limits of operating in real-time, is the concept of *anticipated information*. In DIG, anticipated information consists of projected potential situational and pattern developments. As such, it is based on experience, familiarity with the information being accumulated and with overlearned patterns stored in memory. At any given moment, this restricts the possible new developments to under a dozen possibilities in most situations. Recognition of anticipated information "nodes" reduces the amount of decision-making and consequently reduces the time needed to effect such decision-making, both of which are crucial for operating in real-time. Examples illustrating a number of these concepts are presented in the next section.

4. DIG AND UNIVERSAL GRAMMAR

Although developed independently from generative grammar, DIG addresses a number of concerns relevant to the questions raised by the Chomskian program, especially in the area of explanatory adequacy. It also proposes a partial solution to some problems which may be "beyond explanatory adequacy", though for different reasons due to the different focus of the approach.

While the focus in generative grammar has been on "the language faculty" more or less as conceived by Chomsky — that is, as a rationalist, mathematical model (and not as an empiricist model) with concentration on brain activities as re-

[8]Currently, DIG uses established results from other theories, notably phrase-structure grammar. It may well turn out, however, that these structures may be better analyzed and represented as structures of another type, probably centered on the notion of (localized) context.

flected by linguistic structures and mechanisms which will map onto brain activities and functions — DIG has developed as a result of trying to model the linear accumulation of information which occurs during natural language processing. This has required that the notion of "information" be dealt with in unambiguous, falsifiable terms. It is DIG's information-centric view that provides a bridge between DIG and Chomsky's interest in problems of explanation in Linguistics. Information is constant in both approaches to the study of languaging. If one accepts some version of the representational theory of mind, it is necessary for that mind to have information to represent to itself. The mind, whether it be modular, distributed, a vast storage-and-retrieval device, comprised of central and peripheral systems, or just one enormous multidimensional network of connections, or even if it is operating along quantum lines, requires information to work. All of the metaphors that cognitive scientists use to talk about the mind are information-based. Moreover, if this is true and the language of the mind, or mentalese, as some have called it, is the means that humans use to represent external reality in a set of internal states, then those states must themselves find some sort of metarepresentation. The metaphor of information and information processing carries serious and interesting research implications which pose challenges to the remnants of structuralism that still permeate linguistic theory. For, if we are indeed information-processing mammals, a symbolic species, so to speak, then the notion that our mind's functioning is structured in the same way that a machine is structured, that is to say that there are formulae or rules that are somehow primordial, would appear to be the result of looking for innatism in the wrong place. In an information-processing animal, the initial state may not be defined as a set of principles and parameters, but perhaps rather a set of informational predispositions, algorithms and learning/acquisition/storage strategies. If this were true, then formal syntax ceases to be the central concern of the theoretical linguist — syntax remains, but is simply one stream of information among many — and is replaced by the search for a highly abstract set of processing and storing strategies which form the essence of (a) universal grammar.

4.1. DIG and explanatory adequacy

To reiterate: the lexicon used by DIG is built from lexical entries which, in turn, consist of complex structures of attribute-feature pairs, each in various stages of feature specification. Example (8) shows a typical entry template for nominals (see Sévigny 2000, 2002a, 2003, for details):

(8) NAME: ⟨ ⟩
CATEGORY:
INDEX: gender [ ], number [ ], person [ ]
STRUCTURE-TYPE: [ ]
SEM: {...}

The NAME field simply refers to the written form of a word or the phonetic form in the case of spoken language. Although the concept is problematic, of what

exactly a "word" is, this difficulty will be glossed over in this paper and the traditional view of "word" will be used. Eventually, however, this problem must be addressed and solved in terms of formal representation criteria. The CATEGORY field refers to the part of speech to which a word is assigned. In DIG homophones such as "purchase" (verb) and "purchase" (nominal) are treated as different words. There are difficulties with this classification but given the way DIG works, this seems the most economical way of dealing with cross-category homophones. STRUCTURE-TYPE is treated as part of lexical information. This models our ability to know almost immediately whether a new input belongs to the current structure being built up or not. In other approaches, this is dealt with strictly as a syntactic phenomenon, but in DIG the dividing line between lexical and functional information becomes blurred, given that both aspects of languaging are information-based, rather than strictly rule-based. The lexical field INDEX must be viewed broadly: basic information pertaining to nominal-structures (ns). In many languages, this involves gender, number and person, but other categories are also possible such as animate, shareable, visible, class, etc. The field SEM has been denoted with curly brackets and as an open field to represent an indefinite, open field. There are several reasons for this: words are very fluid and can acquire significant alterations to their basic semantic field values especially in the case of figurative and/or imaginative language; also, it is not possible at present to specify exhaustively the semantic properties of most words. Perhaps it is not even necessary or desirable to do this, given universal human ability to play with the meaning(s) of words and to create new metaphorical meanings on the fly. In this instance, the optimal rule of specification seems to be Simon's (1996) notion of "satisficeability", which states that goals or objectives are satisficed when there is sufficient positive evidence for their resolution and little negative evidence, and unsaticeficeable when there is sufficient negative evidence and little positive support for their satisficeability. Thus satisficeability demands a view of language processing as the attainment of critical thresholds in the on-line activation of concepts represented as collections of information. A DIG-based theory is explanatorily adequate when it is capable of modelling the languaging ability and also how this languaging ability develops. The critical difference is that structure (i.e., syntax), represented in generative grammar as a series of mathematical abstractions distinct from semantics or pragmatics, is represented in DIG as just another information parameter interacting with other information parameters and not as a fundamental and underlying module separate from the other parts of the languaging faculty.

As we will see shortly, lexical information enters into all information accumulation processes. Let us begin with a simple utterance (9):

(9) My neighbour's cat ate our goldfish last night.

In DIG, utterances are processed word by word beginning from the first. Thus, we input "My". At this point, "my" is merely a token. It has no information attached

to it. The first process to be triggered is *lexicalization()*.[9] Briefly, this involves a comparison against a list of lexical entries until a matching entry is found.[10] Once the lexical entry is found, *lexicalization()* attaches information to the token. At this point, it becomes a word, or a lexicalized token, as represented in (10).

(10) NAME: ⟨my⟩
CATEGORY: adjective
INDEX: gender [ ], number [ ], person [+1st?]
STRUCTURE-TYPE: [ ] nominal-structure
HEAD?: –
SEM: {OWNERSHIP: speaker; ... }

So far, not much information has been accumulated. We know (and anticipate) that a nominal-structure has been initiated. We have virtually no indexical information since, in English, possessive adjectives are unspecified for gender and number. We also know that this word indicates some form of ownership on the part of the speaker/writer. We should note also that in DIG, within a nominal-structure, INDEX feature person[ ] refers to the HEAD of the nominal-structure. Hence, the tentatively specified person [+1st?] will likely be overridden once the head of the structure is processed. Psychologically, we are now in a position of anticipation: we expect either a modifier of some sort, or a nominal head.

Next, we input and process "neighbour's". For simplicity of exposition, I am disregarding homophonous readings of *spoken* neighbours (i.e., "neighbours", "neighbours"). This ambiguity would be resolved with the next input, 'cat', since *"my neighbours cat" is incorrect. Resuming, we lexicalize "neighbour's" and obtain the analysis shown in (11).

(11) NAME: ⟨neighbour's⟩
CATEGORY: common noun: concrete
INDEX: gender [ ], number [+sg], person [+3rd]
STRUCTURE-TYPE: nominal-structure
HEAD?: –
SEM: {[+object], [+human], [+animate], [+ownership], ... }

This is a simplification, of course, but there is enough here to illustrate how information is being accumulated. In DIG, a common operator casts the CATEGORY value to meet the needs of the structure being built up. In this case, the semantic feature [+ownership] requires that this common noun be recast as an adjective. This is not a universal feature but it is a common one in the grammar of English. Presumably, in the initial state, the child is predisposed to accepting any specification as normal. Once a specification becomes fixed, however, it usually becomes more difficult to override such specifications as in the case of learning a language

[9] A pair of parentheses () is used to indicate a process as opposed to a lexical entry.

[10] In this paper, we will ignore the problem of dealing with a non-occurring match.

with a (radically) different specification. In French, for instance, this type of CATEGORY casting is much less common. Another feature of DIG which cannot be fully described in a short paper is the concept of the TEMP() process. Basically, TEMP() collects all modifying features of a nominal-structure and attaches them to the HEAD of the structure as soon as it is processed. This feature-accumulating process is needed in order to account for such utterances as:

(12) We spent the night in an old, abandoned, rat-infested, stinking, musty, roof-dripping, disgusting shack ...

where the attributive features accumulated in the chain "old, abandoned, rat-infested, stinking, musty, roof-dripping, disgusting" are all "held onto" (by being stored in TEMP) and eventually tagged onto the HEAD "shack". Although languages differ in their mechanisms for doing feature accumulating and tagging, the process does seem to be a universal feature of human language and hence of human cognitive information processing.

At this stage, then, we have the following information accumulated (13):

(13) TEMP:
INDEX: gender [ ], number [+sg], person [+3rd]
STRUCTURE-TYPE: nominal-structure
SEM: {OWNERSHIP: speaker , [+object], [+human], [+animate], [+ownership], ... }

Note that the SEM field contains a combination of feature specifications for both "my" and "neighbour's". Notice also, that some of the INDEX fields have now become specified. If we were interrupted at this point and questioned, we could answer that the utterance concerns the speaker's/writer's neighbour and something or someone belonging to this particular neighbour.

Continuing, we input "cat" and lexicalize it to obtain the results presented in (14).

(14) NAME: ⟨cat⟩
CATEGORY: common noun: concrete
INDEX: gender [ ], number [+sg], person [+3rd]
STRUCTURE-TYPE: nominal-structure
HEAD?: +
SEM: {[+object], [+animate], [+animal], [+feline], ... }

At this point, through the process of unification, TEMP will attach its content to the word "cat" (15).

(15) NAME: ⟨cat⟩
CATEGORY: common noun: concrete
INDEX: gender [ ], number [+sg], person [+3rd]
STRUCTURE-TYPE: nominal-structure
HEAD?: +

SEM: {OWNERSHIP: speaker, [+object, +modifier], [+human, +modifier], [+animate, +modifier], [+ownership, +modifier], {[+object], [+animate], [+animal], [+feline], ... } }

It is important to realize that there is very little that is new pertaining to structural information since we are still within the initiated nominal-structure. We note, however, the changed status to [+modifier] of feature traits attached to "neighbour's". Although not indicated here, this change was triggered by the CAST() operator. There are two general possible paths normally expected at this point: (a) continued description in the form of, say, a relative clause or closure of this structure, as would be the case in a simple utterance fragment, or (b) the initiation of a new structure, probably a predicate structure. This reflects the common pattern situation that once we have a complete structure, we anticipate that it will have a function of some sort. Mentally, we are now wondering something like "What about [your] neighbour's cat?" or a bit more remotely: "which neighbour's cat are you referring to?" In other words, a situation is beginning to take form and we know or at least we strongly anticipate that: (a) it will involve the speaker's neighbour's cat, and (b) it will make a comment concerning the cat.

It can be asked whether these anticipations are legitimate and if so, what gave rise to them? It is one of the claims behind DIG that part of the "primary linguistic data" from which a child derives its language involves a large number of these situations, called "localized contexts" in DIG. The situations themselves are not necessarily bound to any particular structural form. However, actual language activity must be via a concrete language and with increasing instances and situational repetitions, situations and formal patterns become (strongly) associated with the result that eventually pattern mastery and overlearning result. After this, patterns become fossilized and parameters become set. Once the initial state becomes set, two complementary phenomena occur: (a) the speaker finds it more difficult to express a well-known cognitive local context in a pattern different from the one which has become fixed, and (b) slight nuances in the cognitive structure of the cognitive local context are blurred due to the "satificeability principle". All of this is normal development since we effect languaging at a considerable rate and cannot afford to re-analyze and map formal patterns onto local contexts every time we need to do so. Moreover, it appears that local contexts and their formal expression(s) are not in a one-to-one relationship. Some contexts may be so rare as to elude formal expression; others may have several expressions for essentially the same local context. Once again, we seem to be faced with providing satisficeable results rather than absolutely precise, fully specified conditions.

If we continue our input with "ate", we have, following *lexicalization()*, a word whose CATEGORY value is incompatible with nominal-structure. This will trigger closure of the nominal-structure "my neighbour's cat" and initiate a new structure type: verb-structure. Among the properties of verb-structure is the need to fulfill several roles: subject, complement and object. The word "ate" is derived

from "eat(ns1, {ns2}, {complement})", where ns1 and ns2 are assigned certain functional roles. In English, it turns out that ns1 is assigned the subject role and ns2 the object role. In some languages, the nominal-structures may be marked for functional roles as in Latin or Russian, while in others, word order or even particles will be used to indicate such roles. The underlying fact is that, regardless of language, verb-structures are marked lexically as requiring certain parametric functional roles. Traditionally, these roles have been studied in isolation but there is growing consensus that functional role information overlaps lexical information (cf. Schalley 2004). If this is correct, then part of what a child acquires through exposure, possibly as an innate part of its "language organ", is an instinctive recognition of a fundamental distinction between nominal-structures and verb-structures. This is another claim being explored in DIG. Among the unexpected results to date is the growing understanding that functional roles are simply part of the information accumulation which goes on during the languaging process. This allows us to represent functional roles in terms of the same features which are used to describe lexical items. Feature structure compatibility and specification become the criteria by which functional roles are either assigned or blocked. In the latter case, the utterance is said to be ungrammatical in terms of one or more of a set of incompatibilities. It could be incompatibility in INDEX values, or SEM{...} values. There could even be structural incompatibility (see Sévigny 2000, 2002a, 2004 for examples).

When "ate" is lexicalized, we have several things happen. First, the current nominal-structure is closed and receives an attachment of a set of functional roles compatible with nominal-structures. These include among others: subject, direct object, object of a preposition, topic, appositive, subjective completion, etc. There are about a dozen or so functional roles compatible with nominal-structures. Again, these roles appear to belong to universal grammar since they seem to occur in all human languages, with occasional variations here and there. Then, the verb-structure is initiated. Included in its lexical information is the fact that it requires at least one nominal-structure marked [+animate], among other things because "eat(ns1, {ns2}, {complement})" requires that ns1 be animate and ns2, if it occurs, to be compatible with the SEM feature [+edible]. Where does this information come from? DIG claims that accumulation and storage of such features is part and parcel of a child's "primary linguistic data". For instance, it is not uncommon for a child to claim that its cat "ate its teddy bear" because the child has not yet learned the rule concerning eater and edible objects referred to above. Thus, part of a child's "primary linguistic data" includes storage and application of an open set of SEM features, as well as lexical and functional information associated with lexical items.[11]

We may now input "our", whose CATEGORY value is nominal-structure. This is

[11] It seems highly likely that SEM{ ... } features can be put into a mapping relation with neuronal sequence sets as described in Pulvermuller (2003). This is one of the areas of current interest in the DIG approach. We hope to arrive at positive results in the near future.

incompatible with the current verb-structure and therefore triggers closure of verb-structure and initiates a second nominal-structure. With the closing of the verb-structure, it becomes possible to specify a possible functional role for the closed nominal-structure. In this case, because of factors specific to English grammar, the most likely candidate role is subject. Since "cat" is marked [+animate], among other things, and since it is indexically compatible with "ate(ns1, {ns2}, {complement}), it can and does unify with the verb-structure's first nominal argument and assumes the role of "subject". This triggers a local context which in turn brings in other informational parameters. We can schematize the information accumulated to date, as in (16).

(16) SITUATION: Doer: ns1 = ⟨our neighbour's CAT_{head}⟩
NAME: ⟨cat⟩
CATEGORY: common noun: concrete
INDEX: gender [ ], number [+sg], person [+3rd]
STRUCTURE-TYPE: nominal-structure
HEAD?: +
SEM: {OWNERSHIP: speaker, [+object, +modifier], [+human, +modifier], [+animate, +modifier], [+ownership, +modifier], {[+object], [+animate], [+animal], [+feline], ... } }
EVENT: eat ns1[+animate], {ns2[+edible]}, {complement})
COMPLEMENT: ?

In DIG, local contexts are represented with Situation-schemas. It is another claim of DIG that a child's "primary linguistic data" involves exposure to a large store of situations. Once a situation structure has become fixed, it lends itself readily to anticipation and hence speeds up the process of processing incoming information, a necessary requisite condition if we are to maintain normal communication within real-time constraints. It is interesting to note that everything slows drastically if either the situation developing is new to the listener as would be the case if the topic of discussion were unfamiliar or if the words being input were not found in the listener's/reader's lexical database. It is also worthwhile noticing that commonsense knowledge and world knowledge are involved during the process of local context specifications and resolutions. As a child becomes more familiar with his/her world, all information areas are expanded.

Finally, let us skip ahead a bit and process "goldfish last night". This will yield the following updated Situation (17):

(17) SITUATION: Doer: ns1 = ⟨our neighbour's CAT_{head}⟩
NAME: ⟨cat⟩
CATEGORY: common noun: concrete
INDEX: gender [ ], number [+sg], person [+3rd]
STRUCTURE-TYPE: nominal-structure
HEAD?: +

SEM: {OWNERSHIP: speaker, [+object, +modifier], [+human, +modifier], [+animate, +modifier], [+ownership, +modifier], {[+object], [+animate], [+animal], [+feline], ... } }

EVENT: eat(ns1[+animate], , ns2[+edible], (nominal-structure3modifiertime:past)[+past]

OBJECT: ns2 = ⟨my goldfish$_{head}$⟩

CATEGORY: common noun: concrete

INDEX: gender [ ], number [+sg], person [+3rd]

STRUCTURE-TYPE: nominal-structure

HEAD?: +

SEM: {OWNERSHIP: speaker, [+object, +modifier], [+human, +modifier], [+animate, +modifier], [+ownership, +modifier] [+object], [+animal], [+fish], [+animate], [+edible], ... }

COMPLEMENT: ns3 = ⟨last night$_{head}$⟩

CATEGORY: common noun: abstract

INDEX: gender [ ], number [+sg], person [+3rd]

STRUCTURE-TYPE: nominal-structure

HEAD?: +

SEM: { [+object], [+time], {[+past], ... }

The final input is the period (".") which, in DIG, is one of several terminators. Terminators trigger complete closure which, in turn, triggers several mop-up processes the final result of which is the generation of a discourse unit with additional information parameters. These parameters, among others, indicate the general type of information which has just been generated. For instance, it could be narration, description, argumentation. Given that the exact nature of these general categories can be problematic, but less so if viewed from the perspective of satisficeability, DIG usually labels these information parameters using composite terms such as description-narration if the lexicalized words making up the discourse unit consist mostly of descriptive words rather than of words denoting action and flow of time.

5. Explanatory Adequacy, Grammaticality and Strategies

Earlier, we mentioned that Chomsky (2000: 7) wrote of descriptive adequacy in terms of "a full and accurate account of the properties of the language" which a speaker uses. He continued and explained that this amounted to detailing "what the speaker of that particular language knows" (p. 8). However, as we noted earlier, the expressions "a full and accurate account of the properties of a language" and "what the speaker of that particular language knows" are very abstract. What is meant by these expressions is not at all clear. We need to concretize such expressions in order to be able to build models with which to predict languaging outcomes and performances. The information-based approach developed in DIG brings some of these needed concrete details to language theory.

Assuming that the work presented in section 4.1 is correct, we now have at least a working model with which to explore both descriptive adequacy and explanatory adequacy. Thinking about languaging as a function of information processing, memory and storage brings about a shift in focus in linguistic theorizing away from syntax proper towards strategies for gauging and channeling the flow of information in discourse. This allows us to list and briefly comment on a few feasible properties of the language faculty, which would appear to include at least the following (18):

(18) a. The ability to store information in lexical entries which can become networked to yield structural, functional, contextual, semantic and discourse information. Such informational knowledge accumulation is not hardwired; rather, it is the result of experience. We also note that a lot of this information can be represented in terms of feature bundles which are usually of a binary nature although there can be higher arities. Moreover, a lot of information is unspecified by default but open to specification as soon as local contexts and/or functional relations become specifiable. Such an approach implies that there is no actual line of division between syntax and semantics, nor even necessarily with pragmatics, a claim which supports Sperber and Wilson's Relevance Theory (1994).

b. Another property of language is the innate orientation of the linguistic neophyte toward the recognition, construction and storage of localized context situations. It is a claim of DIG that a person masters his/her language in terms of such an accumulation of localized contexts and their accompanying mappings onto overlearned patterns. It is also a claim that these localized contexts or situational constructs are universal in so far as they can be observed at work in every normal human language. In all probability, they form part of human perceptual and organizational structures and limits. In essence, DIG claims that human information expression is limited to such situational schemas.[12]

c. A third property of language is our universal use of patterns which become overlearned because of their limits and frequency. Such a claim can be easily tested because it states that if a speaker encounters an unusual context, situational construct or unusual pattern, then processing will slow down, possibly even be temporarily rejected as "incorrect".

d. A fourth property of language lies in our universal ability to create metaphor from context, a natural emergent property resulting from our innate predisposition to view information accumulations in terms of fixed, overlearned localized contexts.

e. A fifth property of language lies in the special status of functional roles which are: (a) subject to structural patterns which become overlearned, and (b) effected subject to compatible feature specifications. Common mechanisms at work here are word order, functional endings on words, use of particles. The neophyte does not know that word order is important in English, that functional endings are important in Russian or particles are important in Japanese. As Chomsky would say, these are mere idiosyncracies. The interesting points are two: (a) that these mechanisms are quite limited, and (b) that they are quickly seized upon. Indeed,

[12]See below for a further implication of this claim.

> the rapidity with which first language learning takes place certainly reduces the possibility that functional properties of language are open ended in terms of the number and type which could be expected were the mind a linguistic *tabula rasa*.

The short list presented in (18) represents an incomplete, but measurable set of language properties. If it can be established that properties of this nature are indeed the bricks and mortar of language, then we can get on with the business of developing models which will eventually encompass unknown implications and further properties of language.

With regard to Chomsky's remarks on explanatory adequacy, namely that a theory of language with this property can show how each particular language can be derived from a uniform initial state under the "boundary conditions" set by experience, we note that the situational contexts described in section 4.1 and further commented upon in (18c) are clearly abstract constructs common to all languages regardless of a particular language's concrete structure and lexicon and that humans acquire these situational constructs at an astonishing rate and at a surprisingly slow rate when it comes to acquiring another language later in life. These observations lead us to theorize that part of the "boundary conditions" faced by experience require that (a) languaging be done in terms of information gathering, processing, storage and generating, and (b) that once experience sets in, the "boundary conditions" result in patterns which effect mappings between patterns, information and situational schema closures. A further claim hinted at above is that this kind of information is innate to humans in general and seems to be a conceptual limit of the species. As far as "idiosyncratic" varieties are concerned, this simply reduces to the observation that it does not matter how and in what order situational schemas become filled. It is also worthy of note that once these schemas and their realizations become fixed (or overlearned), they can be processed more easily and more rapidly, a necessary property if people are to communicate within stringent time limitations.

Finally, we note that we have already initiated a reply to Chomsky's third question where he asks why language properties are the way they are. DIG answers that once we look at language from its fundamental purpose rather than from its structure or origin, we see that the properties examined briefly in this paper fall into place if information is to be acquired and processed within realistic real-time limits. It seems that humans do not think in terms of structures. Rather, they think in terms of overlearned patterns which themselves are merely mappings of highly similar situational and informational schemas concatenated and networked into discourse streams. Far from abstracting a complex grammar from an "incomplete set of data", the child is barraged by a steady flow of thousands of (repeated) and/or similar situational contexts. Eventually, overlearning occurs, habits set in and generate overlearned patterns which are used to anticipate other situational developments, most of which turn out to be correct. From all of this experience, the child acquires the system, or the grammar.

6. CONCLUSIONS

When information — including its representation and its patterns of flow — is taken as the principal perspective from which to study language, a new outlook for linguistic enquiry follows naturally. One of the requirements of this new outlook is that 'information" be defined in a manner compatible with the process of languaging. In this paper, we have shown, though only in a cursory fashion, that from the perspective of languaging, information comes from several traditional domains: lexical, functional, structural, situational, pragmatic and discursive. Furthermore, we have also shown that it can be captured using a consistent set of complex feature-value pairs, subject to a small number of simple operators and organizational principles, notable among which are networking of verified, compatible feature specifications.

As regards representational phenomena, there are two main generic ways information accumulation could be described: (a) as a set of multilayers where each layer i acts as the base for the next layer $i + 1$ which, in turn, "defines" the lower layer; and (b) by incremental growth through accumulation and networking of possible relations and/or specifications, subject to consistency and compatibility constraints. DIG has focused on the latter approach because it seemed more natural to a realistic accounting of such information accumulation because of constraints imposed by natural, dynamic, linear, time-constrained processing.

Several suggestions were also made indicating that language processing and/or acquisition may involve considerable cognitive activity. This led to a third hypothesis: namely, that it may turn out to be possible to establish precise connections between such cognitive activities with brain activities using mappings based on neuronal sequence sets. A fourth area that seems promising, though highly speculative at present, is the possibility of coming to grips with how words might be stored. A fifth area which also seems very promising is the conception of language in terms of cognitive informational networking achieved via localized contexts, each with a definite structure, yet originally based on potential mappings onto actual utterances. Should such a hypothesis turn out to be correct, it would be consistent with the commonplace observation that physical organs used in the production of language do not appear to have been originally designed for such activity. An indirect consequence of such a discovery would be a (partial) vindication that language followed cognitive development as claimed, for instance, by Bickerton (1990, 1995) and Deacon (1997). Finally, it is worth noting that language layering cannot account for language activity simply because such activity is primarily an activity of information accumulation and organization rather than a set of rule-based abstractions operating independently of contextual information.

REFERENCES

Alexander, Christopher. 1977. *A pattern language: Town, building, construction.* New York: Oxford University Press.

Anderson, S. and D. Lightfoot. 2002. *Linguistics as cognitive physiology.* Cam-

bridge: Cambridge University Press.

Bickerton, D. 1990. *Language and species*. Chicago: University of Chicago Press.

Bickerton, D. 1995. *Language and human behavior*. Seattle: University of Washington Press.

Cann, R., R. Kempson, and L. Marten. 2005. *The Dynamics of language*. Amsterdam: Elsevier.

Chomsky, N. 1981. *Lectures on government and binding*. Dordrecht: Foris.

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

Chomsky, N. 2001. Beyond explanatory adequacy. *MIT Occasional Papers in Linguistics* 20, 1–28. Linguistics Department, Massachusetts Institute of Technology, Cambridge, MA.

Chomsky, N. 2002. *On nature and language*. New York: Cambridge University Press.

Deacon, T.W. 1997. *The symbolic species: The co-evolution of language and the brain*. New York: Norton.

Hausser, R. 1999. *Introduction to computational linguistics*. New York: Springer.

Hornstein, N. 1998. Noam Chomsky. In *Routledge encyclopaedia of philosophy*, gen. ed. E. Craig. London: Routledge. On-line version.

Kempson, R., W. Meyer-Viol, and D. Gabbay. 2001. *Dynamic syntax*. Oxford: Blackwell.

Milward, David. 1994. Dynamic dependency grammar. *Linguistics and Philosophy* 17: 561–605.

Pulvermullerm F. 2003. Neuroscience of language: On brain circuits of Words and Serial Word Order. Cambridge: Cambridge Univesity Press.

Schalley, A.C.. 2004. Cognitive modeling and verbal semantics: A representational framework based on UML. Berlin: Mouton de Gruyter.

Sévigny, A. 2000. Lexically-driven discourse analysis and construction. Doctoral dissertation, University of Toronto.

Sévigny, A. 2002a. Discourse information grammar. *Linguistics Association of Korea Journal* 10(4): 65–91.

Sévigny, A. 2002b. Information flow in excerpts of two translations of Mme Bovary. *Linguistica Antverpiensia* (n.s.) 1: 259–272.

Sévigny, A. 2003. Towards a procedural grammar for natural language understanding. *Canadian Journal of Applied Linguistics* 6: 187–217.

Sévigny, A. 2004. Coordination from a procedural, time-linear perspective. *LACUS Forum* 30: 425–439.

Simon, H. 1996 *The sciences of the artificial.* Cambridge, MA: MIT Press.

Shin, G. 1987. Linearized phrase structure grammar. Doctoral dissertation. Chonnam National University, South Korea.

Sperber, D. and D. Wilson. 1994. *Relevance: Communication and cognition.* Oxford: Blackwell.

Tomasello, M. 1998. Introduction: A cognitive-functional perspective on language structure. In *The New Psychology of Language*, vol. 1, ed. M. Tomasello, vii–xxiii). Mahwah, NJ: Lawrence Erlbaum.

Tugwell, D. 1993. Dynamic syntax. Doctoral dissertation. University of Edinburgh.