Papers in Linguistic Science
No.2(1996), pp.87-117.

Pattern Matching Analysis
—Toward Cognitive Syntax—

Kow Kuroda
Kyoto University
kuroda@hi.h.kyoto-u.ac.jp

1. Introduction

I want to present in this paper a preliminary sketch of a new, systematic method for syntactic descriptions that I am developing now under the name of **pattern matching analysis**. The basic idea came to me when I read Lakoff (1993) in which he presents the framework of "cognitive phonology," on the one hand, and Karttunen (1993) in which he advocates "two-level (declarative) rule systems," on the other. There are, of course, far from trivial differences in Lakoff's and Karttunen's claims. Lakoff discusses the matter from a PDP-conscious point of view, and Karttunen from a more orthodox computational point of view, though he, of course, advances to parallel computation.[1] Nevertheless, their claims converge at an interesting point: "derivations" are dispensable if correspondences (e.g., between underlying and surface forms) are stated both at input and output.[2] I find this was a very interesting idea that deserves import into syntactic descriptions. Trying in and out to give body to this idea, the method, pattern matching analysis, grew out. This method is, as we will see below, specialized to describe multiple correspondences among as many surface forms as we wish to compare to provide topological structure of pattern network. I believe this method is not only useful but also provides us a new way of viewing syntactic phenomena. This paper is a sort of status report of this new approach.

1.1 What is *Cognitive Phonology*?

Let us begin by looking at how Lakoff's (1993) cognitive phonology is different from generative phonology.

Generativists like Halle and Clements (1983, p. 121) claim that Mohawk surface phonetic form D_6 in (1) is "derived from" underlying form D_0 in (1) by six generative rules, $R_1 - R_6$, of the form $X \Rightarrow A/Y\_\_Z$ in (2) applied in that order. D_i indicates that it is *i*th derivational step.

(1)

D_0:	y	e	+	ʌ̃	k	+	h	r	e	k	*	+	ʔ	#	[=underlying form]
D_1:	y	*		ʌ̃	k		h	r	e	k	*		ʔ	#	[by R_1]
D_2:	y	*		ʌ̃́	k		h	r	e	k	*		ʔ	#	[by R_2]
D_3:	y	*		ɔ̃́	k		h	r	e	k	*		ʔ	#	[by R_3]
D_4:	y	*		ɔ̃́	k		h	r	e	k	e		ʔ	#	[by R_4]
D_5:	y	*		ɔ̃́	k		h	r	e	g	e		ʔ	#	[by R_5]
D_6:	y	*		ɔ̃́	k		*	r	e	g	e		ʔ	#	[by R_6]

(2)
R_1 Vowel deletion: $V \rightarrow * / \_\_V$

R_2 Stress Assignment: $V \rightarrow [+\text{stress}] / C_0 \_\_ C_1 V C_0 \#$

R_3 Vowel change: $[\tilde{\text{ʌ}}] \rightarrow [\tilde{\text{ɔ}}] / \_\_ \begin{bmatrix} -\text{syll} \\ -\text{ant} \\ +\text{nasal} \end{bmatrix}$

```
R4            Epenthesis:   * → e / C ___ ʔ #
R5   Intervocalic Voicing:   [+stop] → [+voiced]/V___V
R6             h-deletion:   h → * / ___ r
```

where null elements are indicated by * for readability.

In Lakoff's cognitive phonology, by contrast, the same data is characterized on a completely different basis. He outlines his approach as follows.

(3) Architecture of Cognitive Phonology:

i. "Cognitive phonology characterizes correspondences between morphemes (as stored in the mind) and phonetic sequences. We will refer to those as the morphemic and phonetic "levels" — *M-level* and *P-level*, respectively. In addition, we posit one intermediate level, the phonemic level, at which among other things, constraints on word level phonology are stated; we shall call this the *W-level*." I take these to be minimal collection of necessary dimensions of phonological structure;[3] I also think they are all that is necessary." (*ibid.*, p. 118)

ii. "Cognitive phonology is set within a general autosegmental phonological framework. I assume that the ultimate formal framework for this approach will be a form of PDP connectionism [...] which is a simultaneous constraint satisfaction system." (Lakoff 1993, p. 118)[4]

(4) Mechanism of Cognitive Phonology:

i. "Generative rules (by which I mean rules of the sort familiar to us from Chomsky and Halle 1968) are replaced by *constructions*, which state well-formedness constraints within levels and correlations across levels." (*ibid.*, p. 118)

ii. "In the default case, there is identity across levels. Cross-level constructions override such defaults. Other default cases, both language-specific and universal, are possible." (*ibid.*, p. 118)

iii. "Cross-level constructions are *direction-neutral*, and are indeed to be used directly in either production and recognition." (*ibid.*, p. 118)

iv. "Constructions combine by *superposition*. That is, each construction imposes a set of constraints, and the constraints of the various constructions are simultaneously satisfied." (*ibid*, p. 118)

Under these general assumptions, Lakoff claims that the correspondence between the underlying and surface forms in (1) can be expressed in terms of correspondence as stated in (5).[5]

```
(5)   M:  y  e  +  ʌ̃  k  +  h  r  e  k     +  ʔ #
             |     |
                                                    C1 and C2 satisfied
      W:  y  *     ʌ́  k     h  r  e  k  *     ʔ #
                   |        |        |  |           C5–C8 satisfied
      P:  y  *     ɔ́  k     *  r  e  g  e     ʔ #
```

In this case, the correspondence between M- and W-level forms results from the "simultaneous satisfaction" of two constructions, C_1 and C_2, that Lakoff states as in (6), on the one hand, and the correspondence between W- and P-level forms from C_5–C_7, on the other.

```
(6) C1  M:  V    +    V             C2  W:  X C0 V C0 #; if X=V, then X=[+stress]
            |
        W:  *

               ┌ −syll ┐
    C3  W:  ʌ̃  │ −ant  │            C4  W:  h    r
               └ +nas  ┘
            |                               |
        P:  ɔ́                           P:  *

    C5  M:  ʔ    h    r             C6  P:  V X V; if X=[+obstruent], X=[+voiced]
            |         |
        W:  *         *

    C7  W:  [ ]  *   [ ]
             |   |    |
        P:   C   e    ʔ    #
```

According to Lakoff (1993, pp. 120—121), C_1 is Vowel Omission, stating that "the first of two consecutive vowels at level M does not appear at level W." C_2 is Stress, stating that "At W, a vowel that is penultimate is stressed." C_3 is Vowel Change, stating that "W-level [ʌ̃] corresponds to P-level [ɔ́]." C_4 is *h*-Deletion, stating that "an /h/ preceding an /r/ at W does not appear at P." C_5 is Cluster Simplification, saying that "a sequence /ʔhr/ at M appears as /h/ at W." C_6 is Voicing, saying that "at level P, intervocalic stops occur voiced." C_7 is Epenthesis, saying that "when C precedes ʔ# at level W, an /e/ absent at level W intervenes at level P (The W-level representation "[ ] [ ]" indicates that the /C/ and /ʔ/ are adjacent at W)."

Furthermore, constructions C_1–C_7 fall into two major categories: "cross-level" and "intralevel" constructions. More specifically, C_1 in (6) is a construction across M- and W-levels. C_2 is a W-internal one. C_3, C_4, C_5, and C_7 are constructions across W- and P-levels. C_6 is a P-internal one.

Crucial points are: (i) representations at different levels are "associated" to one another, rather than derived from one to another, and (ii) M–W and W–P correspondences in (5) embody no intermediate steps for derivation. M- and W-level representations are associated to one another in a "single" step (superposition of C_1 and C_2), and so is the association of W- and P-level representations (superposition of C_3–C_8);

On this basis, Lakoff asserts that "the Mohawk data do not in themselves show the need for ordered rules and derivations, since they can be accounted for with equal generality using cognitive phonology. What this shows is that ordered rules and derivations may well be an epiphenomenon of assuming the existence of generative rules that apply in order and manipulate symbols one step at a time" (*ibid.*, 123).

1.2 What Makes Cognitive Phonology "Cognitive"?

The replacement of (1) by (5) is not merely a technical improvement, or a "notational variant" of the same conception. Let us see why.

Lakoff begins his paper by posing the following two questions about the cognitive reality about generative phonology.

Q_1: Do people go through these step-by-step derivations in their heads every time they pronounce a sentence?

Q_2: What makes derivations cognitively real if the immediate stages never occur in the minds of speakers?

He recollects that he got an absolutely negative answer to Q_1, one the one hand, and "performance-competence" answer to Q_2, on the other.

Lakoff argues, then, that such justification is quite dubious if we take it into consideration that "[c]onnectionist cognitive science suggests that the brain *is* the mind and that anything that is cognitively real is represented in the brain" (*ibid.*, p. 117; see also McClelland *et al.*, 1986; Rumelhart *et al.*, 1986; Davis, *ed.* 1992 for parallel distributed processing (PDP) framework). In short, Lakoff's cognitive phonology is cognitive if (or as far as) PDP models are cognitive.

1.3 If Phonology Can Be Done "Cognitively," So Must be Syntax

I think that Lakoff's challenge should affect on the study of syntax. In this sense, my attempt is to develop a framework of "cognitive" syntax in conformity with Lakoff's cognitive phonology. But before going on into details directly, let us turn for a while to another, seemingly "less cognitive" aspect of the issue.

1.4 A Strange Convergence

Since Karttunen's (1993) discussion may seem to be too technical to most linguists, let me skim his main points here.

Karttunen suggests that what Lakoff calls constructions are "two-level rules" in the sense of Koskenniemi (1983). Two-level rules are statements of correspondence which may specify conditions at input and output at the same time. By contrast, he points out, generative "rewriting" rules are one-level rules which may not specify conditions at output. He remarks that there is no difference in weak generative capacity between one-level and two-level rule systems. But, by virtue of this equivalence, two-level systems are more useful, since they can dispense with rule ordering altogether (of course, at the expense of more complexity in statements). Karttunen remarks, drawing on other important works by Johnson (1972) and Kaplan and Kay (1981), that this almost miracle is attainable as far as cross-level correspondences constitute "regular" or "type 3" relations,[6] and he calls this constraint "finite-state constraint."

I said just above that Karttunen's discussion is "less cognitive." But please do not take this literally. Its real meaning is that it is "indirectly cognitive," and, for evidence, Lakoff's and Karttunen's arguments against generative phonology converge at an interesting point.

Some of us might be surprised if I say that Lakoff's "connectionist" and Karttunen's "orthodox" computational approaches converge; but they really do, if I do not misunderstand. Their claims converge at their emphasis on the reality of computation, irrespective of neural or not. In this respect, it is necessary to recall that neural networks are not only cognitively real but also computationally real. The reason is obvious: neural nets are computational systems in and by themselves.

1.5 Finite-State Syntax?

It is of course very risky to assume that syntactic relations can be basically handled in terms of regular relations, since the descriptive power of regular expressions is very restricted. This assumption in effect runs counter to the motivation for transformational analysis in the period of early generative grammars. Chomsky (1957, 1975) conducted arguments to show that natural language, e.g., English, is somewhere within context-sensitive (type 1) language, if not unrestricted (type 0) language.[7] It follows that most of natural languages, if not all, are not even context-free (type 2), let alone regular (type 3).[8]

We have a crucial question here: What does it mean that natural language is a type n language (n ranges between 0 and 3, possibly fractional)? More exactly, if our "competence" can be identified with a type 1 grammar in formal sense, does it mean that our "performance" produces a type 1 language? And, if not, is there any significance in showing that our competence equates with type 1 grammar, while our performance equates with type 2, or 3 grammar?

We have a good piece of evidence to show that finite-state constraint is also working in syntax. Elman (1992) argues that so-called "subjacency condition" (Chomsky 1973), which, roughly speaking, blocks movements across more than two bounding nodes (NP and $\bar{S}$ (= CP)), could never be operative if grammar performs infinite, "real" recursions. Two problems arise: first, why an element in pushdown stores is "visible" if real recursion takes place. Second, why there can ever exist such an upper limit on the number of bounding nodes.

Elman argues that, if subjacency condition is ever operative, which we assume is true, then the fact by itself proves that grammar does not perform real recursion. What it performs is "leaky" recursion. We can even draw another conclusion: subjacency condition, if anything, is not part of "competence" theory, but part of "performance" theory, since competence will never know any such upper limit. In short, it is a contradiction both to require, on the one hand, that the grammar of English permits infinite recursion in terms of competence, and to require, on the other, that the subjacency condition is true, if the latter condition is part of "universal grammar."

In this respect, Elman's (1992) argument for finite means echoes with Karttunen's finite-state constraint. Indeed, finite-state property may have an important bearing on the behavior range of neural networks. This is, I believe, a piece of evidence to show that some sort of Karttunen's finite-state constraint is operative in syntax, too. It is not unlikely that most, if not all, of syntax can be characterized by finitary devices.

1.6 And We Will Have "Cognitive" Syntax Some Day ...

Since cognitive phonology is a both neurally and computationally realistic performance theory of phonology, if we have a neurally/computationally realistic performance theory of syntax, it must be "cognitive syntax." Cognitive syntax will be part of "cognitive grammar," which Fauconnier (1994), Lakoff (1987), Langacker (1987, 1991), and Yamanashi (1995) are representatives of.

But there is not cognitive syntax yet. Without any malicious intention, I would like to say that cognitive grammar lacks syntactic component, while we already have semantic and pragmatic components almost in maturity.[9] I wish to fill this gap somehow, the best that I can.

I shall talk in what follows as if there *is* already cognitive syntax as a virtual whole of syntactic theories that we will have some day, to which the following sketch of an approach is intended to be a contribution.

2. How "Syntax" Is Conceived in Cognitive Syntax

The primary goal of cognitive syntax is, like Lakoff's cognitive phonology and Karttunen's two-level rule systems, to reduce, or eliminate if possible, the role of derivations from the area of syntax. But this would be a mere day dream without an alternative method, as powerful as derivational analysis. In this first preliminary section, I want first to settle some conceptual and terminological issues that may cloud the way to reach that goal.

2.1 Formations and Correspondences

Syntax is not a monolithic object. Most syntactic studies seem to conflate two essentially different kinds of objects, namely **formations** of linguistic expressions (e.g., sentences) and **correspondences** among formations. Simply put, formations are linear arrangements of units on the selectional basis. Linguistic forms are results of formations. Sentence *John sings strange songs every day*, for example, is a result of formation over the set of units U= {*John*, *sings*, *strange*, *songs*, *every day*}, and, on smaller scales, of U'= {*John*, *sing*, *-s* (for tense), *strange*, *song*, *-s* (for plural), *every*, *day*}, though relevant scales are not coherent. Thus, more generally, if we let $U= \{u_1, ..., u_n\}$, then we could say that forms (e.g., sentences, phrases, words) of a language are appropriate linear arrangements over U on respective scales of units. Of course, it is lexical selection, rather than arrangement by themselves, that determines a set of units relevant to formations.

It is clear, however, that formation is not everything of syntax. Consider the relation between *John sings strange songs every day* and *What John sings every day are strange songs*. None of us would deny that this relation is also part of syntax. But the relation between the two sentences is not of formation. It is another kind of relation about "pairs of formations." More exactly, this kind of relation need not be confined to pairs of formations. Just like a formation is defined as a relation over an appropriate set of units $U=\{u_1, .., u_n\}$, this, the latter kind of relation is a relation over an appropriate set of formations $F=\{f_1, ..., f_n\}$, usually of sentences. This relation is exactly what we will call correspondences among forms.

2.2 Nature of Correspondence

We will not be concerned with formations in this paper. We will not be concerned with all kinds of correspondence, either. Let me first make this point clear.

Correspondence that can be handled in syntactic terms has two aspects: correspondence **between form and meaning**, on the one hand, and correspondence **between one form and another.** Let us call the former kind **form-meaning (fm) correspondence**; and the latter **form-form (ff) correspondence.** To make this point clearer, suppose we have a pair of sentences, s and s'. Obviously, s and s' are in the relation of fm correspondence. There are, in addition, correspondences between f and f', one the one hand, and between m and m', on the other. Such two-dimensional relationship can be illustrated in Figure 1.

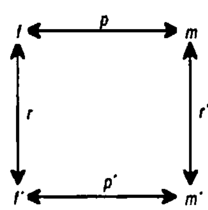

Figure 1

The two fm correspondence in this figure, namely $p(f, m)$ and $p(f', m')$, mediate an fm correspondence, $r(f, f')$, and a **meaning-meaning (mm) correspondence**, $r'(m, m')$.

Note incidentally that what Lakoff (1993), in conformity with Goldberg (1995), Langacker (1987), calls grammatical constructions are fm correspondences in our sense, not ff correspondences. Our ff correspondences will fit the notion of "network of constructions," as will become clear. With the distinction above, what we will try to deal with in cognitive syntax is ff correspondences **of whatever sort.**

It will be helpful to give names to the two dimensions of correspondences. As Figure 1 illustrates, we have dimensions P (p, $p' \in P$) and R (r, $r' \in R$). P is the dimension in which fm correspondences are definable. *R* is the one in which ff and mm correspondences are definable. We may furthermore assume that *P* is the dimension of **predication**, and, for this reason, we may also assert, if we adopt Langacker's (1987) terminology, that *P* is the dimension of **symbolization** in which *m* is symbolized by means of *f*.

For illustration, let *f*= *John sings strange songs every day* and *m*= JOHN SINGS STRANGE SONGS EVERY DAY; *f*′= *What John sings every day are strange songs* and *m*′= WHAT JOHN SINGS EVERY DAY ARE STRANGE SONGS. For *f*, we have such pieces of fm correspondences as p_1(*John/__ sings strange songs every day*, JOHN/__ SINGS STRANGE SONGS EVERY DAY), ... , p_5(*every day/John sings strange songs __*, EVERY DAY/JOHN SINGS STRANGE SONGS __). Also for *f*′, we have similar pieces of fm correspondences, e.g., p'_1(*what/__John sings every day are strange songs*, WHAT/__JOHN SINGS EVERY DAY ARE STRANGE SONGS).[10] These pieces of fm correspondence is clearly what Langacker (1987, 1991) calls "symbolic links."

So far so good, but it is quite unfortunate that we don't have a proper name for *R*. It seems that this lends to the claim that *r*'s are independent of *r*′'s. Most of disputations over the so-called "autonomy of syntax," which I find are totally sterile, will be avoidable only if we identify *R* so that *r* and *r*′ are two "modes" of *R*'s manifestation. Let us call *R* **(syn)tactic dimension**. In this regard, *r* may be called any one of phono(syn)tactics, morpho(syn)tactics, morphophono(syn)tactics, on the one hand, and *r*′ may be called semo(syn)tactics, or **conceptualization**, if we rely on Langacker's (1987) term again.

2.3 Direct and Indirect Motivations of Correspondence

For clarity, however, it will be helpful to say that ff correspondences are "directly" motivated by *r*, and "indirectly" by *r*′. For the same reason, mm correspondences are directly motivated by *r*′, and indirectly motivated by *r*.

Importantly, if we understand Figure 1 correctly, it is clear that *r* and *r*′ are dependent of each other. Indeed, every *r* is a reflection of *r*′, on the one hand, and *r*′ is a reflection of *r*, on the other. We may claim that (i) there cannot be any morphotactics without semotactics—against those who set priority of syntax over semantics, and likewise (ii) there cannot be any semotactics without morphosyntactics, either—against those who set priority of semantics over syntactics.

3. Pattern Matching Analysis

Based on the notion of ff correspondence defined above, I will in this section introduce the basics of pattern matching analysis. The method is intended to be as much effective as derivational analysis. But before going on to details directly, it will be helpful to make a few remarks.

3.1 Why "Method"?

We should first distinguish two things explicitly: our "research program" in the sense of Lakatos (1970, 1977)[11] and our "method" of analysis. The name of our research program is "cognitive grammar," which Fauconnier (1994), Fillmore, Kay and O'Corner (1988), Lakoff (1987), Langacker (1987, 1991), Yamanashi (1995) are representatives of. The works mentioned here and other works related to or stimulated by them (mine included) have one intellectual attitude in common: they have more interest in "performance" theory than "competence" of grammar.

Whatever research program we may work within, our investigation inevitably has another technical aspect as far as it is a scientific enterprise. We need a method, or algorithm, more or less explicitly defined and proven to work in an expected way. I think this aspect is independent of what research program we participate in. As Lakatos correctly pointed out, research programs determine the "domain" of research, and, more importantly, researcher's way of looking at phenomena. To put it somewhat technically, our observations are "theory-laden" (Hanson 1958).

But methods are relatively independent of the choice of a research program. It is common that the same method is used among "incommensurable" paradigms.[12] So, to put it extremely, we may make use of methods in generative grammars, e.g., tree-based analysis, in cognitive way, though I do not attempt to do it in this paper.

To the method that I will make use of throughout, I will refer as **pattern matching method for syntactic analysis**, or simply as **pattern matching analysis.**

Again, the name of our research program is "cognitive grammar," and what we are trying to develop as an integral part of cognitive grammar is "cognitive syntax." The name of the method that we make use of to achieve this end is pattern matching analysis.

3.2 Why "Pattern Matching"?

Pattern matching is an old-fashioned term, as Yamanashi (p.c.) rightly pointed out to me. He recommended me to use other term if possible. But I decided to retain it, since, as far as I know, linguistics has never exhausted real potentials of pattern matching.

The term pattern matching, I think, has at least two advantages. First, it is a term suggestive enough to remind us that it has certain relevance to "pattern association" which is used commonly in the PDP framework to denote the activity at neural scale (Rumelhart, *et al.* 1986; McClelland, *et al.* 1986). Although the meaning of "pattern" is not exactly the same in pattern matching and pattern association, I find more conceptual similarity than dissimilarity. Second, the term makes explicit its relevance to "regular relations," which result from finite-state constraint in the sense of Karttunen (1993). For these reasons, I decided to retain this old-fashioned term.

3.3 An Overview

Pattern matching analysis is, first and foremost, a comparative method to capture properties of ff correspondences. To see this, it is helpful to look at Figure 1 again.

What we usually call sentences, or more generally **linguistic expressions**, are correspondences of the form $p(f, m)$, which are, in our terms, fm correspondences. Of course, it is possible to state the correspondence as symbolizations of m in terms of f, but it is important to note here that not all fm correspondences are linguistic expressions. An fm correspondence is a linguistic expression only if f is a "linguistic" form, though without a clear definition of linguistic.

It is also important to note that formal strings of symbols without meaning associated with them are **not** linguistic expressions in our sense, let alone being sentences. So, in our framework, the notion of "grammaticality" (Chomsky 1965) plays little role. If there is an analogue to the notion, it must be that of **formations.** But it is very likely that formations of f and m are mutually dependent of each other. So, it is difficult, if not impossible ideally, to isolate f-formations from m-formations, and vice versa. Metaphorically put, trying to achieve it would suffer from the same kind of dilemma as we know Sheylock, a merchant of Venice, experienced when he was allowed to take one pound of breast meat, but without any blood bleeding. So, isolation of f-formation from m-formation is, in short, only ideally possible, but practically impossible.

Let us now turn to the details of ff correspondences. As noted above, any ff correspondences have direct and indirect motivations. An ff correspondence is directly motivated in terms of r, and indirectly in terms of r'. Pattern matching analysis tries to describe direct motivations among formations.

3.4 Partial Matches and Mismatches

Loosely speaking, we may identify forms as patterns, though their exact meaning are different. We will return to this issue in §4.7.

Given this rather loose idea of patterns, we need furthermore the notions of **partial matches** and **mismatches** to state ff correspondences in terms of pattern matching. To make this point clear, let us consider the ff correspondence between (7a) and (7b).

(7) a. They loaded bricks in(to) the truck.
    b. They loaded the truck with bricks.

Leaving aside the fm and mm correspondences, their ff correspondence can be illustrated as in (8), where *F* and *G* are "canonicalized" form of (7a) and (7b), respectively.

(8)

F:	They	loaded	bricks	in	the truck	•	•
			|	|		|	|
G:	They	loaded	•	•	the truck	with	bricks

What (8) embodies are: there are partial mismatches between *F* and *G* with respect to *bricks*, *in*, *with*, and there are partial exact matches with respect to other positions.

Do not be so curious about the details of (8), especially about the strange elements, symbolized by " • " to which I will refer as "phantoms." (8) is presented here only for the sake of illustration. The principal aim of this section is to give foundations to describe ff correspondences such as in (8). I will present the basic assumptions and relevant notions to make sense of (8) in what follows, step by step.

3.5 Three Remarks

Three remarks should be made here, however, to avoid possible misunderstandings of what we will see. We assume specifically:

(9) I. Correspondences like (8) claim no **identity in meaning** between *F*- and *G*-forms. The contrary is true: since *F* and *G* are different forms, they can never be identical in meaning, either. I will discuss this issue under the notion of **default difference** in §4.5.

II. Correspondences like the one above claim no derivation from one form to the other, since in (8) none of *F* and *G* is an "underlying" form.

III. Correspondence like the one above does not rely on the notion of "levels."

The last two assumptions are crucially different from the case of phonology that we have seen in §3. Phonological correspondences are correspondences between forms at different levels, say, morphophonemic, phonemic, and phonetic. Cognitive phonology, for example, succeeds in elimination of "derivations" at the expense of crucial reliance on the identification of phonological levels. Indeed, nothing will remain if there are no such levels.

3.6 Levels as Scales of Units

It is possible to think that ff correspondences are level-internal correspondences rather than cross-level ones. But this brings about a number of questions: What is the level of syntax? Is there a level of syntax, or, alternatively, are there levels of syntax? Do surface forms have underlying forms? To avoid such embarrassing problems, I want to take a radical position. I simply assume that there may be an arbitrary number of levels. More specifically, we identify levels as "scales" on which we can recognize different sorts of "patterns."[13] In this view, what we call syntax does not correspond to some unique level. Rather, syntactic relations are interactions of a variety of scales, though relatively larger ones. *John sings strange songs every day*, for example, consists of U_S= {*John sings strange songs every day*}, on the scale of *S* (for sentences), but it consists of U_P= {*John, sings, strange songs, every day*} on the scale of *P* (for phrases); it consists of U_W= {*John, sings, strange, songs, every, day*} on the scale *W* (for words); of U_M= {*John, sing, -s, strange, song, -s, every, day*} on the scale of *M* (for morphemes). Note that *John*, for example, has the same form on *P*, *W*, and *M* scales. Note also that these by no means exhaust the possible scales: there may be intermediate scales such that the sentence in question consists of U_S'= {*John sings strange songs, every day*}, or U_P'= {*John, sings strange songs every day*} or U_P''= {*John, sings strange songs, every day*}. Patterns have **matching scales** of their own. Thus, we understand the question of linguistic levels as the question of constraints on scaling, which should provide answers to the question of why there may not be scales like *H* on which the sentence consists of U_H= {*John sing, -s strange, songs every day*}. This suggests that "consitituency" is an effect of scaling, and need not be determined "uniquely."

Bearing the three remarks stated in (9) in mind, let us turn to the details of pattern matching analysis.

3.7 Canonicalization of Partial Matches/Mismatches

It is important to note that virtually any pattern matching that we will treat in terms of ff correspondence is "partial." To show this is quite easy. Let patterns *F* and *G* be *ABCD* and *ACBD*. Obviously, matching between *F* and *G* is not total, simply because they are different patterns. A pattern matching is total if and only if two patterns match exactly. So, trivially, no pattern matches anything but itself.

To meet the requirement of finite-state constraint (Karttunen 1993), we also want **bidirectionality** (= what Lakoff calls **direction-neutrality**) to hold between *F* and *G*, so that ff correspondences are descriptions of the relations between two, or more, surface forms. This requirement necessitates, however, that matching is total: otherwise, bidirectionality can never hold. So, we need some special procedure to make a partial matching total. This procedure will be called **canonicalization**.

3.8 Phantoms

One obvious, and I believe plausible, way to achieve this is to make use of a sort of "wild card" that matches anything. Thus, (10) is a possible, though not unique, canonicalization of pattern matching between *F* =*ABCD* and *G* =*ACBD*, where the special symbol " • " is a wild card to match any subpattern. Let us call them "phantoms."

(10) Pattern Matching of *F* and *G* (Canonicalization I)

F:	*A*	•	*B*	*C*	*D*
↓↑	↓↑	↓↑	↓↑	↓↑	↓↑
G:	*A*	*C*	*B*	•	*D*

where bidirectionality is indicated by "↓↑".

In cases like this, two patterns *F* and *G* match partially in that *A*, *B*, and *D* on one side have exact matches on other side, but *C* doesn't. The occurrences of *C*, namely *C*/*AB*__*D* (part of *F*) and *C*/*A*__*BD* (part of *G*), have no "exact" match on the other side. For cases like this, we say that *F* and *G* "mismatch" (only) with respect to *C*.

3.9 Two (and Only Two) Principles of Canonicalization

The canonicalization procedure to obtain (10) tacitly assumes two specific principles, which can be stated explicitly as (11A) and (11B), respectively.

(11) A. *Maximal Partial Matches*
Pattern matching must be so canonicalized that there be a maximal number of partial matches.

B. *Minimal Partial Mismatches*
Pattern matching must be so canonicalized that there be a minimal number of partial mismatches.

As we will see, most, if not all, of our interesting results are consequences of these principles (11A) and (11B).

We will never claim, however, that the principles are part of "universal grammar" (Chomsky 1965, 1981, 1995), partly because to claim so is to claim virtually nothing, and partly because we can interpret that they are rather mere consequences of the mind/brain's inherent tendency toward "optimal" states.[14]

Some of us may feel that (11A) and (11B) are equivalent. But it is not true. For evidence, we need (11A) to exclude cases like (12), one the one hand, and (11B) to exclude cases like (13), on the other.

(12)	*F*:	*A*	*B*	*C*	*D*
		↓↑			↓↑
	G:	*A*	*C*	*B*	*D*

(13)	*F*:	•	*A*	•	*B*	*C*	*D*
	↓↑	↓↑	↓↑	↓↑	↓↑	↓↑	↓↑
	G:	*A*	•	*C*	*B*	•	*D*

It is easy to see that (12) can have more partial matches (with respect to either *B* or *C*), and consequently does not satisfy (11A). By contrast, (11B) can have less phatoms, since • *A*:*A* • can be reduced to *A*:*A*.

3.10 Ambiguity of Canonicalization

It is important to note, however, that (10) is not a unique canonicalization. Principles (11A) and (11B) allow another possibility, namely (14).

(14)	*F*:	*A*	*B*	*C*	•	*D*
	↓↑	↓↑	↓↑	↓↑	↓↑	↓↑
	G:	*A*	•	*C*	*B*	*D*

Note that, although (10) and (14) are symmetrical, we have to think of them as two different pattern matchings. An obvious reason is that they specify different things. (10) specifies that *F* and *G* have analyses *A*•*BCD* and *ACB*•*D*, respectively, while (14) specifies that they have analyses *A B C*•*D* and *A*•*CBD*, respectively.

This point will be made clearer if we look at (15), where *E* and *E′* are different.

(15)

F:	*A*	•	*B*	C_2	*D*		*F*:	*A*	B_1	*C*	•	*D*
↓↑	↓↑	↓↑	↓↑	↓↑	↓↑		↓↑	↓↑	↓↑	↓↑	↓↑	↓↑
E:	*A*	C_1	*B*	C_2	*D*		*E′*:	*A*	B_1	*C*	B_2	*D*
↓↑	↓↑	↓↑	↓↑	↓↑	↓↑		↓↑	↓↑	↓↑	↓↑	↓↑	↓↑
G:	*A*	C_1	*B*	•	*D*		*G*:	*A*	•	*C*	B_2	*D*

The difference of *E* and *E′* is enough to distinguish between canonicalizations in (10) and (14). Thus, any decision between them will be arbitrary unless it is based on other independent reasons.

3.11 Meaning of Bidirectionality

We have no proper interpretation of bidirectionality, however. More specifically, pattern matching in canonicalized forms (10) and (14) can be interpreted in one of the following three ways in (16).

(16) A. Correspondences state only input condition
B. Correspondences state only output condition
C. Correspondences state both input and output conditions

Our rejection of levels makes interpretations A and C invalid, because there cannot be input conditions of whatever sort. If there is an analogue to input condition, it will be constraints on formation. But, it is clear that well-formedness of *F* has nothing to do with that of *G*, even if *G* is unacceptable, and vice versa. For this obvious reason, we are safe to decide to take interpretation B. That is, pattern matching state output well-formedness on both sides, without stating input condition.

3.12 Notational Convention

It is obvious that (10) is redundant for notational purposes, especially when we see it in light of **default specification**. In fact, notations like (10) can be simplified if we assume that exact matches are default. Thus, for expository purposes, we have such notation as (17) for abbreviation of (10), where only partial mismatches are specified as they are, assuming that partial matches are default. " | " indicates that there is a partial mismatch between *F*- and *G*-forms.

(17)

F:	*A*	•	*B*	*C*	*D*
		|		|	
G:	*A*	*C*	*B*	•	*D*

It is easy to see that (17) shares a lot of properties with Lakoff's "constructions" presented in (6). But there is a crucial difference: "default identity" in the sense of Lakoff holds only of *A*, *B*, *D*. But we extend the notion of default identify so that it weakly holds also of *C*. It follows that in canonicalizations like (17), there is no insertion (=epenthesis) or deletion in real sense.

3.13 Pattern Matching Embodies No "Movements"

The *F*–*G* correspondence in (17) is the general scheme of pattern matching which is, as we will see below, powerful enough to enable us to state any kind of ff correspondences. But, since correspondences like (17) may appear as if they describe "movement" of *C*, we have to make a few remarks to reject this conception.

It is sure that pattern matching like (17) looks like a derivation, relative to our point of view. Indeed, if viewed as a mapping from *F* to *G*, (17) gives us the impression that *C* is moved leftward. But the contrary is also true. It gives us the impression that *C* is moved rightward, if viewed as a mapping from *G* to *F*.

Such derivational interpretation involves two misunderstandings in it. First, most crucially, it is incompatible with our rejection of "underlying" forms specified in (9): neither *F* nor *G* is the underlying form of the other. Rather, they are both surface forms of their own status. Second, even if the notion of underlying forms is abandoned, movement of *C* in (17) is an apparent effect at best. The reason is as follows. What (17) minimally specify is rather co-occurrence restrictions such that in *F*, *C* is absent in *A*__*B* and present in *B*__*D*, on the one hand, and in *G*, *C* is present in *A*__*B*, and absent in *B*__*D*, on the other. The correct way to capture this is that *C* is not (or may not be) present in *A*__*B* and *B*__*D* at the same time. A straightforward interpretation of this is thus that *F* and *G* are different optimal solutions of the co-occurrence restrictions on $*E=AC_1BC_2D$, in (18), with respect to the presence of *C*.

(18)
$$\begin{array}{llllll} F: & A & \bullet & B & C_2 & D \\ & & | & & & \\ {}^*E: & A & C_1 & B & C_2 & D \\ & & & & | & \\ G: & A & C_1 & B & \bullet & D \end{array}$$

Given *E* is an inadmissible pattern, *F* and *G* are optimal solutions of *E* so that in *F*, *C* is allowed to appear in *B*__*D* (and accordingly *C* is disallowed to appear in *A*__*B*), whereas in *G*, *C* is allowed to appear in *A*__*B* (and accordingly *C* is disallowed to appear in *B*__*D*).

3.14 A Remark on "Underlying" Forms

We have to concede here that there will be no contradiction in saying that *E* is the underlying form of *F* and *G*. It is not clear what we could benefit from such identification, but it is at least compatible with our position to accept this identification only when *E* can never surface as it stands.

3.15 Pattern Matching Is *Not* Transformation

To convince some skeptical readers of the points discussed so far, it will not be a waste of time (and paper) to try to compare the notion of transformation in early generative grammar with pattern matching method in our analysis. Take the case of correspondence between active and passive forms for touchstone.

In early generative grammar, active–passive correspondence was stated in terms of permutation such as in (19), where I follow the notational convention devised by Rosenbaum (1967) (cf. Ross (1986, p. 2)).

(19)	*F*:	NP	V	NP					
		1	2	3					=> Passive
		0	0	3	*be*	2	*-en*	*by* 1	
	(=	NP	*be*	V	*-en*	*by*	NP)		

As noted, permutation requires "index" (e.g., 0, 1, 2, ...) as auxiliary device, since it changes the **position** of units.

By contrast, the same relation will be stated as in (20), if based on our pattern matching analysis by relying on the notions of partial matches/mismatches with phantoms as an auxiliary device, ignoring the alternate possibility of canonicalization to reverse the relation of *F* and *G* regarding the object NP.

(20)	*F*:	NP	V	NP	•	•	•	•	•
		|	|		|	|	|	|	|
	G:	•	•	NP	*be*	V	*-en*	*by*	NP

where some morphological details, especially NP's value for case, are omitted for expository purposes.[15]

It is inevitable that transformation and pattern matching have certain resemblance. But this is simply because of a factual demand: both try to state the relationship of two forms. To put it differently, the two methods of analysis are similar only to that extent.

Though there are many crucial differences to be pointed out, let me emphasize here only a few of them. First, in transformational analysis, passive form (19*G*) is allowed as far as the rule of Passivization (19) generates it, since the relation between them is that of "derivation" of *G* from *F*. In our pattern matching analysis, by contrast, passive and active forms are simply different patterns whose relation is statable as in (20). Second, as a consequence of the first point, *be*, *-en* and *by* are not "grammatical formatives" that have no meanings.[16] The contrary should be true: they must be meaningful elements that indispensably contribute the meaning construction of *G*. We will return to some relevant details of this correspondence in §4.15.

3.16 Phantoms Are Not "Empty Categories"

It is natural if some of us have a question about characteristics of phantoms: Aren't phantoms "empty categories"? The answer is of course negative. Let me specify this.

Phantoms are not real elements of forms: they are "ephemeral," if I could say so, and indeed I call them phantoms for this reason. The reason is obvious:

(i) Phantoms are not necessitated by so-called "phrase markers" like [$_S$ NP [$_{VP}$ V NP]], whether base or derived; and
(ii) Instead, phantoms are necessitated by the interaction of principles (11A, B), which request that there be maximal matches and minimal mismatches in any pattern matching.
(iii) Phantoms appear only when forms are compared; and
(iv) Once requested by principles of pattern matching, phantoms appear anywhere.

In short, motivation for phantoms is not derivational at all.

3.17 No Derivation but Correspondence Among Forms

Let us summarize here the essential points in this section. Pattern matching analysis is a method to replace the derivational view of syntactic relations, especially those that can be characterized in terms of ff correspondences. To take this position, we have such strong reasons as stated below.

A. Derivations are not real. We need additional assumptions (e.g., competence/performance distinction) to give it reality.
B. Derivations presuppose theoretical constructs such as "underlying" forms.
C. Derivations complicatedly ordered make grammar, if any, "unlearnable."

Under these foundations, let us now turn to concrete analyses.

4. Sample Analyses

This section offers a few sample analyses to display the descriptive power of pattern matching analysis. There will be far more sets of facts that our method could handle, but we cannot be of course exhaustive. Furthermore, the syntactic phenomena are selected fairly randomly, without any intention to be systematic. In addition, all analyses are highly tentative.

By doing this, we will examine several kinds of partial mismatches relevant to syntactic descriptions. We will see (i) the notion of **contrast** defined, (ii) **positional** and **lexical** mismatches distinguished, (iii) the notion of **matching scales** introduced, (iv) the notion of **synchronization** required, (v) the notion of **topology of contrasts** introduced, and (vi) the notion of **pattern network** defined.

4.1 Correspondences and Contrasts

It is trivial to note that in ff correspondences, there are **contrasts** (i) wherever there are partial mismatches, and (ii) as many as partial mismatches. The reason is obvious: any pattern exactly matches no pattern other than itself (reflexivity holds).

This means that virtually all correspondences that we treat in terms of pattern matching will have one or more contrasts. We say that contrast is minimal if the number of such mismatches is one. In (21), *F* minimally contrasts with *G* with respect to *always*.

(21)	*F*:	She	is	•	beautiful
				|	
	G:	She	is	always	beautiful

No matter how trivial, we will try to describe seriously this kind of relation in terms of ff correspondence, not because it is the easiest to describe, but because this is one of the possible correspondences that we can find in the syntax of English and many other languages. We will refer to this kind of partial mismatches as **positional** mismatches.

Note however that the meaning of the correspondence in (21) is not like (22).

(22)	*F*:	She	is	$[_{\text{ADV}}\ e\ ]$	beautiful
				|	
	G:	She	is	always	beautiful

What is rejected here is to identify phantoms as so-called "empty categories" (Chomsky 1981).

As explained in §3.16, phantoms are "ephemeral" elements that emerge only when different forms are "compared." To make this point clearer, let us compare the *E–G* correspondence in (23) and the *D–H* correspondence in (24). The two correspondences mean differently: while we need no phantom in *sliced __ all the bagels* in (23), we need one in the same position in (24).

All examples in (23) and (24) are my modifications of those in McCawley (1988, p. 639).

```
(23)  E:  Marvin      •      sliced                all the bagels   carefully
                                                                        |
      F:  Marvin      •      sliced                all the bagels       •
                      |                                                 |
      G:  Marvin  carefully  sliced                all the bagels       •

(24)  D:  Marvin  carefully  sliced       •        all the bagels   carefully
                      |
      E:  Marvin      •      sliced       •        all the bagels   carefully
                                                                        |
      F:  Marvin      •      sliced       •        all the bagels       •
                      |
      G:  Marvin  carefully  sliced       •        all the bagels       •
                      |                   |
      H:  Marvin      •      sliced   carefully    all the bagels       •
```

Although the correspondences do not by themselves tell us any interesting properties of *carefully*, my point is simply this: the one and the same surface form *Marvin sliced all the bagels* should receive different analyses in (23) and (24). More specifically, there may not be a phantom in between *sliced* and *all the bagels* in (23), whereas there must be one in (24). To escape from a contradiction, we say that phantoms are **activated** by contrast. Thus, it is *H* that activates phantom in *slice* __ *all the bagels* in (24).

In a few respects, phantoms are similar to empty categories. Both specify "potentials" for instantiations of lexical items. But, in a sense, phantoms are more abstract than empty categories. Indeed, all phantoms in *F* =*Marvin* • *sliced* • *all the bagels* • specify potential for the instantiation of *carefully*. But they not only need not be empty categories, but also they cannot be. It is in effect a contradiction to say that Adv (or AdVP) does not occur in *slice* __ *all the bagels* in (23), on the one hand, and that Adv occurs in (24), on the other. One might argue that the exact form of *F* is like one in (24), since it is more general. But this is untenable. If phantoms are to be identified as empty categories, then such a strange expression must be permitted as **Marvin carefully sliced carefully all the bagels carefully*, which is instantiation of *carefully* in (24*F*) at all potential positions at once, though *D* is acceptable for interesting reasons.[17]

Also implied by this remark is that phantoms, unlike empty categories, are free from "labels." The existence of phantoms, if activated, and their names are different matters. To put it more adequately, labeling them is an extrinsic matter. We will return to this point in §4.8.

4.2 Relevance to Language Acquisition

It is remarkable that phantoms, presumably the super class of empty categories, can be predictable from the fairly simple and straightforward principles of *Maximal Partial Matches* and *Minimal Partial Mismatches* stated as (11A, B). But this is only a special consequence of the proposed mechanism of pattern matching.

It is a very speculative idea, but I believe we may even think that pattern matching procedure itself constitutes an essential **learning procedure** for language acquisition, though it is a higher-level description of the pattern association that would actually take place on neural scales. Suppose a English learner, not necessarily a child, who knows only that the pattern instantiated by (23*F*) is admissible. This state of knowledge does not preclude that other patterns are possible. The learner will eventually be exposed to other patterns that (23*E*) and (23*G*) instantiate. We may suppose that what he or she will perform in his or her mind *is* the pattern matching illustrated in (23), which serves as an optimization of the differences among (23*E*, *F*, *G*), for example. The little linguist eventually knows that (24*H*) is also admissible. At this stage, he or she arrives at the correspondence illustrated in (24), which is a further optimization of pattern correspondences. Of course, the exact order of experiences is irrelevant, and furthermore the learning is probably lexically based in conformity with "usage-based" model (Langacker 1988). This picture suggests that generalization of surface forms/patterns are sufficient for language acquisition. I believe this speculative scenario is fully compatible with Elman (1992) who shows that simple recurrent PDP networks arrive at distinction among lexical categories and grammatical relations by learning only positive evidence.

4.3 Positional and Lexical Mismatches

There is another kind of partial mismatches that should be taken into account. Consider the correspondence between *F* and *G* in (25).

(25)	*F*:	Many children	like	Mickey Mouse
			|	
	G:	Many children	love	Mickey Mouse

No phantom is necessary in (25), since the relevant kind of mismatch is not positional. We say that this kind of partial mismatch is **lexical**.

4.4 Pattern Marching Assumes No Identity in Meaning

We need a remark about such correspondences as in (25). Pattern matching is not exclusively based on the notion of "meaning the same," if it is ever definable. Instead, we are willing to admit that there are many other kinds of relations that make patterns related to each other. In this respect, we do by no means claim that *F* and *G* in (25) have no difference in what they say. The contrary is true. We rather say that there is always a difference in meaning, no matter how subtle, whenever there is a difference in form, positional or lexical. What we want to make clear is thus simply that what formal difference leads to what semantic/pragmatic difference. Simply put, difference in meaning is default, since we try to describe different kinds of relations among forms. For this simple reason, none of our assumptions is weakened by the fact that, with similarity on one hand, there is semantic/pragmatic difference between *F* and *G*, on the other. Let me specify this point more clearly.

4.5 Default Difference in Meaning

We have a dilemma: it is trivially true that different forms have different meanings, but it is also true that their meanings are not totally different. *F* and *G* in (25), for example, are different forms which have different meanings. But they also have a portion of meaning in common. To escape from this dilemma, let us interpret the famous slogan, "all different forms have different meanings," as the principle of **default difference**. Radically put, no different forms can never have the same meaning. This must be a principle, or an axiom, simply because it is impossible to account for why different forms ever have different meanings, without going into circularity.

But even if meanings of two forms are not so different, it does not follow that they have exactly the same meaning. So motivated, our dilemma can be stated as the problem of to what extent different meanings can be similar, namely of the degree of similarity in meaning, with "meaning almost the same" as a special case of such similarity.[18] Recall that nothing can be identical to anything but itself. To be more explicit, let us furthermore interpret similarity as the property of "loss of differences." Generally speaking, two different things are categorized as identical only if all of their latent differences are dismissed.

4.6 Matching Scales

But there is yet another, more complex kind of partial mismatch, at least apparently. Compare with cases like (25) those like (26) where lexical and positional mismatches take place at the same time.

(26)	*F*:	Many children	•	like	•	Mickey Mouse
			|	|	|	
	G:	Many children	are	fond	of	Mickey Mouse

It is attractive to treat the sequence *be fond of* as a sort of "lexical decomposition" of *like* (see McCawley (1988, p. 656) for lexical decomposition). If this possibility is pursued, then we may say that in such cases as (27), there occurs "merging," without recourse to phantoms which encode positional mismatches.

(27)	*F*:	Many children	like	Mickey Mouse
			/ | \\	
	G:	Many children	are fond of	Mickey Mouse

Let me emphasize, however, that such decision is not mutually exclusive, and it is quite misleading to claim factually that there is no positional mismatch in (27). Leaving aside the issue of whether they are synonymous, which I claim is not crucial, to dismiss positional mismatches is equal to claim factually that sequences like *be fond of* are single words as *like* actually is, which sounds quite odd.[19]

But this dilemma could be avoided easily if we introduce the notion of **matching scales**, which I discussed briefly in §3.6. Units like *be fond of* and those like *like* have different matching scales. If we compare *F* and *G* on the same scale *W* such that U_W= {*many children, likes, Mickey Mouse*}, U_W= {*many children, are, fond, of, Mickey Mouse*} are definable. (26) is responsible for pattern matching on this scale. But this does not preclude the existence of another scale *V* on which we define U_V= {*many children, are fond of, Mickey Mouse*} and compare *F* on *W* scale and *G* on *V* scale, which (27) is responsible for. Scale shift is motivated (or constrained) by the similarity of the meaning of *like* and *be fond of*.

4.7 Forms and Patterns

It is urgent to add that ff correspondences need not be stated exclusively on the lexical basis. A crucial point is that in cases like (28) we are not forced to say that there are partial mismatches between *F* and *G* at every phrasal unit that follows *that*.

(28)	*F*:	Fred believes that	*Liz Taylor*	*shot*	*her husband dead*
			|	|	|
	G:	Fred believes that	*the earth*	*is*	*made of mushroom*

Though possible (and perhaps necessary in some cases), this does not exhaust the possibility of pattern matching: words and phrases are merely units of possible scales for pattern matching. Pattern matching is also possible on other scales, larger or smaller, and probably there will be no limit on the number of possible scales.[20] On larger scales, we may describe such pattern matching as (29) and (30) specify.

(29) *F*: Fred believes *X* *Y*
                                        |
G: Fred believes *X* *Y′*

(30) *F*: Fred believes *Z*
                                |
G: Fred believes *Z′*

where (i) *X*= *that*; (ii) *Y*= *Liz Taylor shot her husband dead* and *Y′*= *the earth is made of mushroom*; and (iii) *Z*= *X Y*, and *Z′*= *X Y′*. Note that the scales for *X Y* and that for *Z* are different.

A notional distinction between **forms** and **patterns** is helpful, though most linguists are not aware of this, for they use the term "form" in the sense of "pattern." By "forms," I mean results of formations that are exclusively made of lexical items. By "patterns," I mean strings that contain one or more lexically uninstantiated variables with arbitrary labels like *X*, *Y*, *Z*. Patterns have a variety of degrees of "abstractness," and the degree varies with the number of lexically uninstantiated variables. Specification of variables is an empirical problem. So, strictly speaking, *Fred believes X Y*, *Fred believes Z* are patterns, not forms in the exact sense of the term. But I will use the term form as loosely as the tradition of linguistics allows.

4.8 Scales of Patterns and Their Units

Note however that the problem of what "labels" *X* and *Y* should have is a completely different sort of problem. Plainly, it doesn't really matter whether or not *X* is "complementizer" (C), whether *Y* is "sentence" (S), or "inflectional phrase" (IP), or "tense phrase" (TP). Similarly, it is of little interest whether or not *Z* is "complementizer phrase" (CP). Nomenclature will be helpful, and in some cases necessary, but it must be distinguished from a minimal requirement. Minimally required in cases like (29) is the existential claim that there exist certain scales of recognition on which units *X* and *Y*, on the one hand, and *Z*, on the other, are definable: anything more than this must be independently motivated.

This does not mean, of course, that we may not label variables (including phantoms). What I reject here is rather the conception of lexical items as instantiation of "syntactic categories." I claim the reverse is true. Lexical items exist first, and categories, if any, are distributional generalizations of them. This is, I contend, the view of language that is compatible with Elman (1992) who demonstrated the PDP learnability of grammar.

The view of linguistic expressions advocated here also renders drastically the notion of "syntactic structures." If there are ever such structures, they must be interactions of patterns on a number of different scales. I claim that tree diagrams can capture only a limited portion of the complex interactions of patterns across scales.

4.9 Synchronized Mismatches

The notion of **synchronization** of mismatches plays a crucial role to narrow our scope of investigation in a desirable way. Without it, we would be forced to account for such kind of partial mismatches as in (31) in terms of ff correspondence.

(31)	*F*:	Andy	hates	Bill
		|		|
	G:	Bill	hates	Andy

There are, at least apparently, two mismatches in (31), but it is dull to attempt to describe such cases in terms of ff correspondence. It is mainly because it is unlikely that this sort of correspondences display real contrasts.

But how to justify this? The most plausible way to escape from this dullness is to rely on the fact that the partial mismatches are not synchronized. As we have seen in (24), positional mismatches are usually mutually exclusive. But this is not true of (31). In fact, given pattern *X hates Y*, which is instantiated by *F* and *G*, lexical choices for variables *X* and *Y* are basically independent of each other. *F* can contrasts with *G′= Andy hates Carol*, for example, without *Andy* being forced to be other form. In short, two partial mismatches between *F* and *G* in (31) are not constraining each other, and are rather caused by free lexical variations of a single pattern *X hates Y*. It is possible to describe these unsynchronized contrasts, but it would not provide so many insights into ff correspondence. Let me clarify this point.

4.10 A Limitation on Pattern Matching Analysis

We have of course cases that indicate lexical selection is not completely free. To see this, it would be sufficient to look at cases like (32) and (33), on the one hand, and cases like (34), on the other.

(32)	*F*:	Andy	hates	Bill
		|		
	?**G*:	*Gravity*	hates	Bill
(33)	*F*:	Andy	hates	Bill
				|
	?**G*:	Andy	hates	*three hours*
(34)	**E*:	Andy	hates	*Andy*
				|
	F:	Andy	hates	*himself*
				|
	**G*:	Andy	hates	*him*

It is important to note that pattern matching analysis has an obvious limit: since it is a theory of ff correspondence, it cannot account for unacceptability of forms by itself. The reason is that acceptability is a property definable on formations rather than correspondences: acceptability is determined by simultaneous satisfaction of well-formedness conditions on formations. More generally, any comparative method cannot account for well-formedness of what are being compared. Thus, what we can do is to make use of unacceptable forms to describe how different patterns interact with each other, rather than accounting for their acceptabilities. Bearing this remark in mind, let us see some details of the correspondences.

What is responsible for the unacceptability of *G*-forms in (32) and (33) is so-called "selectional restrictions." Simply put, *gravity* is inappropriate for *X* in *X hates Bill*, on the one hand, and *three hours* for *Y* in *Andy hates Y*, on the other. By contrast, *E* and *G* in (34) exhibit so-called "crossover phenomena" (Postal 1971). *E* is usually unacceptable unless *Andy* as subject refers to, say, Andy Watkins and *Andy* as the object refers to another Andy, Andy Baltimore. The same is true of *G*-form: *G* is unacceptable unless *him* refers to some other person than *Andy*. The deviance of *E* and *G* in (34) are usually accounted for by Chomsky's (1981) Binding Conditions C and B, respectively, though it is highly dubious whether it is successful.[21]

As noted above, our pattern matching method does not provide any interesting viewpoint in these cases: the reason is obvious simply that here is a problem of "formation," not of correspondence. But let me make residual remarks.

It is unimportant, at least for our purposes, to distinguish selectional restrictions from "co-occurrence restrictions" shown by the unacceptability of *E* and *G* in (34), though the distinction is usually made according to whether deviance is caused by semantic or syntactic factors. The strongest reason is that it is not clear what we could benefit from the distinction of semantic and syntactic factors, both are possible only in idealized dimensions.

Recall that we have observed that in (24), *carefully* may not instantiate all occurrences of phantoms at once in *F*= *Marvin* • *sliced* • *all the bagels* • . But why such constraint on instantiation? Since phantoms are not empty categories, the reason cannot be purely structural. Rather, there would be a general principle governing co-occurrence restrictions, or pattern composition in general. I guess the principle is basically that no two words may appear in a minimal sentence if they have the same reference, which is a stronger condition than Binding Condition C (+B), since it includes all other categories than NP. So, what we have to account for is why such reflexive forms as *himself* in (34*F*) can escape from this restriction. But, to elaborate this idea, we will need a general theory of pattern composition, which is beyond the scope of this paper. But I am now working out this. I suggest here that such condition is necessary to avoid a short circuit of pattern glues.

4.11 Manifold Correspondences and Topology of Contrasts

Another important feature that pattern matching provides us is that it permits us to characterize manifold correspondences, namely complex interactions of correspondences among as many forms as we want to see. Let us take for example the multiple correspondences in (35).

(35)	*E':	Fred	believes	*of*	*it*	•	*that*	*[Bill*	*is*	*a liar]*
	E:	Fred	believes	•	•	•	*that*	*[Bill*	*is*	*a liar]*
	F:	Fred	believes	•	*Bill*	•	•	•	*to be*	*a liar*
	*F':	Fred	believes	*of*	*Bill*	•	•	•	*to be*	*a liar*
	?*G':	Fred	thinks	*of*	*Bill*	•	•	•	*to be*	*a liar*
	G:	Fred	thinks	*of*	*Bill*	•	•	•	*as*	*a liar*
	?**H*:	Fred	thinks	*of*	*it*	*as true*	*that*	*[Bill*	*is*	*a liar]*
	I:	Fred	takes	•	*it for granted*		*that*	*[Bill*	*is*	*a liar]*

Incidentally, the whole constitutes a sort of ring, since E' and *I* are linked.

Delaying relevant details until §4.13, let me specify our main point in advance. We can see in this manifold correspondence the way patterns partially overlap on each other: *E* and *F* contrast with the partial mismatches (i) • – *Bill*, (ii) *that* – •, (iii) *Bill* – •, (iv) *is* – *to be*. Likewise, *F* and **F′* contrast with respect to • – *of*. *F* and *G′* do so with respect to *believe* – *think*. *G′* and *G* with respect to *to be* – *as*, and so on.

It is easy to see that patterns in (35) interact with each other in the manner that Lakoff's (1987) calls "radial categories," on the one hand, and the manner that Langacker (1987) calls "network," on the other.

Our main point is this: such manifold correspondences as in (35) illustrate how contrasts shift from one pattern to another. Or, more adequately, they provide a chance to catch a glimpse of the topology of contrasts, which makes us to state in what respect patterns are "similar" and "dissimilar." Indeed, pattern matching analysis permits us to state manifold ff correspondences among as many forms as we wish to explore.[22]

By virtue of the capacity to visualize topological structure of patterns, our method is given an excellent quality that, as far as I know, no other theoretical frameworks provide, at least systematically.[23] We are provided with a chance to understand the real diversity of patterns and the complexity of their interactions.

4.12 A Note on the Complexity of Manifold Correspondences

The only undesirable consequence of manifold expressibility of ff correspondences is the difficulty in conceiving their complexity on intuitive basis. For terminological convenience, we will say that correspondences of a given form pair is onefold, since the number of correspondence is one. Similarly, if we want to analyze the correspondences among three forms, then we have a threefold correspondence. More generally, we have an $(n-1)n/2$-fold correspondence for a given group of n forms, without taking degenerate correspondences into account. Since the index $(n-1)n/2$ increases by the order of square, its potential complexity is manifest even in the ff correspondences among five forms, *E*, *F*, *G*, *H* and *I*, as illustrated in Figure 2 which embodies a tenfold correspondence.

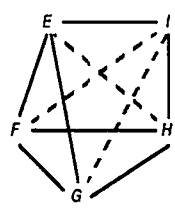

Figure 2

In this figure, we distinguish direct correspondences (e.g., *F–H*) by indicating them by solid lines, from indirect ones (e.g., *G–I*) by indicating them by dashed lines.

By this real complexity, however, I believe we will obtain more insights into syntactic phenomena than lose sight in them.

4.13 What Is Really "Exceptional" in Exceptional Case Marking?

Let us return to the ff correspondences in (35), which reveals an interesting aspect about complementation.

Pattern *F* is notorious for the "exceptional case-marking" of *believe* (Chomsky 1981). The crucial question is why *Bill* in *F* behaves as object of *believe*. There was a controversy as to whether "subject-to-object raising" is involved or not. Generative semanticists (Postal 1974) argued for it, and most interpretive semanticists argued against it, who insisted on $\bar{S}$ deletion, because there is *for*, Comp, that is to be deleted.[24] Of course, such problem is unimportant for our concerns, simply because what we are doing is compare different formations, rather than deriving *F* form *E*. So, even if (35) may seem to support subject-to-object raising, it is a different matter. Most crucially, we do not regard *E* as underlying form of *F*: they are simply different forms constituting a pattern network.

Based on the topology of contrasts from *E* to *H*, we could observe that:

(i) There is a qualitative leap between *E* and *F* due to their difference in pattern scale;
(ii) Verbs of thinking, e.g., *believe, think, conceive*, demand as auxiliary case-marker *of* independently of *for* and *that* for complementation;
(iii) The unacceptability of *F´* is exceptional (cf. *G*);
(iv) On this basis, *for* in *for-to* complement is probably not Comp any more than *of* is;
(v) It follows either: (a) there is a "missing" preposition that the phantom in *believe* __ *Bill* in *F* indicates, which case-marks *Bill*; or (b) the preposition is "incorporated" into the sense of *believe* (cf. *conceive (of) X as Y*).

It is suggested that what is really exceptional, if any, is that *believe* lacks auxiliary preposition to mark NP. The last point of preposition incorporation will be discussed also in §4.17.

4.14 Measuring "Distances" among Patterns

From the manifold correspondence in (35), we can of course only catch a glimpse of how patterns interact with each other. It is not only useful to investigate other interesting properties, we can even measure the "distances" among patterns. In the case of (35), *F= X believes • Y• • to be Z* is distant from *G= X thinks of Y• • as Z*, one the one hand, which is the nearest form that is not unacceptable, by three contrasts, namely: (i) *believe–think*; (ii) • – *of* (iii) *(to be) – as*, and from *E= X believes • • • that Y is Z*, by four contrasts, namely (i) *Bill–* •, (ii) • – *that*, (iii) • – *Bill*, and (iv) *to be – is*, on the other. If we assume for simplicity that each contrast results in the same amount of distance, then even if we do not take into account the disappearance of S-boundary ([...]), it follows that pattern *G* is nearer to pattern *F* than pattern *E* is to *F*. Such measurement of cross-pattern distances, though highly tentative, could be used, if elaborated further, as quantitative, rather than qualitatively and intuitively based, evidence to characterize the "exceptional" behavior of *believe*. I think that this measurement technique seems worthy of elaborations in future research.

4.15 Active, Causative, Passive and Stative

Let us take for another example the active-passive correspondence. Compare the "direct" correspondence between active and passive voices in (36) and the manifold correspondence among active, passive, stative, and other related voices in (37), where the correspondence between *F* and *J* are "indirect."

(36)	*F*:	Bill	opened	the door	•	•	{*by himself* / **by itself*}	Active
		|	|		|	|	|	
	J:	•	•	the door	was	opened	{*by Bill* / **by itself*}	Passive

(37)

F:	Bill	opened	the door	•	•	{*by himself* / **by itself*}	Agentive-Active
G:	Bill	caused	the door	•	open	{*by himself* / ?* *by itself*}	Causative-Active
G':	Bill	let	the door	•	open	{?* *by himself* / *by itself*}	Permissive-Active
H:	•	•	the door	•	opened	{**by Bill* / *by itself*}	Active-Passive
I:	•	•	the door	was	open	{**by Bill* / **by itself*}	Stative
J:	•	•	the door	was	opened	{*by Bill* / **by itself*}	Passive

The difference of (36) and (37) suggests that the active-passive association is a rather simplified picture. Based on the distance measuring mentioned above, active-passive correspondence, if any, contains a number of potential contrasts than expectable from their direct comparison. A straightforward interpretation of (37) would be that active *F* is rather distant from passive *J*, since there are at least four intermediate patterns *G*, *G'*, *H*, and *I* between them. The co-occurrence restrictions on *by*-phrases is very suggestive.

Note however that *G*, *G'*, *H*, and *I* are not intermediate steps for derivation from *F* to *J* (or *J* from *F*, if claimed). They are rather distinguishable patterns that constitute a **network of patterns**, all of which have surface form status of their own.

4.16 No Need for "Logical Forms"

Although pattern matching analysis does not rely on the notion of underlying structure, yet it can treat with the kind of ambiguity exhibited by sentences like (38).

(38) Many students read many books.

It is well known that (38) is two-way ambiguous[25] so that the following sentences are paraphrases of (38), to which McCawley (1981, 1988) calls "pseudo-relatives."

(39) a. There are many students who read many books.
b. There are many books which many students read.

It is a commonplace to disambiguate the meanings of (39a, b) by translating them into so-called "logical form," $\exists x(Fx)$. But the very fact that the relevant ambiguity can be paraphrased by the sentences in (39) is sufficient enough. Indeed, without recourse to logical form, we can employ (39a, b), all of which begin with *there are many X*, to contrast with each other, as in (40).

(40)	*E*:	there are	many books	which	many students	read		•
		|	|	|				|
	F:	•	•	•	many students	read		many books
		|	|	|	|			
	G:	there are	many students	who	•	read		many books

(41)	(	∃	x	)	(Fx	)

It is striking to note the following points. As illustrated in (41), (i) the expression *there are* corresponds to "existential quantifier," ∃; (ii) the expression *many* N to "domain expression" to be quantified by ∃ to form $\exists x$; (iii) if the phantoms inside relative clauses can be identified with "bound variable," then *wh*-expressions function as a "pipe," especially when we employ notation "$\exists x \mid Fx$" which is familiar in UNIX shells.

We will not, however, attempt to account for why sentences like *F* can ever be ambiguous. What we can assert is that pattern *F* activates the pattern network. If its activation settles on *E*, we have the target reading for which *E* is responsible. If, alternatively, the activation settles on *G*, we have the target reading for which *G* is responsible.

In short, if such multiple correspondence as in (40) is correct, then we can achieve the same exactness as logical forms provide us, and, what is more, we can do it almost cost-free.

4.17 Alternation

At this final stage, let us return to our first example, (8), repeated here for convenience.

(8)	*F*:	They	loaded	bricks	*in*	the truck	•	•
				|	|		|	|
	G:	They	loaded	•	•	the truck	*with*	bricks

What this correspondence claims is that, given general patterns $F = X\ V\ Y\ P\ Z$ and $G = X\ V\ Z\ P'\ Y$ (e.g., *P=in(to)*, *P'=with*), *Z* (*=the truck*) serves as the axis of **figure/ground reversal** in postverbal configuration. It is noticeable that, in *F*, *Y* (*=bricks*) behaves as "figure," or "trajector" in the sense of Langacker (1987), but, in *G*, it behaves as "ground," or "landmark," on the one hand, and *Z* (*=the truck*) behaves as ground in *F*, while it behaves as figure in *G*, on the other. This point is evident from the opposition in meaning of *P=in* and *P'=with*. More generally, postverbal alternations like in (8) are not only ff correspondences but also mm correspondences mediated by conceptual transformation (figure/ground reversal) on "container metaphor" (Lakoff and Johnson 1980). In this respect, what we see here is not simply an alternation of patterns but also an alternation of **conceptual images**. This will true of other cases of alternation between *Y P Z* and *Z P' Y*, where *P=in, on*; *P'=of, with*.

We may integrate syntactic and semantic dimensions into such **pattern images** as illustrated in Figure 3, where *V* and *P* are conceived of as "trajectories," or vectors that dominate "domain" D.

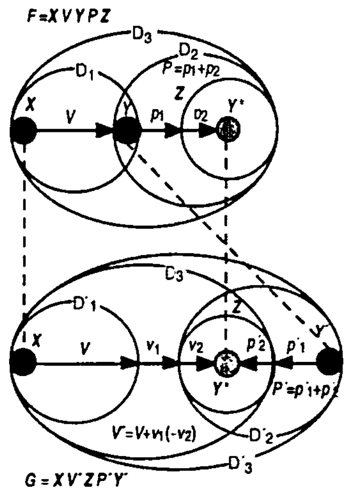

Figure 3

What this figure embodies is that the *F-G* correspondence is an **extension** from *F* to *G* form in that the sense of *P* ($=p_1+p_2$) in *F* is "incorporated" into the meaning of *V'* in *G*, by which *V'* comes to dominate a new domain, D_3 in *G*, for the scope of predication. To describe this special situation, we may say, rejecting the notion of derivation, that *G* is a pattern "based on" *F*.[26]

Though the details of the figure are debatable, I think this is one of the best places where pattern matching analysis meets cognitive semantics.

5. Concluding Remarks

It is unreasonable to want pattern matching analysis, intended as an integral part of cognitive syntax, not to have any resemblance to any other theories of syntax. I had some linguists around me who in fact objected to me, claiming that my version of cognitive syntax is not so different from generative one. Surely, I must concede that this effort is in progress, and there is still much to do. But I believe I have sketched so far a different and promising way of (re)doing syntax. Although there are many similarities with other theories of syntax, there are far more dissimilarities.

Probably, they objected to me, given that my presentation was not too obscure to convey real meanings, only because they dislike whatever is called syntax. I have the impression that they talked as if it were wrong to do syntax in any serious manner. Well, this is a matter of taste, and there is no account of taste.

"No (more) formal syntax" seems to be the slogan bonding strongly current researchers in cognitive grammar altogether, if I could say so. They seem to be gathering to fight a common imaginary enemy, namely "syntax chauvinists."[27] Their reaction is partly reasonable. In most theories of grammar or language or both, syntax was syntax, and semantics was semantics, if anything.

I am willing to admit that cognitive linguistics is a set of elegant and effective theories of semantics and pragmatics (Fauconnier 1985; Lakoff 1987; Langacker 1987, 1991, among many others) that is coming to maturity. But now, syntax and semantics/pragmatics are in mutual segregation. The former does only a little of superficial semantics/pragmatics, and the latter only a little of superficial syntax. But everything needs a good balance, I believe. I am afraid that too strong reaction to generative linguistics might not oscillate everything to the other extreme side, the world of "semantics/pragmatics chauvinists," where any chance of reworking in syntax cannot survive.

I believe Lakoff's (1993) insights into phonology, on the one hand, and Karttunen's (1993) appraisal of finitary means, opened a new way of redoing syntax, doing syntax differently. As stated in introduction, this paper is a status report of my project of reworking in syntax.

I understand that I deviate from Lakoff (1993) in one noticeable point. I replaced what he calls "constructions" by form-form correspondences, as distinguished from form–meaning correspondences, which correspond to what Fillmore, Kay and O'Corner (1988), Goldberg (1995), Lakoff (1987), and many other cognitive linguists call (grammatical) constructions. But this deviation was necessary to meet the special requirement of syntax: to reject the notion of "levels."

Some of us may remark that meaning is precluded (again) in pattern matching analysis. I object to this. Meanings are always behind forms supporting any form-form correspondences. I talked little about meaning only because meaning does not play the leading role of syntax. As form should go background when we talk about meaning, meaning should go background when we talk about form. This is just what "trajector/landmark symmetry" demands.

I am willing to hold, as Langacker (1987) forcefully argues, that grammar has strong symbolic basis. No doubt, any predicational relation is symbolic in its very nature. This is a crucial property that characterizes most of grammar. But I would be more than mildly surprised if it is claimed that all of grammar is symbolic. The reason is obvious: neither morphotactics (= syntax) nor semotactics is symbolic; otherwise, no "information," in the exact sense of Shannon and Weaver (1949), can never exist, nor can language exist, which is a contradiction to the fact.

Straightforwardly, semo(syn)tactics and morpho(syn)tactics are probably two different kinds of tactic relation mediated by symbolization. In this respect, it is quite useful, and I argue necessary, to distinguish form–form correspondences from form–meaning correspondences. Form–meaning correspondences are, I think, too much emphasized in cognitive paradigm to make us blind to the dimension of form–form correspondences. I regret this. If any theory of language wants to be more comprehensive, more explanatory, and more faithful to brute facts, it is sure that both form-meaning and form-form correspondences must be properly recognized and characterized. None of them can be omitted in a well-articulated theory of language.

Notes

1. Karttunen's discussion is more thorough, I think. He not only advocates two-level rule systems, but also assures that they never have more (weak) generative capacity than one-level rule systems. More specifically, the disappearance of rule ordering, cyclicity from rule systems counterbalances the increase in complexity of statements (Incidentally, I could not figure out Lakoff's point until I understood, *after* reading Karttunen, that constructions are two-level rules.

2. This is a rather simplified picture. There is indeed a crucial condition to ensure this. This is what Karttunen (1993) calls "finite-state constraint," which says that phonological alternations must be "regular relations."

3. Although I cannot be confident, Lakoff-Goldsmith style of three-level architecture of phonological system is dubious, since, if Karttunen (1993) is correct, it follows that, if M–W and W–P correspondences are regular relations, then M–P correspondence is also regular. That is, W-level is dispensable, at least formally, since M–P correspondence are statable, according to the results of Johnson (1972), Kaplan and Kay (1982), and Koskenniemi (1983).

4. For details of autosegmental phonology, see Goldsmith (1979).

5. This is not exactly what Lakoff presents. I replace his cross-hatches over corresponding lines by " • ".

6. According to Révész (1991, p. 6), "a generative grammar $G = (V_N, V_T, S, F)$ [V_N is a set of nonterminal symbols,

V_T is one of terminal ones, S is an initial symbol, and F is a set of rewriting rules] is said to be of *type i* if it satisfies the corresponding restrictions in this list:

i=0: No restrictions.

i=1: Every rewriting rule in F has form $Q_1 A Q_2 \Rightarrow Q_1 P Q_2$, with Q_1, Q_2, and P in $(V_N \cup V_T)^*$, $A \in V_N$, and $P \neq \lambda$, except possibly for the rule $S \Rightarrow \lambda$, which may occur in F, in which case S does not occur on the right-hand side of the rule.

i=2: Every rule in F has form $A \Rightarrow P$, where $A \in V_N$ and $P \in (V_N \cup V_T)^*$.

i=3: Every rule in F has form either $A \Rightarrow PB$ or $A \Rightarrow P$, where $A, B \in V_N$ and $P \in V_T^*$.

A language is said to be of *type i* language if it is generated by a *type i* grammar."

It is also known that every type 3 language can be described in form of a regular expression. According to Révész (1991, p. 39), "a regular expression over a finite alphabet V is defined inductively as follows:

1) λ (i.e., the empty string) is a regular expression.

2) a is a regular expression for every a in V.

3) If R is a regular expression over V, then so is $(R)^*$.

4) If Q and R are regular expressions over V then so are $(Q);(R)$ and $(Q)|(R)$,

where symbols "*", ";", and "|" stands for the operations of iteration, concatenation, and set union (or disjunction).

7. But, as Peters an Ritchie (1973) demonstrated, the class of transformational grammars defined in Chomsky (1965) was type 0.

8. Not surprisingly, we can devise a finite-state transducer to disprove Chomsky's (1957, pp. 21–23) argument that sentences like *If S, then S,* and *If if S, then S, then S, If if S, then if S, then S,* cannot be generated by finite-state grammars. Consider the finite-state transducer illustrated in Figure A.

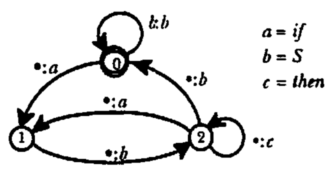

Figure A

This is my modification of Karttunen's (1993, p. 176, Fig. 6.2) transducer implementing • => *ab*/__*b* which generates $a[ab]^*b$ from *ab*. Another component, *either S or S,* is also describable by the same type of transducer. Since finite-state transducers are composable, it is possible to two transducers of this kind to generate *If either S or S, then if either S or S, then S,* though I do not illustrate here.

9. I say this with preparation to have a lot of objections to this assertion. But I cannot help say this, because it seems to me that cognitive grammarians tend to take syntax too lightly, as generative grammarians tend to take semantics and pragmatics too lightly.

10. In general, if f consists of n units, then the number of possible pieces of fm correspondence is $n(n+1)/2$.

11. See also Lakatos and Musgrave, *eds.* (1970) for "research program" and other related notions.

12. See Feyerabend (1970, 1974) for incommensurability.

13. See Mandelbrot 1977, 1983 for relevant notions of "scales" and "scaling."

14. We may furthermore think that an optimal state is a "harmonious" states in the sense of Smolensky (1986), or even "equilibrium" in the sense of Piaget (1971a, b). For this unexpected convergence of ideas, the reader is advised to refer to Bates and Elman (1993), which is also quite useful to get a quick but adequate gasp of connectionism.

15. This analysis is, to tell the truth, wrong, since it does not captures the correspondence between F=NP V NP P NP (e.g., *Sam sent the letter to Jim*) and G=NP *be* V-*en* P NP *by* NP (e.g., ?*The letter was sent to Jim by Sam*) and G'=NP *be* V-*en* NP (*to) (e.g., ??*Sam was sent the letter by Jim*). We would have a better analysis if we assume that passive and active forms are so different patterns that there are no partial exact matches between them, and that what

associates them is not syntax but pragmatically based inferences.

16. More specifically, *-en* is probably an "anticausative" morpheme. This point is also relevant to the alternatin between *YP Z* and *Z P' Y* that we will discuss later. In thise case, *P* (e.g., *in*) and *P'* (e.g., *of*) are conceptual vectors that point to opposite directions.

17. For intriguing semantics of adverbs, the reader is advised to consult McCawley (1988, ch. 19).

18. Note incidentally that transformational analyses, assuming deep structures, are forced to approach the same problem from the opposite direction, since they stipulate "default identity" in meaning. So, it must specify differences if there are any.

19. Incidentally, Lakoff (1993, p. 122) does not hesitate to incorporate the notion of "merging" into his cognitive phonology framework. The exact status of merging is questionable, and we will need more careful treatment, especially as to whether or not it really involves positional mismatches in the case of syntax.

20. It is very important to note that there is invariance of patterns across scales. For evidence, when pattern *X V Y* (*V* is verb, *X*, *Y* are NP) is identified on S-scale (hence SOV-language), patterns on all other scales tend to conform to it. Super-sentence patterns such as *X if Y*, *X when Y*, *X because Y* (*X*, *Y* are S's) conform to S-scale analogue, and P-scale pattern *X of Y* (*X*, *Y* are NP's), too. Also on W-scale, we find pattern *X S Y* (*S* is stem, *X*, *Y* are affixes). Even on far more smaller scale of syllables, we find *X V Y* (*V* is vowel and *X*, *Y* are consonant clusters). This interesting parallelism is usually expressed in terms of so-called $\overline{X}$ theory (Chomsky 1972, Jackendoff 1977), but it is gratuitous to identify such invariance emerges as a manifestation of "universal grammar," since it does not provide its explanation in a serious sense. A milder, and we believe more plausible, assumption is of course that such invariance is a scaling effect (Mandelbrot 1977, 1983) that we can expect from the optimization process in neural networks.

21. Postal (Huck and Goldsmith. 1995, pp. 141–142) criticizes this line of approach as follows: "Chomsky [1981: 193] has argued that Principle C of this so-called Binding Theory explains the strong crossover phenomenon first discussed in Postal 1971. Roughly, this principle claims that a "referring expression" cannot be bound. The claim is thus that the crossover facts can be deducible from this principle. But this fails resoundingly. First, since the Binding Theory only characterizes noun phrases, Chomsky's Principle C treatment entails that no strong crossover effects can be induced by extraction of prepositional phrases, thereby falsely distinguishing the equally impossible **To whom did he say I gave the book* and **Who did he say I gave the book to*, where *he* and *who* or "whom" are supposed to be coreferential. [a note omitted] Second, Chomsky's supposed explanation also ignores the fact that there is an asymmetry in the phenomenon. Roughly, it only exists when an antecedent is extracted, not a pronominal form; thus, a Principle C approach fails to distinguish **Jerome, I convinced him I would hire* (where *Jerome* and *him* are supposed to be coreferential) from *Myself, I can't begin to understand*, wrongly blocking both. [another note omitted]"

22. In this respect, our method could be compared to "comparative anatomy" in biology. Comparative anatomy is a method of biological study by which structures of animal bodies are compared on the homological basis. It enables biologists to explore the diversity of animals not only at the level of species but also at higher levels, e.g., genus, family, class. Thus, if this analogy is correct, our method can be considered as not only "analogically" based in mm correspondence, but also is "homologically based," in terms of ff correspondence.

23. A number of crucial properties of pattern networks may be accounted for in terms of well-articulated semantics and pragmatics, since, as noted earlier, ff correspondences are supported by "hidden" mm correspondences. But I feel it is an overestimation to naively claim that all of the properties of pattern networks could be accounted for in terms of fm correspondences. One obvious reason is that there can be, and I believe is, a distinct kind of similarity among patterns in themselves.

24. Bach (1977) is a good survey of the controversy. To avoid confusion, it must be mentioned here that this position is abandoned in "minimalist program" (Chomsky 1995) which is incompatible in many nontrivial respects with deceased "revised extended standard theory" (Chomsky 1977) and its descendent "government-binding" framework (Chomsky 1981).

25. I omit here two other readings, one of which is a very marginal reading that corresponds to *There are many ways in which a student reads a book*. This will require the notion of generalized quantifier, since *many* binds both of the determiners of *x student* and *x book*. Another is a simple event reading, *There occurs an even in which many students read many books*. This reading is dominant in past tense sentences. For evidence, *When I entered the library, many students read many books* receives primarily this event reading, and marginally receives other logical readings.

26. See Lakoff (1987, p. 555) for the distinction of "base on" from "derived from."

27. This witty locution is brought to me by Yamanashi (p.c.).

References

Bach, Emmon W. 1977. Review article of *On Raising* by Postal. *Language* 53 (3): 621–654.

Bates, Elizabeth A., and Jeffrey L. Elman. 1993. Connectionism and the study of change. In *Brain Development and Cognition: A Reader*, edited by M. H. Johnson. Oxford: Basil Blackwell.

Chomsky, Noam. 1956. Three models for the description of language. *IRE Transactions on Information Theory*, 2 (3): 113–124.

_____. 1957. *Syntactic Structures*. The Hague: Mouton and Co.

_____. 1965. *Aspects of the Theory of Syntax*. Cambridge, MA: MIT Press.

_____. 1970. Remarks on Nominalization. In *Readings in English Transformational Grammar*, edited by R. A. Jacobs and P. S. Rosenbaum. Waltham, MA/Tronto/London: Ginn and Company.

_____. 1973. Conditions on transformations. In *A Festschrift for Morris Halle*, edited by S. R. Anderson and P. Kiparsky. New York: Holt, Rinehart and Winston.

_____. 1975. *The Logical Structure of Linguistic Theory*. New York: Plenum Press.

_____. 1977. *Essays on Form and Interpretation*. New York: North Holland.

_____. 1981. *Lectures on Government and Binding: The Pisa Lectures*. Berlin/New York: Mouton de Gruyter.

_____. 1995. *The Minimalist Program*. Cambridge, Mass.: MIT Press.

_____, and Morris Halle. 1968. *The Sound Pattern of English*. First MIT paperback edition in 1991. Cambridge, MA: MIT Press.

Davis, Steven, *ed*. 1992. *Connectionism: Theory and Practice*. New York/Oxford: Oxford University Press.

Elman, Jeffrey L. 1992. Grammatical structure and distributed representations. In *Connectionism: Theory and Practice*, edited by S. Davis. Oxford, UK: Oxford University Press.

Fauconnier, Gilles R. 1994. *Mental Spaces: Aspects of Meaning Construction in Natural Language*. Cambridge: Cambridge University Press. [First 1985]

Feyerabend, Paul K. 1970. Consolations for the specialist. In *Criticism and the Growth of Knowledge*, edited by I. Lakatos and A. E. Musgrave. Cambridge: Cambridge University Press.

_____. 1974. *Against Method: Outline of An Anarchistic Theory of Knowledge*. London: New Left Books.

Fillmore, Charles J., Paul Kay, and Catherine O'Connor. 1988. Regularity and idiomaticity in grammatical constructions: The case of *let alone*. *Language* 64: 501–538.

Goldsmith, John. 1979. *Autosegmental Phonology*. New York: Garland Press.

_____. 1992. Local modelling in phonology. In *Connectionism: Theory and Practice*, edited by S. Davis. New York/Oxford: Oxford University Press.

_____, *ed*. 1993. *The Last Phonological Rule: Reflections on Constraints and Derivations*. Chicago: University of Chicago Press.

Halle, Morris, and George N. Clements. 1983. *Problem Book in Phonology*. Cambridge, MA: MIT Press.

Hanson, N. R. 1958. *Patterns of Discovery*. Cambridge: Cambridge University Press.

Huck, Geoffrey J., and John A. Goldsmith. 1994. *Ideology and Linguistic Theory: Noam Chomsky and the Deep Structure Debates*. London: Routledge.

Jackendoff, Ray S. 1977. *X-bar Syntax: A Study of Phrase Structure*. Cambridge, MA: MIT Press.

Johnson, C. Douglas. 1972. *Formal Aspects of Phonological Description*. The Hague: Mouton.

Kaplan, Ronald M., and Martin Kay. 1981. *Phonological rules and finite-state transducers (Abstract), Linguistic Society of America Meeting Handbook*. New York: Linguistic Society of America.

Karttunen, Lauri. 1993. Finite-state constraints. In *The Last Phonological Rule: Reflections on Constraints and Derivations*, edited by J. Goldsmith. Chicago: University of Chicago Press.

Koskenniemi, Kimmo. 1983. *Two-level Morphology: A General Computational Model for Word-form Recognition and Production.* Department of General Linguistics. Helsinki: University of Helsinki.

Lakatos, Imre. 1970. Falsification and the methodology of scientific research programmes. In *Criticism and the Growth of Knowledge,* edited by I. Lakatos and A. E. Musgrave. Cambridge: Cambridge University Press.

_____. 1977. *The Methodology of Scientific Research Programs: Philosophical Papers Vol. I.* Cambridge: Cambridge University Press.

Lakoff, George. 1987. *Women, Fire, and Dangerous Things: What Categories Reveal about the Mind.* Chicago: University of Chicago Press.

_____. 1993. Cognitive phonology. In *The Last Phonological Rule: Reflections on Constraints and Derivations,* edited by J. Goldsmith. Chicago: University of Chicago Press.

_____, and Mark H. Johnson. 1980. *Metaphors We Live By.* Chicago: University of Chicago Press.

Langacker, Ronald W. 1987. *Foundations of Cognitive Grammar, Vol. 1: Theoretical Prerequisites.* Stanford, CA: Stanford University Press.

_____. 1988. A usage-based model. In *Topics in Cognitive Linguistics,* edited by B. Rudzka-Ostyn. Amsterdam/Philadelphia: John Benjamins Publishing Co.

_____. 1991. *Foundations of Cognitive Grammar, Vol. 2.* Stanford, CA: Stanford University Press.

Mandelbrot, Benoit B. 1977. *Fractals: Form, Chance, and Dimension.* San Francisco: W. H. Freeman.

_____. 1983. *Fractal Geometry of Nature.* New York: W. H. Freeman and Co.

McCawley, James D. 1981. The syntax and semantics of English relative clauses. *Lingua* 53: 99–149.

_____. 1988. *The Syntactic Phenomena in English, Vol. I and II.* Chicago: University of Chicago Press.

McClelland, James L., David E. Rumelhart, and The PDP Research Group. 1986. *Parallel Distributed Processing—Explorations in the Microstructure of Cognition, Vol. 2: Psychological and Biological Models.* Cambridge, MA: MIT Press.

Peters, P. Stanley, and Robert W. Ritchie. 1973. On the generative power of transformational grammars. *Information Science* 6: 49-83.

Piaget, Jean. 1971a. *Biology and Knowledge.* Edinburgh: Edinburgh University Press.

_____. 1971b. The epigenetic system and the development of cognitive functions. In *Brain Development and Cognition: A Reader,* edited by M. H. Johnson. Oxford: Basil Blackwell.

Postal, Paul Martin. 1971. *Crossover Phenomena.* New York: Holt, Rinehart and Winston.

_____. 1974. *On Raising: One Rule of English Grammar and its Theoretical Implications.* Cambridge, MA: MIT Press.

Révész, György E. 1991. *Introduction to Formal Languages.* New York: Dover Publications.

Rosenbaum, Peter S. 1967. *The Grammar of English Predicate Complement Constructions.* Cambridge, MA: MIT Press.

Ross, John Robert. 1986. *Infinite Syntax!* Norwood, NJ: Ablex. [Publication of Ross (1967) *Constraints on Variables in Syntax.* Unpublished Ph. D. dissertation. MIT].

Rumelhart, David E., James L. McClelland, and The PDP Research Group. 1986. *Parallel Distributed Processing — Explorations in the Microstructure of Cognition, Vol. 1: Foundations.* Cambridge, MA: MIT Press.

Shannon, Claude Elwood, and Warren Weaver. 1949. *The Mathematical Theory of Communication.* Urbana: University of Illinois Press.

Smolensky, Paul. 1986. Information processing in dynamical systems: Foundations of harmony theory. In Rumelhart, J. E. McClelland and The PDP Research Group (1986).

Yamanashi, Masa-aki. 1995. *Ninchi-bumpo Ron* [Studies in Cognitive Grammar] Tokyo: Hitsuji Shobo.