Chapter 1

Introduction

1.1   Theses to Demonstrate

This dissertation lays out a new theoretical framework, called **pattern matching method for the analysis of language syntax,** or **pattern matching analysis** for short, and reports on a set of initial research results. The framework proposed came out of my long-term research project, which ultimately aims at providing a natural framework for a realistic description of language syntax. See Kuroda (1996, 1997) for earlier proposals and results, though some of them are obsolete from the perspective that I will take. Some assumptions made therein turned out to be inadequate.

The purpose of this dissertation is to demonstrate a group of theses that are closely related to each other. The most fundamental theses are about words, in question of what they are. I will try to show:

A. **Existence of word syntax,** or **lexical syntax**— words, as mental representations, are more than "pairs" of sound and meaning, or of form and meaning. **Words have syntax of their own, in addition to** their phonology/phonetics and semantics.

B. Word syntax has an obvious cognitive basis in that it consists in **knowledge of co-occurrence (in a sequence)**, which is presumably a special kind of the "part/whole" relation. More specifically, word syntax is characterized as schematic knowledge of **the environments,** or **contexts,** in which words occur.

C. **Sufficiency of word syntax** — word syntax constitutes the **minimum knowledge of syntax,** and the overall knowledge of syntax is no more than an "emergent" structure out of the complex interaction among pieces of this minimum knowledge of syntax.

A demonstration of these theses is carried out in the framework of pattern matching analysis that I propose for such purposes.

The theses above bring us with some other theses related to them. By equating

grammar, or "knowledge of language", in the Chomskian sense, with knowledge of words, I will further show:

I. Grammar boils down to **concrete, learnable knowledge of words,** under some conceptual revisions specified just below on what words are, rather than to **abstract, unlearnable knowledge** called Universal Grammar.

II. Knowledge of syntax, in the sense defined here, must be described **in addition to** knowledge of semantics (and pragmatics) and phonology (and phonetics), because **it is irreducible to either kind of knowledge.**

The framework proposed distinguishes itself with many other frameworks in some crucial respects. Among others, it also aims to realize another thesis, III, with an auxiliary thesis, IV, stated as follows:

III. Knowledge of language, equated with knowledge of words, must be **described realistically** rather than formally.

IV. A realistic description of the language should be maximally **compatible with connectionist theories and results.**

For expository purposes, the first thesis will be called the **concreteness thesis**. The second will be called the **proper description (of syntax) thesis**; the third will be called the **realistic description (of syntax) thesis**; and the fourth, auxiliary thesis will be called the **connectionism-compatibility thesis**.

The proposed approach to linguistic issues contrasts with most "Chomskian" trends in linguistics, if not with all "generative" ones, because they do not aim to provide realistic descriptions of knowledge of language, which is expected to be compatible with connectionism. Such Chomskian theories hold that knowledge of language, or **grammar** in a technical sense, is an abstract, "unlearnable" system of Universal Grammar (UG for short), rather than a concrete, "learnable" system of words. In short, the first, concreteness thesis, in conjunction with the third, realistic description thesis, makes the approach **UG-free**. I will provides detailed arguments for this statement in Section 1.3.

In equating knowledge of language with knowledge of "words", my position is perhaps in conformity with R. Hudson's framework of **word grammar** (1984, 1990), which claims explicitly that knowledge of language is knowledge of words.[1]

In holding the third, realistic description thesis, the proposed approach is compatible, at least conceptually, with most "cognitive" approaches to language, such as Fauconnier's **mental spaces theory** (1994, 1997), Lakoff's **cognitive semantics** (1987, 1990), Langacker's **cognitive grammar** (1987, 1991a, b), Yamanashi's version of cognitive linguistics (1995, 1999a, b).

Despite this, it should be noted explicitly that this compatibility is sure to be reduced when wanting to be more faithful to the second, proper description thesis. Note that most "cognitive" approaches claim, explicitly or implicitly, that knowl-

edge of language is reducible to **knowledge of meaning** (or **conceptualization**). Admittedly, knowledge of meaning is fundamental, but in holding the proper description thesis, this approach distinguishes itself from any approaches which subscribe, willingly or unwillingly, to an idea that semantics and pragmatics are more important aspects of language, and accordingly they are more responsible for language structure. It is simplistic to hold, without providing any reliable technical detail, that language structure makes its appearance out of language use. Constraints work positively and negatively, on the one hand, and internally and externally, on the other. It is very unlikely if language structure is not constrained internally, as well as externally, unless language use determines every detail of language. My approach inquires into the internal conditions, rather than external ones, since I have seen no evidence that language structure is not internally conditioned. I will provide detailed arguments related to this issue in Section 1.2.

The alleged compatibility of the proposed framework with some "cognitive" approaches does not imply any incompatibility with all incarnations of generative grammar. The contrary should be true. The proposed framework will try to be compatible with research results found in generative linguistics as much as possible, rather than trying to replace them. Translation of descriptions in one formalism, say, of generative grammar, into another formalism, say, of cognitive grammar, is not adequately considered a replacement.

In my view, generative and cognitive approaches are not really incompatible. In many cases, they complement each other in what they describe. If they are ever in some conflict, it is more likely so for ideological reasons. It is sure that most approaches in generative linguistics do not commit to provide a "realistic" description of knowledge of language, thereby failing to meet the realistic description thesis above. Even with this fact, it is not fair to judge that so-called cognitive approaches are more realistic than generative approaches. I want to note that it is metatheoretical and perhaps even ideological matters what actually make descriptions realistic. Without certain implicit "common ground", it is meaningless to compare two different theoretical frameworks that have different motivations, concerns, and goals. Usually, however, there is no such common ground.

Claiming a theoretical framework "cognitive" is obviously insufficient for scientific purposes, unless there are a set of criteria to show independently that it really is. One should assume a higher standard for the realism in description and theory-construction. I advocate the connectionism-compatibility thesis for this specific purpose. One can check whether a proposed framework is really realistic, or more realistic than others, by taking into consideration the compatibility with theories and results in connectionist research.[2] I will provide detailed arguments for this point in Section 1.4.

Another point comes into place if the issue is taken into account of what makes descriptions "look realistic". Let me note first that three things are utterly different. First, in what formalism a description is expressed. Second, how truthful a description is, and third, what is actually described in a description. In this re-

gard, I find it very unfortunate that there are many wrong judgments made on linguistic descriptions by confusing the three matters. Many linguists judge descriptions on what they are made in when they have to judge according to what they describe. In the case of syntax, it does not really matter whether linguists talk mathematically or graphically, whether they use "rules" or "diagrams", and whether they use "tree diagrams" or "image diagrams" to describe linguistically relevant structures such as syntactic structure. What can really matter is nothing but the **problem of correspondence**. Note that any rules or diagrams themselves cannot be identified with what they describe. It is an obvious category mistake. Only **what such diagrams denote can be equated with what they are used to describe**. Thus, if there is something interesting in language structure, semantic or syntactic, it is **formalism-invariant structures** that should be represented in a certain way.[3]

As far as the issue of **how knowledge of words are represented** is concerned, the proposed framework will get along with some notable approaches that came out of the generative tradition. As I will show later, it can be counted as an instance of those theoretical attempts to make use of **redundancy in specification** in representation of linguistic units. As far as I can tell, the idea of **underspecification**, proposed by Archangeli (1984, 1988) for phonological representation, is quite in conformity with the proposed approach.

More generally, pattern matching analysis may subscribe to the **declarative perspective**, especially in the sense of Bird (1995) and Scobbie (1997), which contrasts with the **procedural perspective** on the problem of how knowledge of language (and of words in our case) is represented. The view of **grammar as a simultaneous constraint satisfaction system**, a thesis advocated in Lakoff (1993) for phonology, is best understood from the declarative point of view rather than from the procedural point of view. The reason is that if constraints are satisfied procedurally, their resulting satisfaction fails to be truthfully simultaneous. By simultaneity, I intend an invariance in processing order.

1.1.1   Two views of grammar

One of the recent trends in theoretical linguistics seems to be a restatement of descriptions in terms of "derivations" to descriptions in terms of "constraints". This shift can be seen as a manifestation of the conceptual shift from the "procedural" conception of grammar to the "declarative" conception of it, if not a replacement of the former by the latter.

I feel this shift is quite natural for some reasons. Most importantly, in the procedural conception of grammar, description of grammar is mistaken for grammar itself, which clearly suffers from the **first order isomorphism fallacy** in the sense of Kugler, Turvey, and Shaw (1982). Since it is sure that language results from mental and neural processes like any bodily motions result from coordinations of a number of muscular movements, it is quite tempting, and perhaps partly adequate, to understand grammar in procedural terms. This view is dangerous, however, in that it invites the confusion of a description of a grammar with a

grammar being described.

Despite serious problems arising from the first order isomorphism fallacy, most linguists relied on the misleading processing metaphor, and never stopped viewing grammar as a "factory", which comprises a bulk of mental processes in the brain, to bring "products" called sentences.

The declarative view has appeared to overcome such problems, if I follow a summary by Bird (1995), after a long domination of the procedural view of linguistic phenomena, with only a few notable exceptions.

The first exception is Wheeler (1981, 1988) who proposed, in collaboration with E. Bach, a framework of **categorial phonology**, or **Montague phonology**, which showed that phonology can be made derivation-free by appealing to the idea of **functor** and making use of (logical) implications. As it turns out, the proposed approach freely adopts some of the basic insights of categorial grammar.

The second exception is Diehl (1981) who proposed a **lexical-generative grammar**, which has no base component, and generation takes place "in the lexicon", by making use of redundant specifications, implicitly appealing to the idea of underspecification. In a crucial sense, I should admit, the framework of pattern matching analysis could be thought of a modern version of Diehl's idea, which seemed to be too innovative to be appreciated when it was published, when there were virtually no linguists who were able to understand what the declarative perspective can bring.

There are possibly many other approaches that are compatible with the framework that I am going to expound in this thesis. But, I believe I can claim that the proposed framework of pattern matching analysis is different from the others in one important respect. The framework was conceived of, and has been developed, under strong influences of connectionist results, especially of Elman's (1990, *et seq.*), as I will mention below.

This is probably what one can hardly expect most frameworks in linguistics to be, except **optimality theory** (Archangeli and Langengoen 1997; Prince and Smolensky 1993), though there is a deep conceptual difference in that optimality theory does not try to be a "realistic" theory to meet the third thesis stated above. In particular, the theory is not UG-free, and it is better to say that it merely provides a "formal" description of natural language phonology and syntax. I will touch on this issue again in Chapter 3.

Most linguists believe that they need not worry about whether their descriptions really fit the empirical data, without having no necessity for checking validity of their descriptions. They seem to think as if such validity checking were a "dirty" work for psychologists. They seem to feel comfortable in believing that they made "real" descriptions out of "real" data. I do not want to subscribe to this kind of attitude. I believe that efforts to be truthfully "realistic" necessitate efforts to be responsible for one's theoretical proposals, by providing "readiness for implementation", at least partially.

There is a set of connectionist experiments that the fundamental idea of the proposed framework crucially relies on. It is the series of simulations by Jeff El-

man's (1990*a*, *b*, *et seq.*). He showed that connectionist networks of a specific sort called "simple recurrent networks" are able to learn a small portion of English, a context-free language including center-embedding and long-distance agreement. Through detailed analysis of the networks that have learned the language, Elman shows two important things. The first is that language syntax is learnable by generalization over surface forms alone. Second, syntax of sentences represented in such networks can be formulated as a network of transitions in a high-dimensional state space, whereby words serve as "operators" rather than nodes to be "operated" (Elman 1995). Third, and most interestingly, words in such recurrent networks appear to be encoded in context-sensitive fashion.[4]

Through Elman's research, though controversially, we understand that the abstractness of knowledge of language can be minimum, and more explicitly, basics of language syntax can be **word-based, surface-true generalizations** alone. This is roughly what I have meant by **word syntax** above.

This is an important suggestion. Thus, in my attempt to describe language syntax realistically, I will try not to rely on anything but surface-true generalizations, hoping that this is the best way to make linguistic description of syntax connectionism-compatible, which I hold is the ultimate form of a realistic description of language syntax.

In claiming for the sufficiency of surface-true generalizations, the proposed framework is, in a sense, going to revive insights of **natural generative phonology**, proposed and developed by such authors as Vennemann (1971, *et seq.*) [Bybee] Hooper (1972, *et seq.*), [Bybee] Hooper and Terrel (1976), and G. Hudson (1974, *et seq.*). It is not an accident that proponents of natural generative phonologists objected to the abstractness in phonological descriptions that Chomsky and Halle (1968) admitted generative phonologists to appeal to.

It should be recalled that natural generative phonology was a trend of generative phonology, and proponents of natural generative phonology did not disagree on what formalism linguists should rely on. Rather, they react negatively to the unrestricted "interpretation" of the formalism. In this respect, too, if there was something wrong with generative linguistics, it is not the formalism they used, but the "vision" under which generative linguists interpreted it, which allowed for too much abstractness in formalism, defended under the ghost notion of "competence", which has no relevance to a theory of natural language.

As I noted elsewhere, something is wrong with a theory of natural language syntax that asserts that (1)a, in contrast to (1)b, is **potentially a sentence of English.**

(1)   a. ?*Colorless green ideas sleep furiously*.
       b.   *Revolutionary new ideas appear infrequently*.

My position is that any grammatical theory is an "unnatural" theory of natural language syntax if it equates the possibility of forms like (1)a with the possibility of abstract string (of preterminal symbols) like (1)a (or (1)b).

(2)   a. Adj Adj N V Adv
        b. Adj″ Adj″ N′ V Adv″

Admittedly, it is possible and reasonable in some sense to define a new dimension of well-formedness by devising an abstract system that generate expressions like (1)a, in addition to (1)b. It is not clear at all what relevance the new dimension of so-called "grammaticality" has to the reality of English grammar. My point is this: stating that (1)a is grammatical should not mean more than stating that strings like (2)a and b are grammatical.

   Note that the abnormality of (1)a and the normality of (1)b are lexical in nature. Strings like (1)a and b are different from (2)a and b, composed exclusively of preterminals like Adj$^{('')}$, N$^{(')}$, V, Adv$^{('')}$, which, by definition, are meaning-free and phonology-free abstract variables.

   It is urgent to ask: Is it an empirical fact that preterminals lack semantic and phonological contents? Or, isn't it so by definition?

   This reveals an important point: the artifactuality in Chomsky's argument for grammaticality (judgement) as distinguished from acceptability (judgement) is the artifactuality in the implicit assumption that **abstract strings like** (2)a **and** (2)b **exist independently of lexical items** like *colorless*, *revolutionary*, *green*, *new*, *ideas*, *sleep*, *appear*, *furiously*, *infrequently*.

   In the reappraisal here, I argue against the notion of **competence** in the sense of Chomsky (1965), by which linguists are entitled to say that expressions like (1)a are "grammatical". The point is that a theory (of English) to state that (1)a is potentially a sentence (of English) is too unrestricted for descriptive purposes, and there is no need for such an unrestricted theory. In fact, it would be false if one does not state that (1)a (no longer) is English, even if it is **a sentence of a language other than English,** known to nobody.

   A realistically restricted theory of English syntax, rather than of a language that nobody knows, should definitely state that (1)a is **not a sentence of English,** while (1)b definitely is.[5]

1.1.2   Outlining   arguments for the proposed framework

It is urgent to discuss a few methodological and/or metatheoretical issues in some detail, thereby making understandable the goals, concerns, motivations, and assumptions of the framework I assume because it is unfamiliar to most readers. For this, I present three considerations to clarify and justify objects of the inquiry and place it in a proper context.

   The following statement sums up the goal of the proposed framework.

(A)   Pattern matching analysis is a framework which tries to provide a **serious** account of **knowledge of syntax** which is **free from Universal Grammar**

(**UG-free** for short) and **connectionism-aware**.

I admit that this statement is controversial since it contains several controversial notions "seriousness", "knowledge of syntax", "UG-freeness", "connectionism-awareness". To give validity and entirety to this statement, let me make it more explicit in the following sections.

First, I will give, in Section 1.1.3, a brief sketch of the proposed framework, providing arguments for the controversial claim that **syntax of language cannot be reduced to semantics and/or pragmatics of it**, and accordingly **syntax must be studied for its own sake**, if not in isolation. In addition, it is argued that the proposed framework has nothing to do with so-called universal grammar, though it investigates in connectionistically definable **human language potential**.

Finally, in Section 1.5, we see that primary object of our inquiry is **grammar**, equated with **knowledge of language**, rather than **language *per se***, by noting that it is inadequate to conceive of grammar as neutral to the speaker/hearer distinction, and insensitive to the personal/social differentiation.

1.1.3   What makes PMA different from other approaches?

To make a few subtle points clearer, let me first make explicit specific assumptions about the relation among syntax, semantics, and pragmatics of language that makes the proposed framework different from previous approaches to language and language syntax.

(B)   By **serious account of syntax** in the statement above, I mean that language syntax is something that should receive a due account in and by itself. More specifically, pattern matching analysis keeps itself away from linguistic theories which try to discharge the burden of explanation by apparently "reducing" syntax to other aspects of language such as semantics and pragmatics. Such treatment, in my opinion, is a mere gerrymandering.

(C)   Pattern matching analysis assumes that **knowledge of syntax is inseparable from knowledge of semantics (and pragmatics)**. I will expound this more clearly in following chapters, but let me note here that this point makes the proposed approach basically similar to most approaches in cognitive linguistics and different from most approaches in generative linguistics.

Some may wonder if there is not inconsistency in (B) and (C). As I will expound this more clearly in following chapters, there is no inconsistency. To anticipate, let me note a few relevant points.

Adoption of (B) is virtually an adoption of a weakest form of the notorious **autonomous syntax thesis**. This indeed makes the proposed framework conceptually, if not factually, incompatible with most approaches in cognitive linguistics. But there is no inconsistency in the proposed approach. What pattern matching anal-

ysis claims about knowledge of syntax is not its **independence** from all other kinds of knowledge, but its **irreducibility**. Neurally, it is not mysterious at all if two autonomous processes are **mutually dependent** or **interdependent** even if they remain autonomous.

This is an important point that proponents of leading cognitive linguistics like Fauconnier, Lakoff, and Langacker fail to recognize when they hold that syntax is reducible to the relation between sound and meaning systems, because it is (at least partly) dependent of both, or either. The logic of mutual exclusion that they are utilizing is invalid, at least neurally. They argue that syntax is reducible to semantics and/or pragmatics, based on the fact that syntax is dependent of semantics and/or pragmatics. No evidence of syntax's dependence on semantics ever entails syntax's reducibility to semantics.

First, but not most crucially, if their argument were true, it follows, by the same kind of logic, that semantics must be reducible to syntax, because there is a lot of evidence that semantics is (partly) dependent of syntax. Second, and more crucially, it is very likely, for reasons specified above, that syntax, semantics and pragmatics are differentiated processes that run in parallel and autonomously. This becomes more plausible if we take into consideration the emergence of syntax. If syntax is emergent from language use, which we believe is true, it implies, against an argument by Langacker (1997) to the contrary, that syntax is irreducible to language use. Recall that emergence in general means irreducibility, rather than reducibility. Since this is an issue metatheoretically crucial, I shall discuss it in more detail in Appendix B.

To avoid confusion, it is urgent to conceptually distinguish two-way **interdependence** from **one-way dependence** (as a synonym of reducibility). With this distinction, it is very reasonable and even necessary to assume that syntax and semantics (with pragmatics incorporated therein) are interdependent. There are a number of interdependent components in the brain that are specialized for certain perceptual, cognitive and motor tasks. Such components work cooperatively and interdependently, and more importantly autonomously to each other.

I believe that there is a complex system consisting of a number of autonomous processes to form overall knowledge of language. So, there is no inconsistency in claiming that knowledge of syntax is autonomous, on the one hand, and that knowledge of syntax is inseparable from other kinds of knowledge like one of semantics and pragmatics, as well as it is inseparable from knowledge of phonology and morphology. This view, to be called the **complex systems view of language**, is the most reasonable in that it is compatible, at a higher level, with a lot of contradicting pieces of evidence.

On these grounds, I do not emphasize that pattern matching analysis is a "cognitive" approach in the popular sense of the term. I never intend to claim that pattern matching analysis is an enemy of cognitive linguistics. Actually, the contrary should be true. The proposed framework can be a friend of cognitive linguistic if cognitive linguistics is a friend of connectionism. Pattern matching analysis, I would like to claim, is cognitive even if it is not a subscriber to the recent trend of

cognitive linguistics. What makes the proposed approach qualify as a cognitive approach is, in my view, its **connectionism-awareness,** a feature basically independent of qualification of cognitive linguistics. What makes practices in cognitive linguistic worthy of calling cognitive is perhaps that they are, as Lakoff (1987) claims, **mentally real**. They in fact seem to succeed in incorporating a number of cognitive phenomena like prototype effects, metaphor/metonymy, etc. I would like to remark that these so-called cognitive phenomena are rather relevant to higher-level cognition, and there are other kinds of cognitive phenomena that are due to lower-level cognition that connectionism would capture best.

Further complication is that, despite all of what I have specified so far, pattern matching analysis is a virtual friend of generative linguistics, as well. This is so despite the incompatibility of generative linguistics with connectionism. The reason is that pattern matching analysis agrees with generative linguistics in its claim for due account of syntax. Again, it is hoped that knowledge of syntax receives a due account, which cognitive linguistics fails, or avoids, to do so.

In summary, the skirmish between generative and cognitive linguistics seems, in my view, to be political rather than scientific, and no serious conclusion should be drawn.

1.2   What Knowledge is Knowledge of Syntax?

As we have seen, pattern matching analysis assumes, though controversially, as follows:

(D)   **knowledge of syntax exists;** and
(E)   **this knowledge is distinct from knowledge of semantics and/or pragmatics,** and of course from knowledge of phonology and/or phonetics.

The two statements should be supplemented. In agreement with Hudson's (1984, 1990) word grammar perspective, it is claimed more specifically, by, stepping a bit further, that:

(F)   **knowledge of syntax is knowledge of words.**

This statement raises another question: What are words, then? In my view, words are best characterized as subpatterns which serve as **perceptual schemas** in the sense of Arbib (1989). Discussions in this chapter indicate that knowledge of such "words" must be abstract.

The last point can be reformulated as follows:

(G)   **syntax comprises schemas of its own,** provided (sub)patterns are such **schemas for syntax.**

This is one of a few basic claims of the proposed framework, but this simple statement becomes somewhat controversial especially with respect to a dominant view in cognitive linguistics.

An important issue is whether or not my definition of schema, based on Arbib (1989), is identical to what Lakoff (1987) and Johnson (1987) denote by **image schema**,[6] on the one hand, and what Langacker denote by (**constructional**) **schema**, on the other. My suggestion is that even though the same term "schema" is used, my conception of schema for syntax differs conceptually from theirs. This is a serious problem, because, as we will see below, similar terms are used almost in a complementary manner. This issue will be discussed in greater detail in Appendix B .

1.2.1   Remarks on knowledge of language, with special reference to its sensitivity to the hearer/speaker distinction

As explained at the onset, pattern matching analysis investigates the definition of knowledge of syntax. But this is an understatement of a more specific goal. First, as we have seen so far, orientation for the alleged universality of language and knowledge of language is blurred.

Another consideration that most strongly motivates my development of pattern matching analysis is a dissatisfaction at the definition of grammar as a model of linguistic knowledge which is:

(H)   i. **neutral to the personal/social distinction**, and
         ii. **neutral to the hearer/speaker distinction**

See Chomsky (1965) for relevant discussion. Both idealizations are highly problematic. For (H)i, many sociolinguists (and some pragmatists) indeed object to it by arguing that language is social. As far as I can tell, few scholars argue against (H)ii; I know only Hockett's (1961) argument against it by proposing "grammar for the hearer".

Plainly, I think generative linguistics idealizes the situation overly by assuming the neutralities, thereby concealing a lot of complicated matters. To try to have such a model is to blur many possible goals of linguistics. PMA claims as follows:

(I)   If we need a "psychologically plausible", "cognitively real" theory of language, we should not fail to distinguish between knowledge of a speaker (as an encoder of messages) and of a hearer (as a decoder of messages).

The reason is that we see often such tricks that one first assumes that the speaker/hearer distinction, or S/H distinction for short, is "neutralized" in analysis and account of language, but later he or she provides analysis and account exclusively

relevant to hearer's knowledge (or speaker's). It seems that this is caused by the ambiguity of *speaker* results in a number of disastrous consequences, since the term can mean speaker, on the one hand, and **both** speaker and hearer, on the other. The latter meaning is not what I will assume. For consistency, thus, I require that the S/H distinction may not be neutralized.

   Instead of assuming these kinds of neutrality, pattern matching analysis will be concerned with knowledge of language as (i) **personal rather than social**; and (ii) **one of a speaker rather than of a hearer,**[7] by addressing a specific question:

(J)   What does **a speaker** know **in terms of syntax** when he or she speaks **a language?**

This is a specific version of the following question.

(K)   What does **a speaker know** when he or she speaks **a language**?

These are basic issues based on which I find we need tools for analysis like our pattern matching analysis. Implicit in these questions is concerns of:

(L) a. **internalization of language,** or knowledge of language, more than **language *per se***; in other words, **mechanism** more than **manifestation;**
     b. **individuality** of such knowledge more than its **commonality** or **universality;**
     c. a particular aspect of such knowledge called **syntax** (with part of **mor-pho(phono)logy** included therein) more than other aspects of **semantics** and **pragmatics;**
     d. language **production** more than **comprehension**, or **message encoding** more than **message decoding**.

Admittedly, this is not a necessary setting for a linguistic theory. I have chosen this setting via a number of personal preferences, or "biases".

   Concerns (L)a and b come from my conviction that knowledge of a language is something that everyone **acquires from personal experiences**. Concern (L)b comes also from my dissatisfaction with a number of "fake" explanations that generative linguists offer relying on Universal Grammar (UG for short).

   I am unable to explain exactly where concerns (L)c and d come from. Perhaps, this is a reflection of my biased taste. I will not try, at least so seriously, to answer other possible questions, some of which are exemplified as follows:

(M)   What does a speaker of a language know **in terms of semantics** when he or she speaks a language? Or equivalently,
(N)   What does a speaker of a language do when he or she **understands** a language?
(O)   What do speakers of language (**try to**) **accomplish** by speaking?

These questions will be given secondary importance, if at all, in pattern matching analysis, because they are out of my interest.

By not addressing the semantically and pragmatically motivated questions mentioned above, I am certain to remove myself a little away from "cognitive" trends in linguistics (Ungerer and Schmit 1996). I add, however, that this by no means implies that the problems that are not addressed are of little importance. My point is simply this: we have nevertheless due rights to ask different questions like (J) and (K) that I issued above. My hope is that there should be different answers to questions asked differently.

1.3   Why Explanations in Linguistics Should Be "UG-free"

Pattern matching analysis is a framework that I propose to provide a **cognitively (and/or psychologically) realistic description of natural language syntax**. To explicate this assertion, I shall begin with some general remarks.

By a cognitively realistic description, I mean a description such that all of the constructs used in it have mental reality, directly or indirectly attestable. Our goal of proving a cognitively realistic description of language must be contrasted with a generativist goal of providing a formal description of language. But it is necessary to discuss in some detail what distinguishes formal and cognitive realistic descriptions. I will return to this issue later.

I suggest that the proposed framework is based on an innovative conception of **representation** of linguistic units, which is inspired by well known connectionist results, and thereby leads us to a new conception of **composition** (and decomposition) of such units. I will return to this issue in Section 1.3.2 to discuss it more thoroughly.

A cognitively realistic description should be contrasted with a **formal description of language in general and language syntax in particular**. This is a goal of most generative theories of language. But, even if such a goal is reachable, I find it less appealing in light of recent trends in cognitive science.

The main reason why such a goal becomes less attractive is probably that the paradigm of artificial intelligence became, after long research, less attractive, at least as it stands. Classical artificial intelligence caught many tech people's minds by asserting that human mind can be emulated (if not reproduced) if they are lucky enough to be able to give a formal description of it. In fact, to declare "We shall provide a formal description of …" (*à la manière de* Chomsky) was very fashionable in the 60's and 70's. For example, Lerdahl and Jackendoff (1983: 1) began their book, *A Generative Theory of Tonal Music*, by asserting as follows:

> We will take a music theory to be a *formal description of the musical intuitions of a lister who is experienced in a musical idiom.* (original emphasis by italics).

It is rare to see such powerful assertions these days. I guess that more and more people feel that such goal is not sufficient, even if achievable, and more important-ly even feel that it is not a necessary goal.

Surely, this manifests an air of disillusion that came after too much enthusi-asm. This is not the whole story. Such disillusion was spurred by the advent of the paradigm of **parallel distributed processing (PDP)**, or simply **connectionism**, which serves as a revival of classical associationism in the 50's.

What crucially distinguishes the already established paradigm of connection-ism from the classical artificial intelligence is largely the shift of researcher's inter-est from mind's (mere) **systematicity** to **flexibility** (on a par with its overwhelming complexity). No one would disagree that the human mind is incredibly flexible. It seems to accommodate almost everything to any degree of precision. It is clear that such an amazing property is not captured successfully in the paradigm of classical artificial intelligence. Instead, it brought us a lot of expert systems which show little such flexibility. In most cases, it was necessary to "blur" models to make proposed models realistic in the sense discussed here. Many researchers began to realized that this is preposterous; some of them arrived at the idea of PDP.

Despite some obvious deficiencies (such as "scaling" problem), connectionist PDP models appear to show such a kind of flexibility. If such an impression is not false, then it is reasonable to conclude that most people in cognitive science now find that giving a formal description to a psychological phenomenon is one thing, and giving a realistic model to explain it is another.

This dualistic attitude sharply contrasts with one of the assumptions that classical artificial intelligence (AI) held. It claimed, quite controversially, that **psy-chologically realistic models automatically follows from formal descriptions**.[8]

This statement, or expectation, was, I claim, a core assumption that lay at the heart of the classical (rationalist kind of) AI movement and made it so powerful. So, there is no doubt that recent trends in cognitive science show little interest in formal descriptions in the classical sense. Indeed, many connectionist models were criticized by classical AI supporters for their obvious lack of characterization in terms of formal description. For example, Rumelhart and McClelland's (1986) model for English past tense formation was criticized for this reason.

Such criticism is preposterous, it seems: one of their model's most important implications is that it is possible to make so-called formal descriptions implicit, if not unreal. As far as human mind is concerned, there is a great gap between psycho-logically plausible "good" models and mathematically elegant "good" formal descriptions.

Based on the considerations so far, we may conclude as follows. While formal descriptions are not useless, they are not useful as such.

In summary, return to the assessment of generative linguistics. It should be clear now that formal descriptions of language, which generative linguistics was originally proposed to provide, by no means provide realistic models of language for the reason that the classical AI framework fails to give realistic models of hu-

man mind. The reason is that realistic models do not automatically follow from formal descriptions.

1.3.1   Connectionism-compatibility

As noted earlier, connectionism-compatibility plays an important role in our investigation. Connectionism-awareness dictates that it is irrelevant, for purposes of linguistics, to characterize grammar as "competence", as distinguished form "performance".

Connectionism-awareness challenges generative linguistics to concede, and it is very likely that most generative linguists will keep attacking connectionist, **competence-free** account of knowledge of language. We have to try to base the validity of knowledge of language on its learnability rather than its unlearnability, asking what details are endowed to a self-organizing system called the brain to realize so-called knowledge of language.

1.3.2   New light on the notion of representation and composition/decomposition

Given the general goal of providing a cognitively realistic description, I must ask: In what respect is the proposed framework an alternative to classical generative account of language syntax? To make a long story short, it provides a new theory of **representation**. We will see that it is not only psychologically implausible but also unnecessary for descriptive purposes to have a model of syntax having the following properties.

(P)　　i.　There is an lexicon-independent component (usually called "base") to "generate" blindly a number of formations.[9]

ii.　There follows a process called "derivation" to determine (only) one good, if not best, formation out of them, thereby making all other great many formations useless (because of their ungrammaticality).

The first part realizes the assumption of **purpose-free generation**. The second part realizes the assumption of **selection**.[10]

The picture may be good for a formal description in the sense discussed above; but, it does not fit cognitively realistic models very well. First of all, if they are realistic, it follows that there are **processes** of free generation and derivation running in the head. If so, then everyone wants to ask, "Isn't it a theory of performance?"

The picture that I will rely on is a simpler, more intuitively plausible one. Linguistic forms are generated by combing units of a variety of sizes, with some of them being basic and others derived.

Also it is not necessary to define abstract structures (usually called "phrase markers") independent of specifying "selectional" properties of lexical items.

Roughly, adequate description of complex language syntax is achievable only if we assume that words in the lexicon are already structured so that they have subject (and object(s) if any) of their own. In short, words are themselves schemas. Such schematic knowledge is a generalization of surface formations. Thus, the proposed framework sheds a new light on composition. Analogically, what the proposed framework suggests is that conscious and unconscious construction of linguistic units (such as sentences) is comparable to jigsaw puzzle.

1.3.3   UG-based "explanations" are not explanations

I claimed above that the framework of pattern matching analysis is "UG-free". This term can be rephrased **competence-free**. But what does this really mean?

A theory or approach to language (syntax) is UG-free if it does not account for facts of language (syntax) by assuming **Universal Grammar** (and any of its conceptual analogues) in the popular sense of the term.

Explanations in linguistics must be UG-free and accordingly competence-free. The strongest reason is methodological: so-called "principles and parameters" approach (Chomsky 1995) of language is nothing but a tautology for reasons specified below.

A theory of language syntax should be competence-free to avoid stating that (1)a, repeated here, is "grammatical", no matter how the notion grammatical is defined.

(1)   a. ?*Colorless green ideas sleep furiously.
       b.   Revolutionary new ideas appear infrequently.

As discussed earlier, a theory of English grammar is "unnatural" and unrealistic if it enables us to assert that (1)a is grammatical. There is no need for such a powerful theory.

Suppose that the principles-and-parameters theory of language is correct. Then, UG, to make sense, must be some (formal) device that generates grammars rather than languages. To see this, let me illustrate a few technical matters. Assume, for discussion, that language L is defined by grammar G. On the one hand, we have an infinite set $\mathscr{L}$ of natural languages, $\mathscr{L} = \{L_1, ..., L_n\}$. It is a set of Afrikaans, Bedouin Arabic, Chinese, etc. Correspond to $\mathscr{L}$ is a set $\mathscr{G}$ of grammars such that $\mathscr{G} = \{G_1, ... G_n\}$, where $G_i = G(L_i)$. According to UG, this set comprises grammars of Afrikaans, Bedouin Arabic, Chinese, etc, all as a **parameterized realization** of UG.

A simple consequence of this is that UG is not a grammar that is, at least in the usual sense, relevant to "natural" linguistics, as opposed to "formal" linguistics. UG, as the grammar of grammars, is more like a "metagrammar" in the sense of Gazdar, *et al.* (1985), which is conceptually distinguished from particular grammars.

The appeal to UG is partly gratuitous and partly tautological. It is gratuitous because we have no necessity, other than theoretical or ideological one, to appeal to UG; what we have to assume is merely that, given language L as a whole is a collection of "phenetic" features (e.g., f_1 = [±subject precedes object]), the possibility of language is determined by a "local region" within an n-dimensional space defined by $[f_1 f_2 \ldots f_n]$.

The appeal to UG is tautological unless crucial theoretical claims are more explicitly formulated. Note that it is trivial to claim merely that there **exists** a local region in an n-dimensional state space that contains all the points that represent human languages. Assumed identification of languages as points, or local "blurs", in some n-dimensional space already defines an n-dimensional space as a range of possibilities. Thus, it is tautological to claim that realization of grammars is "parametrized" relative to a finite set of "principles of UG" unless it provides two things independently:

(Q)   First, it must specify an empirical formula $\mathscr{F}$ that well distinguishes, in the given space of possibilities, a properly local region of "realizable" grammars from "unrealizable" grammars.

(R)   Second, it must provide a theory to plausibly account for the formula $\mathscr{F}$.

What makes Chomskian linguistics, as distinguished from generative linguistics, conceptually inconsistent is its incapability in balancing theoretical claims of the first kind against ones of the second kind. More often than not Chomskian linguists appeal to so-called language universals. Witness their claim, "Phrases in all languages obey $\overline{\text{X}}$ theory". These sorts of claims have no significance because they are tautological. Note that $\overline{\text{X}}$ theory is nothing but an empirical formula in the sense defined above. Thus, it is a matter of course that it fits most data, because an empirical formula, by definition, is an abstraction from empirical data. More specifically, $\overline{\text{X}}$ theory is merely one way to stating so-called **language universals**, all of which must be considered as an empirical formula.

"Why are there such and such language universals?" is a question that linguists frequently ask after they found them. An account of such universals is meaningless without an additional, very empirically dubious assumption, namely:

(S)   The sets of **realized and thereby observable languages** and of **realizable, or possible languages** are the same, extensionally or intensionally.

It is crucial to distinguish between what is "actual" and what is "potential" in language. Note that an empirical formula is empirical because it is a hypothesis coming out of a large set of observations. But, if we believe what modern philosophy of science teaches us, even an infinite set of observations cannot determine such a potential.

The possibility that **realization itself may be biased** is also important. Here, the

meaning of realization will be better understood if it is replaced by **sampling** in a statistical sense.

I am well aware that it is a hard problem whether realized, actual languages exhaust the potential for language. We are able to decide its validity only hypothetically. Nevertheless I would like to note this: if linguists accept (S) without readiness to check its validity, which is exactly what Chomskians urge us to do, and thereby claim that, based on a couple of fragmentary observations at hand, such empirical formulas as $\overline{\text{X}}$ theory "support" the postulation of UG as an abstract system of principles and parameters, they might as well run the risk of sociologists who have worked on biased, or rather "extrinsically conditioned", data sampled from people living here and there in this world and conclude that languages of people are "parametrized" so that they are either Arabic, Basque, Chinese, Danish, etc. This conclusion is absurd, because nothing essential is explained.

I want to keep myself away from this kind of absurdity. One of most reasonable ways to do so is probably to define UG in its weakest form. If there is something that we may adequately call UG, it is nothing but **biological potential of human for language**. My question is thus how to determine scientifically such potential.

1.4   Why Description of Natural Language Syntax Should be Connectionism-aware

Given the alternative definition of UG as a biological conception of human language potential, I am now ready to discuss another conceptual link, namely **connectionism-awareness.**

Connectionism is a rising research paradigm that has rapidly spread since publication of Rumelhart, *et al.* (1986) and McClelland, *et al.* (1986), the "bible" of PDP. It is a modern, more sophisticated incarnation of associationism from the 1950s.[11]

Connectionism is a challenge to the "classical theory of mind" in the sense of Fodor and Phylyshyn (1988). Based on "neurally inspired" architecture, it rejects the model of thinking taken for granted in classical, symbol-manipulation paradigm: **rules** and **representations,** or more generally, **control structure** and **data structure** in the sense of theory of computation. It claims that both data and their control are **distributed** over a massive network of "units" (likened to neurons); and moreover, data and their control need not (and can not) be **inseparable,** if it is possible to conceptually distinguish.

Compatibility of a linguistic theory with connectionism can never be achieved without due effort and cost; and I believe this is why most linguists are reluctant to incorporate a number of interesting results, consequences, and suggestions from connectionist theorizing. Leaving technical details aside, one of the most striking things that linguists should give away is the following, apparently quite reasonable

assumption:

(T)   Grammar is a system of rules (= control structures) that apply to representations (= data structures)

Since Universal Grammar is not a grammar in the usual sense, then it makes no sense to claim that grammar is not a system of "rules and representations" but a system of "principles and parameters", the latter definition is relevant only to UG as a metagrammar.

1.4.1   Why knowledge of language of particular users rather than universal

This conceptual link forces us to reconsider the status of language universals. As noted above, it is unlikely that all language universals are worthy of being called so. This is so no matter how well formulated they are. This is precisely because realization of language itself may be biased. An alternative view is this:

(U)   A language universal is a meaningful partial characterization of human language potential if, and only if, it expresses a property that is either compatible with, or predictable from, the neural reality of human brain.

From this criterion follow a number of theoretically important but unwelcome consequences:

(V)   to have a theory to "list up" language universals is not a sufficient theory of human language potential.

(W)   to have a theory of competence (equated with UG in the narrow sense) without having a theory of performance is of no significance.

(X)   to have a cognitive theory of language is not sufficient

Naturally, language universals, understood here as observed common properties of human languages, will play a secondary role in my investigation of language syntax. This point makes the proposed approach different form functionally, typologically oriented approaches to language. Also, this makes the proposed approach different from proponents of UG. As I have discussed above in some detail, we can expect little revelation in asserting that such and such language universals come from UG unless it is defined as an empirical formula $\mathscr{H}$ such that, given each language is represented as a point in $\mathscr{L}$, $\mathscr{H}$ specify a (presumably blurring) region L_{H} (for human language) in an n-dimensional space $\mathscr{L}$ (defined by n parameters, preferably orthogonal to each other). The notion of UG without such specifics is scientifically empty. Such question is virtually meaningless unless the degree of complexity connectionist network are able to learn is known.

1.5   Remarks on the Notion *Language*, as distinguished from the Notion of *Grammar*

My theoretical commitment to a "cognitively realistic description of natural language syntax", so goes the title of the present thesis, leads to another important point. Strictly speaking, the object of my inquiry is **not language itself** but a mechanism that underlies it. Language only reflects such a mechanism.

This does not mean, however, that I am a "mentalist". Rather, I am a mechanist who refuses to appeal to the notion of **mind** to explain properties of human language, and who tries instead to appeal to the notion of **brain**.

It is regrettable and even embarrassing that the long tradition of linguistics encourages us to use the term *language* very imprecisely and inconsistently. The trouble is that language may refer both to **what is phenomenological** and **what is mental**. On the one hand, when you say Arabic, Basque, Chinese, Danish, English, etc are (natural) languages, what you intend to say is that there are objects that one can define more or less phenomenologically. On the other hand, when you say that your friend, say, John M., speaks very good French, German, Hungarian, etc, you don't mean the same thing as the previous sentence. What you intend is that he has a distinguishable mental ability (or abilities) to manipulate the languages. In short, mechanism, which is mental, and its product, which is phenomenological, are not properly distinguished.

This is an old dictum via Chomsky (or Lucien Tesnière). In popular belief, it was Noam Chomsky who established the conceptual distinction between the two quite different senses of *language.* He proposed to use the term **grammar** to denote what is mental by retaining language for what is phenomenological.[12]

The distinction is unfashionable now. One finds less and less linguists and psycholinguists who appeal to the notion of grammar crucially. More specifically, it is recently more fashionable to use *language* in the sense of *grammar*, thereby ignoring the fact that language is also a phenomenon that one can define in an objective, extensional basis.

It seems that such shift of emphasis is arbitrary, if not unmotivated. I encounter, mostly in the literature of cognitive linguistics, a number of pseudo-arguments for dismissal of such an important distinction, largely based on misunderstandings. I am unwilling to subscribe to such an intellectual attitude. Preferring *language* over *grammar* does not solve any serious problem. So I engage myself to reformulate the language/grammar distinction in my version.

I remark that treating language as a mental phenomenon is nothing but assuming grammar in a sense somewhat different from Chomsky. Conceptual need for the notion of grammar never disappears even if we stop using grammar. It makes no sense for some cognitive and/or functional linguists to claim, against Chomskians, that there is no grammar but language alone. Such an argument is at least a red herring, has no grounds, and is even childish, which completely misunder-

stands what motivates the notion of grammar. Whether grammar is formal or not is of little importance. What really matters is the distinction between the two modes of language's being (and denotation): in one mode, language is phenomenological, and in another mode, it is mental and a cause of other mental phenomena.

It is worthwhile to note that without the distinction, one has to run a risk of eliminating either the phenomenological or the mental properties in favor of the other. On the one hand, one runs foolish behaviorism if one keeps language from being mental, thereby disallowing oneself to assert safely that language is a mentally rooted phenomenon. On the other hand, one runs foolish idealism if one keeps language from being phenomenological, thereby disallowing oneself to assert that language is an objectively observable phenomenon.

I believe any analysis should start from the fact that languages exist; but, it is undesirable to stay there too long. It is required to step forward until the way that languages exist is precisely specified. By descriptions, I mean tasks of this sort. The tasks could not be achieved without appealing to what is mental, since what controls the phenomenological is no longer phenomenological. Ultimate questions about language could not be successfully answered unless descriptions arrive at the physical level. But, in any case, the proposed approach is at the beginning of a long road from the phenomenological dimension to the physical dimension. It is likely that first a few steps are more important than thousands steps to follow.

Notes

1.   In fact, I was repeatedly and greatly influenced by Hudson's word grammar ideas while developing my own approach to be exposed in what follows. Of course, our approach is, I believe, more advanced in some important issues.

2.   It is necessary in addition that there should be an explicit way to guarantee whether an alleged compatibility is substantial; if not, such compatibility is merely a "terminological" one. Some claims of connectionism-compatibility made in cognitive linguistics seem to be the case. See Collier (1998) for relevant points.

3.   I do not think that image diagrams are better than tree diagrams, despite the popular view in cognitive linguistics.

4.   For example, the internal representation of *John*, which appears in *John married Ann*, is slightly different from that of *John* in *John hit a dog*.

5.   By this, I never suggest that a theory that states forms like *Colorless green ideas sleep furiously* are "well-formed" is of no theoretical importance. Rather, **this kind of well-formedness should have the least impact on the description of natural language syntax**. More explicitly, the form under discussion is grammatical in any $L(G)$ where $AABCD$ is grammatical, and this $L(G)$ happens to be English on more specific conditions as follows:

i. $A \rightarrow colorless, A \rightarrow green, B \rightarrow ideas, C \rightarrow sleep, D \rightarrow furiously$

ii. $A \rightarrow revolutionary, A \rightarrow new, B \rightarrow ideas, C \rightarrow appear, D \rightarrow infrequently$

Properties of natural language syntax are more determined by the system of "selectional" rules

like (i) and (ii) than by the system, called the "base", of rules by which abstract, meaning- and phonology-free forms like *AABCD* are generated. In any case, no justification is provided for the notion that the system of base rules (or principles and parameters) is independent of the system of selectional rules that include (i) and (ii). In a sense, one of the goals of my framework is eliminate the base component altogether by replacing it by an "enriched" system of selectional rules.

6.   There is a complication. Johnson's (1987) and Lakoff's (1987) versions of image schemas are slightly different; the former is more compatible with our view than the latter, because it overlaps with schemas in the sense of Piaget, on the one hand, and in the sense of Arbib (1989) and his collaborators.

7.   This also shows how my approach differs from most approaches in cognitive linguistics, where the aspect of (linguistic) **comprehension,** a basic task of a hearer, is more emphasized.

8.   This idea may have some conceptual link to so-called Church-Turing thesis. It claims that no distinction need not, or should not, be made between machines, or rather computational processes, performing **the same** algorithm. Of course, a weaker form of this thesis is necessary to guarantee that the two results of the different computations in given two calculators are **the same.** More concretely, the result of 2 + 4 in your mind and the result of calculating 2 + 4 on a computer need not be the same without the aid of the thesis. So, the matter is not simple, because we have a dilemma. The essential claim of the Church-Turing thesis cannot be taken literally, but if it makes no sense, it is also nonsensical. Though controversially, I find it rather reasonable to assume that this dilemma will be solved by making algorithms more device-dependent instead of making them more abstract and implementable.

9.   It should be noted that the crucial notion of 'generate' (and therefore 'generation') has undergone drastic conceptual changes since its introduction. In earlier frameworks, the term was used as a synonym of 'define', or 'determine' in a mathematical sense. But in recent frameworks, it is obvious that the same term is used as a synonym of 'produce', or 'yield'. This shift of usage makes it obscure as to what aspect of grammar is generative, as Pullum (1996) remarks.

10.   Some generative linguists propose, analogically, to identify this double-wheeled mechanism as an analogue of (**biological**) **evolution,** probably under the strong influence of the ideas expressed in Richard Dawkins's books (e.g., 1986). I agree that there is far from superficial similarity between the two processes. In fact, the classical notion of **survival of the fittest** has a certain conceptual relevance to the conception of derivation if it is understood as a survival-like process: only one "winner" form and many "loser" forms. This analogy is interesting, but is sure to face a lot of problems, e.g., What is the counter part of selectional pressure? I do not remark on this sort of "incompleteness". Rather, I want to remark that such an analogy, if taken seriously, should force us to abandon a number of basic assumptions in generative linguistics. First of all, generative linguistics could not remain a **competence** theory. Trivially, generative linguistics must become a **performance** theory that incorporates the effects of processing time through selection (logical analogue of selection). More generally, it makes no sense to compare evolution to "competence" of life. A better conception is, of course, that evolution only can follow the abstract trajectory in an n-dimensional space that is allowed, but not determined, by the potential of life. What determines the trajectory is of course "noises" that each individual or gene encounters. Thus, if the notion of competence makes sense, then it is as this notion of potential of life rather than evolution itself.

11.   Bates and Elman (1993) is an excellent introduction to connectionist ideas. See Bechtel and Abrahamsen (1990) for more detailed historical background.

12.   More recent terminology is I-language (for grammar) and E-language (for language). See Chomsky (1986*b*, 1988) for discussion. It is clear what really matters is the conceptual distinction, rather than the terminological one.