JAKLIN KORNFILT AND NELSON CORREA

CONCEPTUAL STRUCTURE AND ITS RELATION TO THE STRUCTURE OF LEXICAL ENTRIES

1. INTRODUCTION

Recent work in generative grammar and the theory of the lexicon (Guerssel et al. 1983, Hale and Keyser 1987, Jackendoff 1983, Levin and Rappaport 1986) posits a level of linguistic representation, referred to as Lexical Conceptual Structure (LCS), at which syntactic and semantic properties of lexical items, and certain regularities between the two sets of properties are represented. LCS represents an improvement from the simple sets of features used to represent word meanings in early semantic theories (e.g. Katz and Fodor 1963) to a powerful notation for expressing predicate-argument and other kinds of semantic relations in natural language utterances. Current formulations of the LCS system (Jackendoff 1983) include processing mechanisms that capture semantic relations between expressions, such as synonymy, common-sense entailment, etc; the mechanisms include preference rules and default inference.

Independent efforts in artificial intelligence and cognitive science (Hobbs et al. 1987, Pylyshyn 1984, Schank 1973) have developed and argued for systems of symbolic representation as the basic mechanisms which underlie human cognition and action. The view advanced is that world and "common-sense" knowledge are encoded in these cognitive representations; organisms form the representations in response to sensory experience and apply syntactical rules to them to act and infer beliefs about the world. We shall use Schank's and Jackendoff's term *conceptual structure*(CS) to refer to these cognitive representations. LCS is to be seen as part of the larger system of conceptual structure which underlies human cognitive abilities; it encodes the lexical knowledge that native speakers have about the items in their lexicon.

While there is a substantial literature on the subject, we feel that previous work has failed to address the relation that exists between the semantic representation of an item and its most salient syntactic properties, such as its complement structure, the thematic or case relations of complements to the entry, and the alternations in complement structure that the item may exhibit. This observation is especially true of the work in artificial intelligence, like that of Schank, Wilks, and Hobbs, which does not even consider the problem. Jackendoff (1976, 1983, 1987) develops a detailed system of conceptual structure, but does not turn to the problem of realization of arguments from a given LCS expression, the realization of arguments is simply stipulated in Jackendoff's lexical entries. Hale and

Eric Reuland and Werner Abraham (eds), Knowledge and Language, Volume II, Lexical and Conceptual Structure: 79—118.

Keyser (1987), Levin and Rappaport (1986), Rappaport and Levin (1988), and Williams (1981) consider the problem of complement structures of predicates and alternation; however, they do not make a proposal concerning the form of LCS representations; Predicate-Argument Structure appears as the most basic lexical representation of an item in the lexicon. Similarly, in Fillmore's case grammar (1968, 1977), case frames are the basic lexical representations of predicates.

The objective of our work is to develop an explicit system of conceptual structure that may allow us to explore the LCS-syntax correspondence. CS will be used as a system of representation within which syntactic and semantic properties of lexical items may be captured.[1] In particular, we will study the problem of deriving a verb's predicate-argument structure or PAS (Marantz 1984, Levin and Rappaport 1986) from its lexical-conceptual representation. The PAS of a lexical item is closely related to the older notions of its theta-grid (Stowell 1981) and subcategorization frame (Chomsky 1965); the PAS of a predicate consists of a listing of its argument positions, as identified by thematic relation labels, together with an indication of the manner in which arguments of the predicate are syntactically identified with argument positions in the LCS.[2] It should be stressed that we do not take PAS to be a level of lexical representation, but rather a lexical-syntactic property of lexical items, determined by their LCS.

In this paper we draw from previous work on the subject, particularly Hobbs et al. (1987), Jackendoff (1972, 1976, and 1983), Pylyshyn (1987), and Talmy (1985) for the formulation of the system of conceptual structure, and Hale and Keyser (1985), Levin (1985), Levin and Rappaport (1986), Perlmutter (1978), and Rappaport and Levin (1988) for the study of the LCS-syntax correspondence. The paper is organized into two major parts which are relatively independent and may be read separately. The first part (sections 1—3) is devoted to the formulation of the system of conceptual structure. This part is more concerned with problems of knowledge representation for natural language than with linguistic issues per se; it places special emphasis on an adequate and explicit formulation of the CS system. The second part of the paper (sections 4—7) studies the relation between lexical-conceptual representations and the syntax of lexical items. The general syntactic framework that we assume in this work is the Government-Binding (GB) theory of Chomsky (1981, 1982).

Section 2 introduces conceptual structure as a system of mental representation and develops a syntax for it. Our approach is motivated by Jackendoff (1983), who makes the most explicit and linguistically motivated proposal. Section 3 contains preliminary remarks on the semantic interpretation of conceptual expressions.

Section 4 defines lexical-conceptual structure as a subset of conceptual structure and illustrates the correspondence between LCS and the syntax of predicates. Section 5 defines thematic relations as configurational relations on LCS and theta-roles, a syntactically relevant notion, as the collection of thematic relations that a given argument bears to the LCS expression in which it occurs. It thus becomes possible to state a simple projection rule that derives theta-roles, and ultimately predicate-argument structure from LCS. By disassociating the notions of "thematic relation" and "theta-role" we solve the problem noted by Jackendoff (1972, 1987) in connection with the proper statement and validity of the Theta-criterion in Government-binding theory (Chomsky 1981). Section 6 considers the relation between arguments in LCS and their syntactic realization. We state a projection rule that predicts the grammatical functions in which the arguments of a given predicate are syntactically realized.[3]

2. CONCEPTUAL STRUCTURE AS THE LANGUAGE OF MENTAL REPRESENTATION

Conceptual structure is a language used to represent a class of mental entities, namely concepts. The view we take is that at the level of abstraction appropriate for linguistic descriptions, concepts and meanings are not metaphysical objects, but instead have physical instantiations in cognitive codes or equivalence classes of mental states. The problem of justifying and formulating such a language has been approached from a number of disciplines, including psycholinguistics, cognitive psychology, artificial intelligence, and lexical semantics. Due to space limitations, we cannot discuss the relevant literature here (e.g. Pylyshyn 1984, Hobbs et al. 1987, Schank 1975, Wilks 1977, Jackendoff 1983). For a short overview, the reader is referred to Kornfilt and Correa (1990).

2.1. *Relation to Mental Representation*

In this section, we shall very briefly mention Jackendoff (1983). Jackendoff's (1983) theory of conceptual structure provides notions for representing predicate-argument structure, as well as other kinds of conceptual relations between expressions, such as restrictive and non-restrictive modification;[4] CS is identified as a level of representation, with an interface to other systems of cognitive representation. The conceptual well-formedness conditions, as well as the set of primitives in the system are assumed to be innate and universal.

Jackendoff (1983) adds to the hypothesis that language is mentally represented by the "mentalese" of conceptual structures the following cognitive constraint: "There must be levels of mental representation at which the information conveyed by language is compatible with information from other peripheral systems such as vision, nonverbal audition, smell, kinesthesia, and so forth" (p. 16). Implicit in this constraint is the earlier hypothesis that humans act on the basis of cognitive representations; the constraint explains how language interacts with other cognitive systems such as thinking, perception, and action. The constraint is augmented by the following simplifying hypothesis on the number of levels of mental representation: "There is a single level of mental representation, conceptual structure, at which linguistic, sensory, and motor information are compatible". This is known as the *conceptual structure hypothesis* (p. 17).

Summarizing, we take as a working hypothesis that conceptual structure is the language in which concepts and thought are represented. A central component of this hypothesis is the conceptual structure hypothesis. This is a strong hypothesis about the nature of the human mind; it may turn out to be empirically false, but gives, as a starting point, the simplest model of mental organization.

Our objective in the following subsections is to arrive at a more explicit system of conceptual structure than available in the lexical semantics literature. We strive for some degree of formality and explicitness in the formulation of the CS system, on the grounds that any interesting claims that may be made about conceptual structure or its relation to syntax (or rather, the verifiability of these claims) depend to a large degree on the details of the system. This will be seen clearly in the LCS-syntax correspondence rules that we formulate in the second part of the paper; these rules require direct reference to LCS expressions.

2.2. *An Example of Conceptual Structure*

Let us now consider what is often meant by conceptual structure in the lexical semantics literature, in order to provide a concrete idea of it. In (1a) we reproduce the conceptual structure for the verb "cut", taken from Hale and Laughren (1983). The notation is informal, although it could be formalized. The notion of CS implicit in this notation is not far from that found in the artificial intelligence (AI) literature; in (1b) we reproduce the informal rendering of Hobbs et al. (1987) for the verb "hit".

(1) a. cut: x produce CUT in y, by sharp edge coming into contact with y.

b. hit: x move into contact with y, with some force.

An LCS expression is a semantic decomposition of the event or state named by the verb into a number of simpler terms, such as the predicate "produce" and the entity "CUT" in (1a), or the predicate "move" and the notion of "force" in (1b). The constituent elements of the expression are not necessarily primitives of the notation, but rather may be further analyzable in terms of yet simpler concepts. For example, the element "CUT" in (1a) may be analyzed as a "linear separation in material integrity" (Hale and Keyser 1987). Ultimately, however, (the representation of the meaning of) each lexical item should be decomposable into the universal set of conceptual primitives made available by the system.

The decompositions in (1) intend to capture the core aspects of the verb meanings, without implying that all aspects of the meanings are represented.[5] Similarly, the primitive terms in the decomposition are defined by "typicality" and "centrality" conditions, while allowing the possibility of exceptions to some of the defining criteria (cf. Jackendoff 1983, chapter 5).[6]

The LCS expressions (1) reflect important aspects of the syntax of the corresponding verbs. An LCS expression identifies the participants in the situation named and defines, implicitly, their thematic relation to the situation. The expression (1a), for example, identifies two participants *x* and *y* in the event named, with *x* playing the active role and *y* a passive role. Identification of the participants is important syntactically since it determines the number of arguments that the verb selects. We will show below that, with an appropriate choice of LCS notation, the syntactic realization of these participants is also determined by the expression.

Since the work of Gruber and Jackendoff, it is common to identify the arguments of a predicate by the thematic relations they bear to the predicate. We will use thematic relation labels similar to those proposed by Gruber (1976) for the study of verbs of motion and location, including Agent, Theme, Source, Goal, and Location. In (1a), for example, *x* may be thematically related to the predicate as Agent and *y* as Theme.[7] Williams (1981) proposes that lexical representations of predicates include a *predicate-argument structure* (PAS), which is an unordered list of the argument positions of the predicate, identified by the thematic relations defined at those positions. This list has some further structure, to indicate the syntactic realization of the arguments. Williams (1981) includes a distinction between "internal" and "external" arguments of the predicate; other authors (Marantz 1984, Zubizarreta 1987) assume a further element of structure in the lexical representation, an identification of the "direct" internal argument. The argument structure associated with (1) is (2), assuming Williams' convention of underlining the external argument.

(2) cut:(<u>Agent</u>, Theme)

The objective of the annotations on lexical representations of argument structure is to express the manner of syntactic realization of arguments at D-Structure (or any other "initial" level of syntactic representation). For example, Williams' external argument is realized external to the maximal X-bar projection of the predicate; similarly, Marantz' direct argument corresponds to the Direct Object. With this understanding, the two arguments in (1a) are realized as noun phrases, *x* as the external argument in Subject position, and *y* as Direct Object.

Little previous work has addressed the relation between LCS representations and the notion of predicate-argument structure just outlined. Notable exceptions to this are (Williams 1981) and (Carter 1976). We claim that thematic relations and syntactic realization of arguments (i.e., PAS) need not be stipulated in lexical representations, since they are determined by LCS expressions, which must independently be represented in lexical entries.[8] LCS thus becomes the chief element in the lexical entry of a predicate, from which much of its syntactic behavior follows.

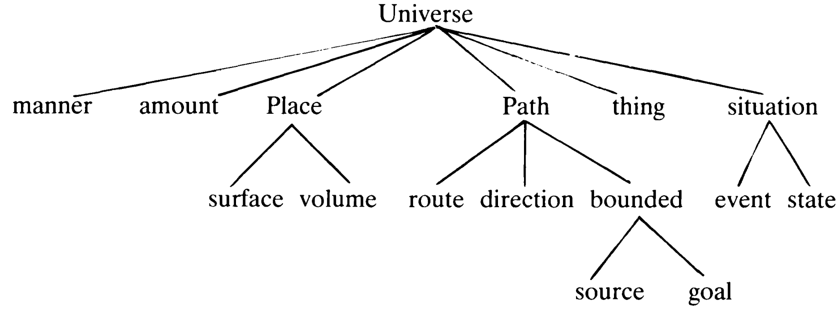
2.3. *Conceptual Domains*

In choosing a notation for LCS one must consider the kinds of entities to which the expressions may refer. The ontology assumed in LCS, as well as in most of the semantics of natural language literature is, as Hobbs (1985) points out, very close, if not isomorphic to the way we actually talk about it.[9]

For the purposes of this paper, and following previous approaches in lexical semantics, we will distinguish two dimensions of the LCS ontology: *Conceptual types* and *semantic fields* (cf. Gruber 1965, Jackendoff 1972, Talmy 1985). Conceptual structure is thus a system of expressions typed along two dimensions. The dimension of conceptual types defines the categories "event", "state", "thing", "path", "property" and "place"; expressions belong to one or another of these categories. The second dimension identifies *semantic fields* such as the "spatial", "temporal", "possessive", "circumstantial", and "identificational" (Gruber 1965, Jackendoff 1972, Talmy 1983). To these five fields we could add others, such as the "mental", to refer to abstract movement of information between cognitive agents (talk, tell, hear), between an agent and a medium (read, write, see), or within a cognitive agent (think, imagine, dream). Conceptual expressions are also identified with one of the semantic fields, thus achieving a cross-classification of expressions with conceptual types. A *conceptual domain* is any subdomain of conceptual structure. Conceptual types and semantic fields provide an initial set of domains for the classification of conceptual expressions.

The ontology of types implicit in Jackendoff (1983), and which we will assume, is shown in (3). The semantic fields assumed are the six listed above.

(3)



Note that the ontology (3) focuses on "situations", "paths", and "places", elaborating them much more than the other types. On the other hand, it ignores the (sub-)ontology of "things", which has been of major concern in AI research. The focus on the former set of types is a reflection of the chief concern of research in lexical semantics, to explain the syntax of verbs and prepositions. Situations correspond loosely to the kind of entities referred to by verbs, and paths and places to the entities named by prepositional phrases (cf. Gruber 1965, Jackendoff 1972, Talmy 1985).

The ontology (3) has linguistic import. There are correspondence principles, partly language specific, that determine which conceptual types may be expressed by which syntactic categories. First, clauses correspond to situations.[10] Within this class, there is a further subdivision into events and states. There are clear syntactic tests to distinguish between the two. Jackendoff (1983) gives the following test: Only events may occur in the context after "What happened/occurred/took place was that". This yields the contrasts in (4).

(4) a. What happened was that John got lost.
?What happened was that John was here.

b. What happened was that the window broke.
?What happened was that the window was broken.
(broken = adject.)

A preposition in English can express either a place or a path. A place is a point or region in space; in contrast, a path is a connected sequence of such points or regions. Places are identified by reference to an object, as with the prepositions "in", "on", "at" and "under" in (5a), or by the intransitive preposition "here" in (5b). Paths are rarely fully specified; their partial identification is accomplished by providing one or both end-points (source and goal), as in (5c), or a point intermediate, as in (5d). The points on a path may be identified by reference to an object, as in (5c—d), or to a place, as in (5e). As with places, paths may also be identified by intransitive prepositions, as in (5f).

(5) a. John is in/at/on the house.

b. John is here.

c. I went (from Boston) to New York.

d. I went via Schenectady.

e. Beth came running from under the bridge.

f. Mark went north/south/away.

The linguistic import of the six semantic fields assumed is considered in section 2.6.

2.4. *Conceptual Primitives*

We continue developing the Conceptual Structure Hypothesis by now considering the elements out of which the cognitive representations of word meanings are made.

We postulate that underlying all representations is a system of linguistically relevant primitive concepts. We shall use the term "conceptual primitive" to refer to the elements of meaning into which word and utterance meanings may be decomposed.[11] Our usage of the term "conceptual primitive" follows its normal use in lexical semantics (cf. Jackendoff 1976).

For discussion of semantic primitives and their nature in artificial intelligence and cognitive science literature, see Schank (1973), Wilks (1977), and Olson and Bialystok (1982). In contrast, Jackendoff (1976, 1983), while one of the chief advocates of the use of primitives in lexical-semantic representations, does not explain the method that might be used for choosing a set of primitives or the alternatives there may be to primitives. The primitives are simply postulated in cognitive representations and their functioning only illustrated by way of examples.

In discussing the primitive elements of cognitive representations we must distinguish the primitive concepts just alluded to from other logical symbols indispensable in the notation, such as variables and operators. We assume that the notation for conceptual structure has at least these three classes of elements: primitives, variables and operators. The primitives discussed in this section correspond to the constant, function, and predicate symbols of the notation.[12] Variables and operators are discussed later on.

Beginning with the work of Gruber (1965) and Bierwisch (1967), lexical semantics first focused its attention on the semantics of expressions of spatial location and motion. Accordingly, the vocabulary of spatial cognition was taken for analyzing the meanings of lexical items in this domain.[13] This vocabulary includes terms like "on", "in", "under", "at", "up", "down", "front", and "back" to refer to physical places relative to objects; "to", "from", "towards", and "away" to refer to points or orientations of paths; and predicates like "cause", "stay", "go", "orient", and "be" to refer to events and states of location and motion. Among these terms, the predicate "cause" represents the potential element of causation in spatial events. We may thus postulate the basic vocabulary (6) of conceptual primitives, corresponding to the terms just listed; this vocabulary is found in Jackendoff's (1983) system.

(6) a. ON, IN, UNDER, OVER, AT, BETWEEN, UP, DOWN, FRONT, BACK, . . .

b. TO, FROM, TOWARD, AWAY, VIA, . . .

c. STAY, GO, ORIENT, BE, CAUSE, LET, . . .

The primitives (6a) and (6b) are unary functions; they require an argument or reference object to identify a place or a path, respectively. The place-functions (6a) may identify different places with reference to the same object (7a—c). Likewise, the path-functions (6b) partially define paths by specifying their source or goal endpoints (7d—e), or a middle point (7e).

(7) a. on the table.
b. under the table.
c. at the table.
d. from the table.
e. to the table.
f. by the table.

The event and state primitives (6c) are used in the analysis of the verbs of location and motion in (8). In these examples, "remain" may be analyzed as 'STAY at location'; "fall" as 'GO along path downwards'; "point" as 'ORIENT in direction of path'; "lay" as 'BE at location'; and "throw" as 'CAUSE to GO away'.

(8) a. The rock remained at the bottom of the hill.
b. The rock fell from the cliff.
c. The sign points north.
d. The vase lays on the table.
e. John threw the rock.

Much of the interest in the hypothesis that word meanings are mentally represented comes from the ability to formulate and postulate a collection of primitive concepts. The interest in the hypothesis is even greater if the collection can be shown to be both innate and universal. What is important are the combinatorial possibilities of the primitives and the semantic operations they sanction (e.g. inferences), rather than the names chosen for the primitives. Conceptual structure is an algebraic system.

The identification of the set of primitive concepts is one of the major open problems in theories that assume the existence of a primitive set. The exact set of primitives in the human conceptual system has not been identified in any of the work we are aware of, and we shall make no concrete proposal in this regard.[14] The primitives identified above from the spatial field arguably play a major role in the human conceptual system, but clearly a larger set is needed. We expand on the meanings attached to the primitives listed in (6) in tables (9) and (10). See Gruber (1965) and Jackendoff (1983) for discussion of some of these primitives.

The principal primitives that yield expressions of types "path" and "location" are those in (9).

(9)

Primitive	**Arguments**	**Intended denotation**
(paths)		
PATH	none	Some unspecified path.
TO	x	A path whose goal end-point is at thing or location x
TOWARD	x	A path whose goal end-point is in the direction of thing or location x.
FROM	x	A path whose source end-point is at thing or location x.
AWAY	x	A path whose source end-point is in the direction of thing or location x.
VIA	x	A path which passes by thing or location x.
(locations)		
HERE	none	The place of the speaker.
AT	x	The location occupied by thing x.
IN	x	The volume contained by thing x.
ON	x	The upper or outer surface of thing x.
UNDER	x	The space under or inside thing x.
OVER	x	The space over or outside thing x.

The primitives that under lie spatial events and states are those in (19).

(10)

Primitive	**Arguments**	**Intended denotation**
(states)		
BE	x, y	The state of thing x being at location y.
GOext	x, y	The state of thing x extending along path y.
ORIENT	x, y	The state of thing x being or iented in the direction of path y.
(events)		
GO	x, y	The event of thing x traversing path y.
STAY	x, y	The event of thing x staying at location y.
CAUSE	x, y	The event of thing or situation x causing situation y.
LET	x, y	The event of thing or situation x enabling situation y (e.g. permissive causation).

The lists of primitives in (9—10) is not complete, nor is it intended to be; it will be extended below, with the addition of some new primitives, and we believe that further extension will still be necessary. An important feature of the lists is, however, that the number of primitives assumed is small, when compared to the number of verbs that may be found in particular languages for naming spatial situations. In English, for example, a rough estimate would put the number of such verbs in the order of several hundreds (the total number of English verbs listed in Longman (1978) is nearly 8,000 (Boguraev and Briscoe 1987)).

Although it is an empirical issue, the work of Carter (1976) and Jackendoff (1972) supports the claim that the number of primitives is small, significantly smaller than the number of lexical items whose lexical meanings may be encoded using the primitives.

2.5. *Primitives as Partial Functions on Conceptual Domains*

It is implicit in the informal definitions of the conceptual primitives in (9) and (10) that a conceptual primitive is a function on the available set of conceptual domains. The function is in general a partial function, since the information it encodes is partial information about specific situations in the world and, furthermore, people don't have perfect knowledge about all situations.

A primitive may thus be a constant (O-ary function) belonging to a given domain, or a function of one or more arguments, from the Cartesian product of a number of conceptual domains into another domain. This is shown in (11). The primitives PATH and HERE in (9) are constant functions; the respective maps are (11a—b); the primitives are in the domains "path" and "place" respectively. The primitive TO of (9) is a function from the domain "thing" or "place" into "goal-paths", as the map (11c) shows. To give one more example, the event primitive GO in (10) is a function of two arguments, from the product (thing × path) into the domain "event", as in (11d).

(11) a. PATH: $\emptyset \rightarrow$ path

b. HERE: $\emptyset \rightarrow$ place

c. TO: thing $\cup$ place $\rightarrow$ goal

d. GO: thing $\times$ path $\rightarrow$ event

According to Jackendoff's (1983) referentiality principle, the default use of a conceptual primitive is referential; a primitive is used to identify objects, manners, places, paths, events, and states in the space of concepts, and hence its role is that of a constant or function symbol in a logical language. A predicative or non-referential use of a primitive is obtained in the system by addition of a conceptual marker TYPE, which indicates that the primitive is used for categorization purposes.

2.6. *Generalization of Primitives to Non-spatial Semantic Fields*

An important insight in the work of Gruber (1965) on lexical semantics is that expressions from the spatial domain in natural language (12a) appear used in other semantic domains, as the possessive (12b), temporal (12c), or identificational (12d).

(12) a. The bird flew to the tree-top.

b. The inheritance went to Bill.

c. The class was moved to 3:30.

d. Paul changed his profession to lawyer.

In these examples we want to focus attention on the use of the preposition "to". In all cases, the preposition serves to indicate the goal endpoint of a path. However, depending on the nature of the reference object and the verb of which it is complement, the path may be said to be spatial (12a), possessive (12b), temporal (12c), or identificational (12d). The primitive function TO which is lexicalized by the preposition, then, is being applied in different semantic fields. Similarly, the verbs "go", "move", and "change" used in (12b—d) in non-spatial fields are typical verbs of (spatial) motion. The primitive GO which presumably underlies all these verbs, then, is also applying different fields.

We may formalize the observation above by noting that the signatures assigned to the primitives in (11) are specified in terms of conceptual types alone. The primitives defined may be said to be non-specific regarding semantic field. We obtain a new set of primitive concepts by restriction of the signature of the functions (domain and possibly range) to narrower conceptual domains, as given by specific semantic fields. This is illustrated in (13), where the variants $\mathrm{PATH}_{spatial}$ and PATH_{temp} of PATH; and $\mathrm{GO}_{spatial}$ and TO_{poss} of GO and TO, respectively, have the signatures given. The examples (12) serve to illustrate the use of the new primitives (13c—d).

(13) a. $\mathrm{PATH}_{spatial}$: $\emptyset \rightarrow (\text{path} \cap \text{spatial})$

b. PATH_{temp}: $\emptyset \rightarrow (\text{path} \cap \text{temporal})$

c. TO_{poss}: $(\text{thing} \cup \text{place}) \cap \text{possessive} \rightarrow (\text{goal} \cap \text{possessive})$

d. $\mathrm{GO}_{spatial}$: $\text{thing} \times (\text{path} \cap \text{spatial}) \rightarrow (\text{event} \cap \text{spatial})$

The inferences and other conceptual operations that a primitive sanctions may well depend on the particular semantic field specialization of the primitive. This can be seen clearly in the primitives PTRANS and MTRANS of Schank's conceptual dependency, which may be analyzed in our primitive collection as $\mathrm{GO}_{spatial}$ and GO_{mental}, respectively, assuming the semantic field of "mental" activities. The inferences that these two primitives sanction are different. While in the case of spatial motion (14a) it is a valid inference that at the beginning of the event the theme object is at the source end of the path (i.e. Boston), and at the end of the event it is at the goal (NYC) and no longer at the source, in (14b) the theme (the proposition "that Mary came") is both at the source (John) and goal (Bill) endpoints of the mental path at the end of the event.

(14) a. The train went from Boston to New York.
b. John told Bill that Mary came.

The parametrization of conceptual primitives by semantic fields yields collections of related conceptual functions, like GO_x for x equal to the different semantic fields. Parametrization increases the number of primitive functions available, by a factor equal to the number of semantic fields assumed. Some verbs like "arrive" lexicalize a primitive that may apply in several semantic fields; the two possibilities in (15a) are used in "the train arrived (to the station)" and "a time will arrive (for revolution)". Other verbs may lexicalize primitives in a particular semantic field and no others. For example, the conceptual structure of "tell" may be given as (15b), so that the verb always refers to an event in the mental domain.

(15) a. arrive: GO_{sf}(x, TO_{sf}($HERE_{sf}$)), for sf = spatial, temporal (x arrive)
b. tell: CAUSE(x, GO_{mental}(REP(y), $PATH_{mental}$[FROM(x), TO(z)])) (x tell z that y)

The core set that we assume in the domains of paths, locations, and situations is of about 30 primitives; it is essentially that assumed by Jackendoff (1983). We shall make no proposal regarding the primitives there may be in the domains of things, properties, and manners (roughly corresponding to nouns, adjectives, and adverbs). Since our concern in this paper is mainly with the LCS and syntax of verbs, we leave untouched that section of the system of conceptual structure.

2.7. *Composition of Conceptual Primitives*

Conceptual primitive functions may be combined with one another to form conceptual expressions. We shall assume three forms of composition, functional, conceptual, and operator-function composition, which differ in basic respects. *Functional composition* is the basic type. A CS expression X may be composed with a primitive P at position i if X is contained in the domain of position i, as given by rule (16). Functional composition is obligatory, in the sense that every primitive in a well-formed expression must appear with the appropriate number of arguments; that indicated by its arity. We say that P is the *head* of the expression in (16); the conceptual domain is determined by P.

(16) P(. . . , X, . . .), X at the i-th argument position of P,
if D(X) is contained in the domain D_i of the position.

The second form of composition is *conceptual modification.* This form of composition has two variants, restrictive and non-restrictive, which differ in their semantics, as in natural language expressions. A conceptual primitive function may be composed with an unbounded number of modifiers, which further restrict the denotation of the primitive (restrictive modification), or add comment-type information to the denotation (non-restrictive or appositive modification). For example, event and state functions may be composed with expressions of manner, place, purpose, and time. The notation we use for conceptual modification is (17), to distinguish it from functional composition. The modifying items are inside square brackets and are all optional. We say that E is the head expression, and that the conceptual domain of the entire expression is that of E.

(17) $E[X_1, \ldots, X_n]$, where E and X_i, $i = 1, \ldots, n$, are conceptual expressions.

An important use of restrictive conceptual modification is in sharpening the denotation of path functions. These functions identify paths, but only partially. The generic PATH constant, for example, denotes an arbitrary member of the collection of paths and hence is highly unspecific. The primitive, however, may be modified by path expressions that specify source, goal, and direction. Thus, while in (18a) the path is syntactically left completely unspecified and discourse context has to supply the intended reading, it is considerably sharpened in (18b).

(18) a. John went.
GOspatial (JOHN, PATH)

b. John went from Boston to New York via Great Barrington.
GOspatial (JOHN, PATH[FROM (BOSTON), TO(NY), VIA (GB)])

The third form of composition is *operator-function* composition. The notation we adopt for this is (19), in which the operator O appears next to the variable it quantifies. The operators we assume include generalized quantifiers and interrogatives. The operator-function expression in (19) is obtained from the simpler expression $F(\ldots, x, \ldots)$, where x is a free variable in the expression. The variable is said to be *bound* by the operator in (19).

(19) $F(\ldots, O : x, \ldots)$, where O is an operator and x a free variable in $F(\ldots, x, \ldots)$

Notice that the quantified expressions that may be built according to (19) are quantified in situ; i.e. the quantifiers appear at the location of the LCS variable they bind. These expressions thus resemble the D-Structures defined in Government-binding theory, where operators have not yet been raised. The scope of these quantifiers is *not* defined by the notation; instead it is established syntactically, e.g. according to Williams' (1986) proposal.

The system of conceptual structure we have sketched so far closely follows that outlined by Jackendoff (1983). Two minor differences with the system presented here are the simplified notation we assume and the provision for operator-function composition. Regarding the first, Jackendoff customarily subscripts the conceptual type of the constituent to each constituent in a CS expression.[15] We may omit the conceptual domain subscripts from our notation since the domain of a constituent is always clear from the primitive that heads the constituent.

The second difference between the two systems remedies a major drawback in Jackendoff's system as a "language for thought" and natural language semantics. Jackendoff's theory provides no mechanisms for expressing relations of quantification in conceptual expressions. We see this drawback more as a matter of omission than of principle or approach; we have sought to overcome it with the provision for operator-function composition.

2.8. *Conceptual Variables*

As noted earlier, a system of conceptual structure includes an inventory of variables, which are primitives of the CS notation and may appear in conceptual expressions like expressions of any other type. A variable may be bound by an operator expression, or it may be substituted by another expression, for example, as in "theta-role assignment".[16] In the latter case the variable acts as a place-holder in the LCS expression. We assume that the inventory of variables is typed, so that variables may be specified to belong to certain conceptual domains (e.g. the domain of ANIMATE things); selectional restrictions on variables arise by this means.

A variable in a conceptual expression is distinguished uniquely by its name. It is necessary to distinguish between the notions of a variable in an expression and *occurrences of the variable* in the expression. A conceptual expression may contain several variables, with perhaps several occurrences of some variables. When a variable is substituted by some value or bound by an operator, each occurrence of that variable is substituted by the value or bound by the operator, respectively.

A variable is *free* if it is not bound by an operator. We will adopt the convention of preceding every LCS expression with at least one free variable in it by the symbol LAMBDA and a list of the free variables in the expression. As in the lambda calculus of Church (1940), LAMBDA is the abstraction operator and the list of variables following the operator is the *abstraction* of the expression. This notational convention is illustrated in (20).

(20) LAMBDA x y z F(x, y, z)

An expression with free variables in it is said to be *open*; otherwise it is said to be *closed*.

Since conceptual structure is a typed system, each expression in it is associated with a conceptual domain. Variables have always a conceptual domain associated, which places a restriction on the class of conceptual expressions that may be substituted for the variable. Substitution of a variable by an expression in the wrong conceptual domain would result in an expression which is syntactically ill-formed, according to the rules for LCS formation which we now consider.

2.9. *Syntax of Conceptual Structure*

The above remarks are summarized in the following definitions, which characterize the syntax of conceptual structure.

(21) Let C be the universe of concepts. A *conceptual domain* is any subset D of C.

The universe C of concepts includes individuals, collections of individuals, collections of collections, and so on.

(22) A *conceptual primitive function* is a map F: $D_1 \times \ldots \times D_n \rightarrow D_0$,
where $D_0, \ldots, D_n$ are conceptual domains; $n \geqslant 0$. $D_1 \times \ldots \times D_n$ is called the *domain* of the primitive and D_0 its range or *type*.

If $n = 0$ in (22) we say that F is a constant function. The definition (22) of primitives covers what are known as constant, function, and predicate symbols in the predicate calculus. We will assume the details of the particular map effected by a primitive are not part of its definition; these details are perhaps relativized to situations, and have to be acquired in the course of experience.

Typically, the signature of a primitive is defined indicating only the conceptual types of the signature. As we have seen, the types include things, events, states, paths, etc. (cf. the ontology (3) above). A generic definition of a primitive, in terms of types alone, may be modified by restriction of one or more of the signature types to a semantic field S, in which case we get a new primitive function, specialized to field S. For example, the primitive GO: thing x path → event may be specialized to the "possessional" or "spatial" fields, as in (23). The function GO_{poss} might be the one involved in capturing the meaning of "The house went to John", while the analogous function $GO_{spatial}$ would be involved in the meaning of "John went to the house".

(23) a. GO_{pass}: thing $\times$ (path $\cap$ posessive) $\rightarrow$ event $\cap$ posessive
b. $GO_{spatial}$: thing $\times$ (path $\cap$ spatial) $\rightarrow$ event $\cap$ spatial

The denotation of a primitive may be restricted or predicated upon by conceptual modifier expressions attached to the primitive. A system of conceptual structure must hence state, for each primitive, the range of conceptual modifiers that it allows. We will assume that the modifiers that a primitive allows are determined by the type of the primitive, so that it is not necessary to list the range of allowed modifiers individually for each primitive, but only for each conceptual type. The signature and range of modifiers of a primitive completely determine its syntactic behavior at the level of conceptual structure.

(24) *A system of conceptual structure* is a five-tuple (D, P, V, O, R), where D is a collection of conceptual domains, including conceptual types T and semantic fields S; P a finite collection of primitives; V an infinite collection of typed variables; O a collection of logical operators; and R a set of associations $\langle T_i, M \rangle$; T_i a type in T and M a subset of T. We take M to be a specification of the conceptual modifiers that primitives of type T_i allow.

Definitions (21—24) capture the essentials of a system of conceptual structure in the sense of Gruber and Jackendoff, and allow us to state explicitly a syntax for conceptual expressions.

(25) Let CS = (D, P, V, O, R) be a system of conceptual structure. A *conceptual expression* of CS is any expression that may be built by application of rules (i)—(v) a finite number of times:

(i) Every constant primitive function in P is a conceptual expression.
Likewise, every variable in V is a conceptual expression.

(ii) Let C_i be a conceptual expression with domain $D(C_i)$, for i = 1, . . . , n, and let F: D_i x . . . x D_n $\rightarrow$ D be a conceptual primitive function, with D (C_i) contained in D_i. Then F $(C_1, \ldots, C_n)$ is a conceptual expression.

(iii) Let F(. . . , x, . . .) be a conceptual expression, with x a free variable; let O be an operator. Then F(. . . , O : x, . . .) is a conceptual expression and x is said to be bound by O. The scope of O in F is undefined.

(iv) Let C_i be a conceptual expression with domain $D(C_i)$, for i = 0, . . . , n. Let M be the domain of conceptual modifiers

allowed by D_0, as given by R, and let D_i be contained in M, for $i = 1, \ldots, n$.
Then $C_0[C_1, \ldots, C_n]$ is a conceptual expression, with head concept C_0, and modifiers C_i, $i = 1, \ldots, n$.

(v) Nothing else is a conceptual expression.

The LCS syntax (25) takes into account the fact that the denotation of an LCS expression E need not be precisely known. Thus, in step (ii) of the LCS formation rules we do not use the condition that E be in the domain D_i, but rather that the conceptual domain D(E) of E be contained in D_i. In general, all we know about the expression E is the domain D (E) to which it belongs; this domain is given by the type of the main functor of the expression.

3. INTERPRETATION OF CONCEPTUAL STRUCTURES

The notion of conceptual structure developed in section 2 is purely syntactic; conceptual expressions are formulas built from the set of primitives in the system, according to the formation rules (25). Neither the primitives, nor the formulas built out to them were given precise semantic interpretations in that section.

It is important to point out that the denotational interpretation of primitives in a system of conceptual structure is an issue orthogonal to whether that system is based on lexical decomposition into a system of primitives or not. A theory of conceptual structure need not concern itself exclusively with the form of conceptual representations and the formal relations between them, what we might call an "operational" semantics, but rather is to be more highly valued if, in addition, it provides a framework for assigning denotations of its primitives and expressions — i.e., a "denotational" semantics for its constructs.

Having the interpretations of primitives, an interpretation function $I: CS \to C$ may be defined from the set CS of formulas to the domain C of interpretations (concepts), that assigns an interpretation I(c) to each expression $c \in CS$. This is carried out in much the same way that interpretations are assigned to formulas in predicate logic, assuming the principle of compositionality. We shall not consider this in our exposition.

It is important to observe that conceptual structure expressions cannot be interpreted directly, without consideration of the syntactic component of the grammar. The reason is that the syntax of LCS notation does not define the scope of operators in the notation. This makes it necessary to postulate a syntactic level of representation, like Logical Form (LF) in Government-binding theory, at which the interpretation of quantified expressions is defined.

4. LEXICAL CONCEPTUAL STRUCTURE

The lexicon in a generative grammar is a list of associations between a lexical form, its lexical-conceptual representation, and other information, such as lists of syntactic, morphological, and phonological features. The conceptual expression associated with an entry is its *lexical-conceptual structure* (LCS); this terminology is that used by (Levin 1985); we say that the lexical item *lexicalizes* the LCS associated with it.[17]

In this section we illustrate the relation of the lexicalization patterns of English predicates to the syntax of the predicates.

4.1. *Lexicalization Patterns: A Sample LCS*

Let us start by considering the lexical entry for the verb "come", as in (26).

(26) ⟨**come**, +V . . . ,
LAMBDA x GO(x, PATH[TO(HERE)]), . . .⟩

The entry lists the lexical form as the first element of the entry, the syntactic category of the entry and remaining syntactic features as the second element, the LCS and semantic features as third element, and all remaining features (morphological and phonological) in the fourth position.

The LCS in (26) has a single abstraction variable x. The reading of this variable is supplied syntactically, by the syntactic constitutent realized at the position (grammatical function) specified for this variable. For example, if the variable is to be realized in direct-object position (assuming an unaccusative analysis of the verb), then the reading is supplied by whatever constituent appears at or binds the direct-object position. This is shown in (27), where "e-i" is the trace of the Subject NP "John."

(27) a. John-i came e-i.

b. GO(x, PATH[TO(HERE)]), with x = JOHN.

Since the variable x in (26) appears as first argument of the primitive GO, its conceptual domain is limited to the type "thing", and any reading bound to x must be of that type. We argue that the syntactic realization of LCS variables is predictable by a lexical redundancy rule. The realization rule formulated in section 6 takes into account the conceptual domain of the variable to determine its syntactic realization.

The second argument of GO in (26) is the primitive PATH, which we let denote an arbitrary entity of conceptual type "path". This primitive is modified by having one of its end-points partially specified, by the expression TO(HERE). The latter expression denotes an arbitrary member of the class of paths whose goal endpoint is at one of the locations denoted by the spatial constant HERE. The denotation of HERE is, in turn, the immediate place of the speaker, which may be identified as the place at which he stands, or also the room, city, or country of the speaker, depending on pragmatic context information.

4.2. *Optional Arguments and Adjuncts*

Since the conceptual modifiers of an LCS expression are optional and their number has no upper limit, the denotation of PATH in (26) may be further specified by supplying syntactically *optional arguments* to the verb, as in (28). We postulate that a verb has optional arguments only when their readings may be attached as conceptual modifiers to a subexpression inside the verb's LCS. Optional arguments may be attached to the subexpression as long as their readings do not conflict with the modifiers already attached to the same subexpression (i.e., yield an empty denotation, when taken together). The semantic anomaly in (28c) results from the inconsistent specification of the goal endpoint of the path.

(28) a. John came to the party.
GO(JOHN, PATH[TO(HERE), TO(THE PARTY)])

b. John came from New York.
GO(JOHN, PATH[TO(HERE), FROM (NEW YORK)])

c. ??John came to New York. (when uttered in California)
GO(JOHN, PATH[TO(HERE), TO(NEW YORK)])

The primitive GO in (26) is of type event, and thus may be modified by expressions of time, manner, instrument, accompaniment, etc; these are the typical conceptual modifiers for expressions of type event. We let this kind of modifiers be supplied syntactically by *adjuncts* of the verb, as in (29). Time and aspect markers of the verb lexicalize some of the conceptual modifiers attached to the primitive GO. The marker PAST in (29a), for example, is given by the past form of the verb. Notice that since the readings of adjuncts appear as modifiers in conceptual expressions, they are always optional and their number in a given sentence unbounded.

(29) a. John came at 3 o'clock with a friend.
GO(JOHN, PATH[TO(HERE)])[PAST, AT(3PM), WITH(FRIEND)]

b. ??John came tomorrow.
GO(JOHN, PATH[TO(HERE)]) [PAST, TOMORROW]

As before, the conceptual modifiers determined by adjuncts must be compatible with modifiers already in place. The semantic anomaly in

(29b) is due to the disjoint readings of the markers PAST and TOMORROW.

4.3. *Typed Variables*

Next we will consider the possibility of restricting the class of readings that may be substituted for a variable in an LCS. Since conceptual structure provides an infinite collection of typed variables, this possibility is already built into the system. We saw in (26) that the reading of the argument bound to the variable x must be of type "thing", due to the position the variable occupies. Hence, the domain required of x in the expression must be "thing" or one of its subdomains.

In general, we will allow the LCS of an entry to restrict the conceptual domain of any variable in it, as long as the specified domain does not conflict with the position the variable occupies. All these restrictions turn out to be *selectional restrictions* on the variable. The notation we will use for the restrictions is attaching a list of conceptual modifiers to the variable, each of which denotes a particular conceptual domain. This is shown in the lexical entries of (30). The restriction on x in (30a) indicates that only entities in the domain of animate things may be appear as arguments of "swim." Similarly, the restriction in (30b) indicates that the argument of the preposition "in" must be a bounded area or volume.[18]

(30) a. ⟨**swim**, + V . . . ,
LAMBDA x CAUSE(x[ANIMATE],
GO(x, PATH[VIA(WATER)])) [BY . . .], . . .⟩

b. ⟨**in**, + P . . . , LAMBDA x IN(x[BOUNDED AREA OR VOLUME]), . . .⟩

It might be that some of the selectional restrictions on a variable arise because of restrictions on certain occurrences of the variable in the expression. In (30a), for example, if we assume that the LCS given captures the sense "swim" given by "*To move through water by moving limbs and/or tail*" (Longman 1978), there are certainly occurrences of the variable inside the BY subexpression, and it might be that the overall condition of animacy on the reading substituted for x is established by one of the later occurrences of x (e.g. a real-world condition that only animate things have limbs and move them purposefully). In this case, the selectional restriction stated directly on the first occurrence of the variable is superfluous.

We will let LCS state directly any linguistically relevant selectional restrictions on variables, as in (30). Continuing with the hypothesis of "fully-specified lexical entries" in the lexicon, proposed by Jackendoff (1975), we might let common-sense knowledge act as a redundancy system that captures the regularities between specified and implicit lexical information.

Consider the first approximation (31) to the LCS of the verb "extend".

(31) ⟨**extend**, + V . . . , LAMBDA x y GOext(x, PATH[y]), . . .⟩

The LCS in (31) contains two variables x, y, which by their positions must belong to the domains "thing" and "path", respectively. The use of the variable y in (31) indicates that the expression PATH has an obligatory conceptual modifier; example (32a) is ill-formed since no reading is supplied to y. The syntactic realization rule might indicate that the variables are to be realized as (unaccusative) direct object and prepositional complement. This then allows for the use of the LCS in the expression (32b); variable x ends up bound to the time period "the summer" (a sort of "thing"), while y is bound to the goal-path "into October".

(32) a. *The road extends.

b. The summer-i extended e-i into October.
GOext(THE SUMMER, PATH[TO(IN(OCTOBER))])

However, nothing in the LCS of (31) prevents it from being used in the expressions (33), since in both cases the readings bound to y are in the domain "path". This shows that not every kind of path is a valid reading for the second argument of "extend"; we must then hypothesize that the verb lexicalizes a restriction on the domain of its second variable, e.g. as given in (34).

(33) a. *The summer-i extended e-i from June.

b. *The summer-i extended e-i away.

(34) ⟨**extend**, + V . . . , LAMBDA x y GOext(x, PATH[y[GOAL]]), . . .⟩

The ability to impose conceptual domain restrictions on variables extends the descriptive power of LCS a great deal. This power seems needed, but it must be used with extreme care, in order not to trivialize the LCS-syntax correspondence.

5. THEMATIC RELATIONS IN LINGUISTIC THEORY

We show in this and the following section that predicate-argument structure (PAS) is fully predictable from LCS (notice that this claim is significant only to the extent that there is concensus on the PAS and LCS representation of lexical items — a non-trivial requirement in current linguistic theory, as has already been suggested in the previous sections).

5.1. *Thematic Relations*

The arguments of a predicate enter into certain *thematic relations* with the predicate, each by virtue of being identified as an argument of the predicate.

The core set of thematic relations we assume is that proposed by Gruber (1965) for the analysis verbs of motion (35). The notion of causation present in the definition of Agent is a naive one, that which may be involved in "common-sense" explanations of causation in physical phenomena (cf. Hobbs 1985).

(35)	Agent	An entity which causes an event or state.
	Theme	The object undergoing motion or being located.
	Location	The location of the object located.
	Source	The initial location of the object undergoing motion.
	Goal	The final location of the object undergoing motion.

The relations (35) must be augmented with some additional ones, such as Instrument, Experiencer, Situation, and Path, to label arguments which do not fit well any of the initial set of relations.

The thematic relations postulated for the spatial semantic field are generalized, so that they also apply in fields outside the spatial one. The arguments of verbs naming situations which do not have to do with motion or location are also identified with the same basic set (35). In the possessive field, for example, the Theme may be identified as the object whose possessor or change of possessor is identified.

5.2. *Projection of Thematic Relations from LCS*

We adopt here the view that thematic relations are non-primitive notions of linguistic theory, which are configurationally determined in conceptual representations, in a manner similar to how grammatical functions (like subject and object) are determined by syntactic representations in certain theories, like Government-binding.[19] In particular, we claim that each subexpression in a conceptual expression bears a thematic relation to the whole expression. The thematic relation is determined by the primitive the subexpression appears as an argument of, by the argument position within this primitive, and possibly also by the conceptual type of the subexpression. This is stated precisely in (36).

(36) **Thematic Assignment rule**:
A subexpression E in LCS L is assigned thematic relation *r* if
(i) P is the conceptual primitive immediately containing E,
(ii) the argument position of E in P is i, and
(iii) $r = \mathrm{R}(\mathrm{P}, \mathrm{i})$, where R is the map (37) that follows.

(37)

Primitive	**Position**	**Thematic relation**
BE	1	Theme
	2	Location
GOext	1	Theme
ORIENT	1	Theme
GO	1	Theme
	2	Path
STAY	1	Theme
	2	Location
LET, CAUSE	1	Agent
	2	Situation
AT, IN, ON, UNDER	1	Location
TO, TOWARD	1	Goal
AWAY, From	1	Source

The projection rule (36), together with the associated map (37), define for each subexpression in a conceptual expression the thematic relation that the subexpression bears to the complete expression. This is illustrated in (38), where the subexpressions JOHN, A LETTER, and TO(NYC) are assigned the relations Agent, Theme, and Path; similarly, the subexpression NYC is assigned the relation Goal.

(38) John sent a letter to NYC.
CAUSE(JOHN, $\text{GO}_{\text{spatial}}$ (A LETTER, TO(NYC)))

We have omitted from the projection rule (36) the possibility that the thematic relation of a subexpression depends not only on its LCS position, but also on the specified conceptual type. The definition of thematic relations is purely configurational. For example, in the LCS of (38), the subexpression TO(NYC) is assigned the thematic relation Path, although it could also be called a Goal, since the type of the subexpression is "goal-path."

It is also worth noting that the configurational definition (36) of thematic relations depends only on local LCS context, so that it is possible for two distinct subexpressions of an LCS to bear the same thematic relation to the whole expression. This is akin to how a single syntactic representation may contain several configurationally defined subjects, objects, etc. For example, the subexpressions BREAD and SHARP EDGE in (39) are both identified as Themes.

(39) John cut the bread.
CAUSE(JOHN, GO_{ident}(BREAD, TO(AT(CUT))))
[BY(CAUSE(JOHN, $\text{GO}_{\text{spatial}}$(SHARP EDGE, TO(IN(BREAD)))))]

Finally, (36) and (37) do not define a thematic relation for every possible LCS configuration. The thematic assignment rule is not defined, for example, for arguments of the primitives BY and VIA. Likewise, the rule does not define some intuitively plausible relations like Experiencer or Patient. Experiencer is presumably "an argument of an as yet unexplored state-function having to do with mental states" (Jackendoff 1987, p. 378). The rule is hence defined only for particularly prominent conceptual configurations.

5.3. *Theta-roles and the Theta-criterion*

In the case of a conceptual variable x occurring in the LCS expression, with occurrences $x_1, \ldots, x_n$, say, each occurrence of the variable induces an independent thematic relation of the variable to the expression. We call the collection of thematic relations thus associated with the variable the *theta-role* of the variable (40). This theta-role is said to be assigned to the argument that is syntactically identified with the variable.

(40) **Theta-role rule**:
The theta-role associated with a variable x, with occurrences $x_1, \ldots, x_n$, is $\Theta = \{r_1, \ldots, r_n\}$, where r_i is the thematic relation defined by (36) for x_i.

The manner in which (40) operates is illustrated in (41), where the theta-roles associated with the variables x, y, and z in the LCS (41a) of "buy" are listed in (41b).

(41) a. ⟨buy, + V . . . ,
LAMBDA x, y, z CAUSE(x, GO(y, PATH[TO(x), z[SOURCE]]))
[BY(CAUSE(x, GO(MEANS, PATH[FROM(x), TO(AGE_1(z))]))],
. . .⟩

b. ({Agent, Goal, Agent, Source}, {Theme}, {Source, Goal})

c. (Agent, Theme, Source)

In (41b) we have listed the thematic relations associated with the variables in the order of the occurrences of the variables. Notice that according to the proposed method for associating thematic relations, and hence theta-roles with abstraction variables in LCS, each variable may be associated with several thematic relations. In fact, each variable will be associated with as many relations as there are occurrences of the variable in the LCS expression. In (41), for example, variable x is associated with four thematic relations; two Agents, one Goal, and one Source; this collection constitutes the theta-role of x.

It might be argued that the occurrences of a variable in an LCS expression are ranked in some fashion, so that exactly one of the occurrences is more "prominent" than the others. For example, the ranking might be done according to the hierarchical organization of the conceptual primitives occurring in the expression, so that the primitive naming the main event or state in the concept denoted by the expression is more prominent than the others, and so on. If such ranking of occurrences of variables is assumed, it becomes possible to distinguish a "chief" or most prominent thematic relation of a variable to the LCS expression, and the theta-role associated with the variable might well be identified with this thematic relation. Assuming this approach to the identification of theta-roles, the theta-roles associated with the LCS (41a) for "buy" are those in (41c).

We will ignore the possibility of ranking the occurrences of variables in this section, but we will return to it in section 6, on syntactic realization, where it is indeed needed.

The association of theta-roles in LCS with arguments is done by lexical insertion at D-Structure, according to the Theta-criterion (42) of Chomsky (1981).

(42) **Theta-criterion**:

(i) Each theta-role must be assigned to one and only one argument.

(ii) Each argument must be assigned one and only one theta-role.

The notion of theta-role defined in (40) is different from that found in the literature (Chomsky 1981, Jackendoff 1987). Chomsky (1981, p. 35) and Jackendoff identify the notions "thematic relation" and "theta-role", which means that a single abstraction variable may end up associated with several "theta-roles," in contradiction with the requirement of the Theta-criterion. This problem was pointed out by Jackendoff (1972, 1987).[20]

By defining theta-roles as collections of thematic relations, we obtain the right empirical possibilities, namely, that an argument of a predicate may bear several thematic relations to the predicate, as pointed out by Jackendoff, while at the same time we capture the requirement of the Theta-criterion. The Theta-criterion is a constraint on the linking of abstraction variables in LCS with arguments at D-Structure, rather than on the association of the thematic relations of a predicate with arguments. Notice that the Theta-criterion does not pay any attention to what sort of formal objects theta-roles are, or to their particular content (i.e., the labels involved). This is in line with the program of the GB framework, where syntactic principles like the Theta-criterion are not sensitive to factors of a semantic nature, like thematic relations, but rather to syntactic ones, like complement structure (Levin and Rappaport 1986). We might dispense with the derived notion *theta-role* and keep the more basic one *abstraction variable of a predicate.*

5.4. *Predicate-argument Structure*

The configurational definition (36) of thematic relations provides the mechanism needed for identifying the theta-roles associated with a predicate. However, to establish the initial syntactic configuration of the arguments, it is necessary to provide a specification of how these theta-roles are assigned. We will consider in this section the minimal distinction proposed by Williams (1981), into *external* and *internal* arguments of a predicate. A predicate may have at most one external argument, which is realized external to the maximal projection of the predicate at D-Structure. The upper bound of at most one external argument on a predicate is motivated by Williams' (1980) theory of predication. The internal arguments are realized as complements of the predicate, under its first X-bar projection. We shall adopt here Williams' convention of underlining the external argument, if any. This is shown in (43), the PAS of the verb "buy" in (41a).

(43) buy PAS: ($\underline{\{\text{Agent, Goal, Agent, Source}\}}$, {Theme}, {Source, Goal}

The principles which determine the initial linking arguments of a predicate to its abstraction variables are a topic of significant research in current linguistic theory.

In the work of Jackendoff (1972—1987), it is assumed that each lexical entry contains, among other items, the LCS representation of the entry along with its subcategorization frame. The manner of realization of each argument is stipulated by coindexation of variable positions in the LCS with categories in the subcategorization frame.[21] Notice that this approach, due to its stipulative nature, does not predict any possible vs. impossible linkings (cf. Carter 1976). Any variable (at most one) which is not coindexed must then be realized as the external argument, since this would be the only option left open for realization. A sample lexical entry from Jackendoff (1987), assuming our format for LCS expressions, is shown in (44).

(44) a. John entered the room.

b. $\begin{bmatrix} \text{enter} \\ [-\text{N} +\text{V}] \\ [\_(\text{NP}_k)] \\ \text{GO}(x_1, \text{TO}(\text{IN}(x_k))) \end{bmatrix}$

Most approaches that attempt to exploit the regularities between variables in LCS and their syntactic realization (Carter 1976, Levin and Rappaport 1987, Williams 1981) embody somehow the notion of a thematic hierarchy, which is used to state constraints on the assignment of theta-roles to the available grammatical functions. An instance of this hierarchy is (45). We shall apply the hierarchy to theta-roles also, with the understanding that when it is used on theta-roles each role is identified with the "most prominent" thematic relation it contains, relative to the LCS on which the role is defined, as discussed above.

(45) **Thematic relations hierarchy**:
Agent > Goal, Source > Theme > Path, Location, Situation

The basic observational generalization on theta-role assignment is (46) (Fillmore 1968, Levin 1985, Williams 1981). Thus, if a predicate has an argument thematically identified as Agent, it has that argument external, and its PAS, in the sense of Williams, is completely determined.

(46) If there is an Agent in a thematic grid, then it is realized as the external argument.

Beyond this generalization, however, there is only limited further understanding. When there is no Agent in a predicate's thematic grid, many other thematic relations may appear external to the predicate's maximal projection, at surface structure. Thus, the examples in (47a—c) exhibit Theme, Goal, and Source arguments in surface Subject position, respectively.

(47) a. Gorillas exist.
b. John received a letter.
c. John lost a letter.

It should be borne in mind, however, that the strings in (47) are surface strings, and that syntactic realization of arguments is defined at D-Structure (using GB theory notions). In this respect, Perlmutter (1978) has proposed that it is necessary to distinguish two classes of intransitive clauses, those whose surface Subject is an underlying Object, and those whose surface Subject is also an underlying Subject. This is the Unaccusative hypothesis, whose statement is shown in (48a) as given by Perlmutter, and in (48b) in equivalent GB terminology.

(48) **Unaccusative Hypothesis**:
a. Certain intransitive clauses have an initial 2 but no initial 1.
b. Certain intransitive clauses have an S-Structure Subject which is a D-Structure Object, and a non-thematic D-Structure Subject.

Perlmutter argues that the unaccusativity of a clause is predictable from its semantics. The predicates determining unergative (roughly, non-unaccusative) clauses correspond closely to activity predicates, which include predicates describing willed or volitional acts (work, play, walk, swim, etc.), or certain involuntary bodily acts (cough, sneeze, cry, etc.). The class of unaccusative predicates listed by Perlmutter is very large, and we shall approximate it here to the class of predicates whose highest ranked argument is a Theme (assuming the thematic hierarchy (45)). Indeed, the unaccusative predicates given by Perlmutter includes predicates whose "initial nuclear" term is a Patient, predicates of existing and happening, which may be analyzed by primitives BE_{exist} and GO_{exist}, which determine Theme relations in their first argument position, and non-voluntary emission of stimuli (shine, smell, stink, etc.) and aspectual and durative predicates (begin, stop, continue, last, remain, etc.), which may be analyzed by GO_{circ} and $STAY_{circ}$, which also determine a Theme relation at their first argument position. We advance (49) as a generalization to predict initial unaccusativity.

(49) A predicate is unaccusative if it includes a Theme theta-role and Theme is its highest-ranked role.
(Theme is defined configurationally on LCS, as in section 5.2)

Notice that the unaccusativity of a predicate is entirely predictable from its LCS representation, given the configurational definition of theta-roles.

The Unaccusative hypothesis has been very influential in explaining certain syntactic phenomena, such as the possibility of impersonal passives of predicates (Perlmutter 1978, Kornfilt 1989) and the selection of auxiliary verbs by predicates in languages that exhibit this phenomena (Burzio 1986). We will incorporate the hypothesis into our projection rules that derive predicate-argument structure, by truncating the selection of external argument at Theme, according to the hierarchy (45). It follows that a predicate whole highest relation is Theme is predicated to be unaccusative.

Our rule for derivation of PAS from LCS is thus (50).

(50) The PAS of LCS L is Q if

(i) $(\theta_1, \ldots, \theta_n)$ is the list of theta-roles derived from L by (40),

(ii) without loss of generality, θ_1 is the highest ranked theta-role in $\theta_1, \ldots, \theta_n$, according to the hierarchy (45), and

(iii) $\Theta = (\theta_1, \ldots \theta_n)$, if $\theta_1 >$ Theme;
$\Theta = (\theta_1, \ldots, \theta_n)$, otherwise ($\theta_1 \leqslant$ Theme).

Rule (50) for the derivation of predicate-argument structure assumes the

Thematic Hierarchy (45) in (50ii) to rank the theta-roles of a predicate. Notice, however, that in (50iii) a threshold is set on the Thematic Hierarchy so that a given argument will surface as an external argument only if its theta-role value exceeds the threshold. According to (49), which we advanced as a generalization that predicts initial unaccusativity of a predicate, the threshold on theta-roles is set at "theme".

The use of the Thematic Hierarchy, together with the threshold on the theta-role values, thus yields the syntactic distinction between unaccusative and unergative predicates, and makes it entirely predictable from the LCS representation of the predicate.

6. ASSIGNMENT OF GRAMMATICAL FUNCTIONS TO ARGUMENTS

The mechanism (50) for the identification of the external argument of a predicate constitutes part of the linkage we have sought to establish between LCS and the syntactic realization of arguments. Thus far we have identified the thematic grid of a predicate, in the sense of Williams (1981) and Stowell (1981). In order to complete the linkage, however, we must find out how LCS representations determine the grammatical functions at which the internal arguments of a predicate are assigned.

6.1. *Previous Approaches*

In what follows, we consider the approach in Williams (1981), which is more restrictive than the one of Jackendoff (1972, 1987). In Williams (1981) the syntactic realization of an argument is (partially) predicted from the thematic relation it bears, by "realization" rules sensitive to relation labels.[22] Williams argues that while syntactic realization rules based on thematic relations cannot be stated universally for all languages, it is possible "to avoid specification of realizations for particular lexical items in particular languages." The rules developed by Williams specify the alternative realizations of an argument from the thematic role it bears, without reference to the particular lexical item that assigns the thematic role. These rules are thus stated once and for all for a language, and individual lexical items cannot specify their own realization rules in the lexicon.

The realization rules set up by Williams are for the internal arguments of predicates; the external argument, if any, is identified by predication (e.g. in Subject position). The rules are of the form (51), where T is a thematic relation and R a sequence of maximal X-bar categories. The sequence R of categories specifies, in effect, the VP-internal grammatical function in which an argument with thematic relation T may be realized. A sequence such as (XP) refers to "the XP category dominated by VP", and (XP, YP) means "the XP category dominated by YP, in turn dominated by

VP".[23] The X-bar categories in the sequence may also include specification of internal features, such as the head of the phrase labelled by the category, a Case feature, or others.

(51) T : R

Some specific realization rules for verb arguments are listed in (52).

(52)	Agent:	(NP, PP-by)
	Theme:	(NP)
		(NP, PP-of)
		(NP, PP-about)
	Goal:	(NP, PP-to)
		(NP-2)
	Source:	(NP, PP-from)

The syntactic realization rules (52) are predicate-independent and define a one-to-many mapping from thematic relations into, effectively, grammatical functions. The interpretation associated with these rules is that an argument with thematic relation T may receive any of the realizations specified for T. Thus, a Goal argument may be realized as object of the preposition "to" or as indirect (dative) object of a verb.

The interpretation associated with (52) has a well-known problem, namely that when an argument with certain thematic relation has several alternate realizations, different verbs which assign the thematic relation may idiosyncratically accept some of the realizations, but not others. The problem is illustrated by Williams with the verbs "give" and "donate", both of which assign a Goal relation, as in (53). However, while "give" accepts the two alternate realizations (NP, PP-to) and (NP-2) specified in (52), "donate" accepts only the former ("give" is a dative verb, while "donate" is not).

(53) a. I gave a book to Bill

b. I gave Bill a book.

c. I donated a book to Bill.

d. *I donated Bill a book.

The problem is partially addressed by Williams by assuming a theory of "markedness" of realization rules. For example, of the two realizations of Goal, (NP, PP-to) might be the unmarked one. The mechanism by which the marked specifications of realization may be selected by verbs, however, is not given by Williams.

6.2. *Projection of GFs from LCS*

We propose in this section that the grammatical functions specifying the categorial realization of internal arguments at D-Structure are determined by lexical-conceptual representations. The exact mechanism will be reviewed in detail, and is similar to how thematic relations were determined in Section 5 by LCS. That is, the syntactic category of the realization of an argument is determined configurationally by the LCS in which the argument is to be substituted.

One technical detail is important here: Recall from section 5 that abstraction variables in LCS may have several occurrences in an expression. It is this fact which determines that a given argument may bear several thematic relations to the LCS. For example, the variable x in the LCS (40) of "buy" has four occurrences. In order to let the configuration surrounding the occurrences of this variable uniquely determine the syntactic realization of the argument, it is necessary to introduce the notion "main occurrence of variable x" in the LCS. This notion was informally discussed in section 5, where it was used to establish the "most prominent" thematic relation contained in a theta-role, relative to a given LCS.[24]

We will let, as an empirical generalization on the set of LCSs we have surveyed, the pre-order enumeration (Knuth 1973) of the occurrences of a variable define the ranking of the occurrences. The pre-order is defined on the syntactic structure (tree) which shows the structure of the expression. This structure may be defined in a number of equivalent ways, based on the LCS syntax of section 2. We will omit from the structure tree the LAMBDA abstraction operator and the occurrences of the variables that immediately follow it. The ranking is defined in (54).

(54) Let x_1 and x_2 be variable occurrences in LCS L.
Then $x_1 < x_2$ if x_1 occurs before x_2 in the pre-order enumeration of the nodes of L.
The preorder enumeration of the nodes of L is defined by the following procedure:

(i) visit the root node;
(ii) for each $i = 1, \ldots, n$, where n is the number of children of the current node, visit the subtree dominated by child i in pre-order.

This ranking of occurrences of variables implies, among other things, that occurrences of a variable inside the head of an LCS expression are ranked higher than occurrences on the same variable inside conceptual modifiers of the head. For example, in the LCS (40) for "buy", repeated for convenience as (55), the occurrence of x as first argument of CAUSE in the head of the expression is ranked higher than the others. Thus, this occurrence determines the syntactic realization of x, and also determines the main thematic relation of x, namely Agent.

(55) ⟨buy, +V . . . ,
LAMBDA x, y, z CAUSE(x, GO(y, PATH[TO(x), z[SOURCE]]))
[BY(CAUSE(x, GO(MEANS, PATH[FROM(x), TO(ARG_1(z))]))],
. . .⟩

The rule for assigning a syntactic realization to abstraction variables in LCS is now given by rule (56).

(56) **Grammatical function assignment rule**:
Abstraction variable x in LCS L has syntactic realization [V′, XP] if

(i) P is the conceptual primitive immediately containing x,

(ii) the argument position of x in P is i,

(iii) T is the required conceptual type of x in L, and

(iv) XP = S(P, i, T), where S is the map (57) below.

(57)

Primitive	**Position**	**Type**	**Category**
BE	1	thing	NP
		situation	CP
	2	—	PP-place
GOext	1	—	NP
ORIENT	1	—	NP
GO	1	—	NP
	2	—	PP-path
LET, CAUSE	1	—	NP
	2	—	CP
AT	1	thing	NP
		property	AP
IN	1	—	NP
TO	1	thing	NP
		place	PP-place
PATH	mod.	goal	PP-goal
		source	PP-source

The right-hand column of the table in (56) lists the syntactic category of the argument. Thus, if XP is the category listed, the grammatical function in which the argument is realized is [V′, XP]. The category of prepositional phrases has been augmented with notations such as PP-place (locative PPs) and PP-path (path PPs), to partition the class and state the kind of selectivities that verbs exhibit regarding PP complements.

The grammatical functions that rule (56) predicts for the realization of the two internal arguments of "buy", according to LCS (55), are those listed in (58). In the terminology of Marantz (1984), y is identified with the direct argument, and z with an indirect argument whose theta-role is to be assigned compositionally by the verb and a preposition of type source (i.e., "from"). Variable x is linked to the external argument, since it is associated with the Agent theta-role.

(58) y : [V′, NP] z : [V′, PP-source]

An important dependence of the GF projection rule (56) on LCS expressions is revealed by the above example. Unlike the projection rule (36) for thematic relations, rule (56) is sensitive to the specified conceptual type of the main LCS occurrence of the variable to be realized. This sensitivity on conceptual type permits lexical-conceptual structure to encode syntactically subtle distinctions in realization of arguments. This was illustrated in section 4.3.

Due to the sensitivity of rule (56) to the conceptual type of variables, the rule should be taken with a grain of salt; if the LCS expressions that encode the meaning of lexical items are not well motivated, conceptual types might be used to directly encode the grammatical functions of arguments, at the level of LCS, thus trivializing the posited correspondence between LCS and syntax.

7. LEXICAL ALTERNATIONS AND THE UNACCUSATIVE HYPOTHESIS

7.1. *Lexical Alternations*

A theory of lexical organization is of interest to the extent that it allows to capture, not only the meanings of lexical items and the regularities in their syntactic behavior, but also the alternations in complement structure that the items exhibit.

These alternations, which we will call *lexical alternations*, are due to several sources: (i) the possibility of having alternate syntactic realizations of arguments, (ii) the optionality or obligatoriness of the arguments selected, and (iii) the possibility of "adding" an extra argument to an argument structure. The most common example of lexical alternation due to (i) is given by the dative verbs, which select internal Theme and Goal arguments, and may realize either argument as the direct one.[25] This is shown in (59a). Another instance of alternation due to (i) is provided by the class of "spray/load" verbs (Rappaport and Levin 1986). Like the dative verbs, these verbs select two internal arguments, a Theme and a Location or Goal argument, and have two alternate syntactic realizations of them, as in (59b). Yet another instance of this kind of alternation is the conative alternation, exhibited in (59c) (Guerssel et al. 1985).

(59) a. I sold a book to John.
I sold John a book.

b. John sprayed paint on the wall.
John sprayed the wall with paint.

c. The police shot John.
The police shot at/toward John.

Besides the alternations shown in (59), verb complement structures contrast in whether they require syntactic realization of their arguments. While a verb like "eat" allows optional realization of its complement (60a), the verb "put" requires realization of both its complements (60b). Furthermore, as Levin and Rappaport (1986) point out, the optionality of realization of arguments depends on the pattern of theta-role assignment, for verbs with more than one pattern. This is illustrated with the verb "sell", which has the two assignment patterns (60c-d) given by the dative alternation.

(60) a. I ate the food.
I ate.

b. Mary put the keys on the table.
*Mary put the keys.
*Mary put on the table.

c. John sold his motorcycle to David.
John sold his motorcycle.
*John sold to David.

d. John sold David his motorcycle.
*John sold David.

The last form of alternation mentioned above is the possibility of adding an argument to the predicate-argument structure of a verb. This we assume is the case in the causative/inchoative alternation shown in (61).

(61) a. The window broke.

b. John broke the window.

The representations of predicate-argument structure used thus far list the theta-roles associated with a predicate, and are slightly augmented with diacritics and other notational conventions, to indicate the manner of realization of arguments (cf. the notations for marking the external and direct arguments). In these representations, it is implicitly assumed that all arguments must be syntactically realized, in order to satisfy the Theta-criterion and Projection principle. It has frequently been observed, however, that many verbs allow optional realization of some of their arguments.

The terminology we have used to describe the sources of alternation, including the notions of "alternate syntactic realization", "optionality of arguments", and the possibility of "adding an extra argument", needs a precise interpretation in our framework. This terminology is taken from Levin and Rappaport (1986). In part, the source of the alternations was discussed in section 4. In future work we shall give this terminology a precise interpretation; we shall also investigate the relationship between Lexical Conceptual Structures and Perlmutter's (1978) Unaccusative Hypothesis (integrated into the GB framework mainly in Burzio (1986)) and address the question of how our projection rules proposed in the present report can deal with the Unaccusative Hypothesis and with alternations in the complement structure of predicates.

8. CONCLUSION

The present chapter has proposed Lexical Conceptual Structure (LCS) as a valid level of linguistic representation and developed an explicit syntactic account of it that allowed us to explore the LCS-syntax correspondence. Our formulation of this correspondence is a three state process which in its first stage defines thematic relations and, derivatively, theta-roles as configurational relations on conceptual representations. In the second stage, predicate-argument structure (PAS) is defined for each lexical item by means of a projection rule, based on the thematic hierarchy we have postulated, and which may select an external argument. At this stage certain properties of a predicate such as its accusativity or unaccusativity surface, depending on the highest ranked role of the predicate. The third and final stage assigns the grammatical function in which each internal argument of the lexical item is realized; this in turn determines the full range of the syntactic behavior of the item.

In the approach to conceptual structure we have developed LCS is a formal syntactic system with its own set of formation and relation rules. This approach does not exclude, in principle, a denotational account of the semantics of conceptual primitives and expressions. As pointed out in section 4, we believe that the denotational approach has much to contribute to the study of natural language semantics and that it serves to put in perspective the approaches based on meaning postulates, lexical decomposition, and several other approaches to the description of the various entailment relations between lexical items.

The present study will benefit from a more detailed and carefully documented series of case studies of lexical items and their lexical properties and alternations than was possible to carry out here. A cross-linguistic study of variation and applicability of the approach also seems desirable and is left for future investigation.

NOTES

[1] In this work, we shall alternately speak of "using LCS to capture syntactic properties of lexical items", and of "deriving the syntactic properties of a lexical item from its LCS representation". Given that LCS is, a priori, a more basic property of a lexical item than any of its syntactic features, it seems more appropriate to view the LCS-syntax correspondence as a directed relation; in such way syntactic properties are determined by lexical-conceptual representations (cf. Levin and Rappaport 1986). However, lexical-conceptual representations of lexical items are not generally available, and it is instead the objective of this work to arrive at a framework for such representations and at the representations for a representative sample of English verbs. In arriving at those representations, semantic criteria play the dominant role, but in unclear cases we will also let syntactic tests help determine the correct conceptual representation for a given item. Thus, in these situations it is more accurate to speak of LCS as a representation encoding syntactic and semantic properties of the item.

[2] The standard terminology in GB theory for referring to the identification of arguments with argument positions of predicates refers to assignment of thematic relations; thematic relations are said to be "assigned" to the arguments of the predicate (Chomsky 1981).

[3] Throughout we assume that grammatical functions are configurationally defined on syntactic representations (cf. Chomsky 1965).

[4] An important element missing from Jackendoff's theory is a mechanism for representing relations of quantification in conceptual expressions. However, given that the system already has the notion of (lexical) "variable" in it, the major element missing is the introduction of a class of operators, including quantifiers, which may be related to variables by a predicate-calculus notion of "binding". We will tackle this in section 2.7.

[5] The pitfall that this approach avoids is arbitrarily providing criteria for determining when all aspects of a predicate's meaning are captured in an LCS formula. The approach thus contrasts sharply with that advocated by Schank (1973), who claims that conceptual decompositions of a word should capture all elements of the meaning of the word.

[6] The second kind of pitfall that the approach avoids is to provide an unambiguous or prescriptive characterization of the conditions under which a given concept qualifies as a "CUT", a "sharp edge", etc.

[7] The use of thematic relations to identify the semantic relations of arguments to predicates stems from the work of Gruber (1965) on lexical relations and Fillmore (1968) on Case grammar. It poses a number of difficult problems, of which the two most often pointed out are difficulty of characterization and justification of the set of thematic relations that exist, and the justification of particular thematic relation assignments to arguments. Comrie (1981) takes an interesting position concerning thematic relations, indicating that there is "not so much a set of discrete semantic relations, but rather a continuum, the [thematic relation] labels representing different points along this continuum" (p. 53).

[8] Alternatively, we may view the rule that relates LCS to PAS as a lexical redundancy rule which captures the regularities between the two representations. This rule thus reduces the amount of independent information in the lexicon and contributes to the evaluation metric for the proposed model of lexical organization (cf. Jackendoff 1975).

[9] Other approaches to natural language semantics try to formalize this ontology by restricting it to an artificially delimited collection of kinds of entities, such as "physical objects", "events", "times", "propositions", and "possible worlds".

[10] Clauses in natural language always involve a predication on a number of arguments; within the GB framework, this is motivated by Williams' theory of predication.

[11] It may be that the conceptual "primitives" postulated are analyzable or related to yet other kinds of cognitive representations. Hence, the elements assumed are primitive only relative to the level of lexical decompositions, and need not be primitive below this level.

[12] Jackendoff's theory of conceptual structure does not have a logical category of predicates. The distinction between referential and predicative uses of symbols is marked by the TOKEN/TYPE alternation.
[13] This fact lends validity to Hobbs' (1985) observation that most approaches to semantics of natural language use a vocabulary which is isomorphic to the one we use to talk about meanings.
[14] The work we have in mind here is that reported in the lexical semantics literature. Schank's (1973) theory of conceptual dependency identifies a vocabulary of eleven primitive actions. To our knowledge, Schank's is the only fully explicit and concrete proposal made regarding the set of primitives in the human CS system; for an evaluation of Schank's work see (Sproat 1985).
[15] This is probably a hangover from syntactic representations, in which part-of-speech ambiguity of terminal symbols is the rule rather than the exception.
[16] As noted in the introduction of the paper, we view theta-role assignment as the identification of arguments of a predicate with variables in the LCS of the predicate. This formal view is to be contrasted with the more standard one, which assumes assignment of theta-role labels to the arguments.
[17] The exact format of the representation of the information in the lexicon is open to a variety of options, which will not be addressed here. One possibility, for example, is to assume "fully specified lexical entries", in which all the semantic, syntactic, etc. information of an entry is fully specified, ignoring regularities that may exist between the different sorts of information. Such regularities would be captured by redundancy rules of the sort proposed by Jackendoff (1975), thus reducing the amount of independent information in the lexicon and contributing to the evaluation metric of the lexicon in question.
[18] As observed by Jackendoff (1983), "the place-function IN requires its reference object to be regarded as a bounded area or volume" (p. 162).
[19] This position on thematic relations is present in Gruber (1965) and Jackendoff (1972, 1987). However, the position is not exploited in most other work on lexical semantics (Fillmore 1968, Levin and Rappaport 1986, Hale and Keyser 1987) or syntactic theory (Chomsky 1981, Bresnan 1982, Gazdar et al. 1985); thematic relations are simply taken as diacritics or features specified in the lexical entries of predicates.
[20] The statement of the theta-criterion in Jackendoff (1987) is ambiguous, since it refers to "argument positions in CS." This term might be interpreted as referring to abstraction variables or to occurrences of the same variables; we get the correct interpretation only in the first case.
[21] Jackendoff (1987) mentions the possibility of general linking principles but does not actually formulate any of them; the problem is left open.
[22] Williams identifies thematic relations with theta-roles.
[23] It is to be noted that, as in (Chomsky 1965), Williams does not distinguish between the first and maximal projections of a category in the specification of GFs. Thus, the notation for the Direct Object is [VP, NP], rather than [V′, NP] as it should be in a two-level X-bar system.
[24] Recall that a theta-role is formally defined as a collection of thematic relations.
[25] The analysis of the D-Structure representations involved in the alternation are not settled, however (cf. Chomsky 1981, Levin and Rappaport 1986, Stowell 1981).

REFERENCES

Bierwisch, Manfred: 1967, 'Some Semantic Universals of German Adjectivals', *Foundations of Language* **33**, 1–36.
Boguraev, Bran and Ted Briscoe: 1987, 'Large Lexicons for Natural Language Processing:

Using the Grammar Coding System of LDOCE', *Computational Linguistics* **13**(3–4), 203–218.

Bresnan, Joan (ed.): 1982, *The Mental Representation of Grammatical Relations*, MIT Press, Cambridge, Massachusetts.

Burzio, Luigi: 1986, *Italian Syntax: A Government-Binding Approach*, Reidel, Dordrecht.

Carter, Richard: 1976, 'Some Linking Regularities', manuscript, Université de Paris, Vincennes; also in 'On Linking: Papers by Richard Carter', Beth Levin and Carol Tenny (eds.), *Lexicon Project Working Papers No. 25*, 1988. Center for Cognitive Science, MIT, Cambridge, Massachusetts.

Chomsky, Noam: 1965, *Aspects of the Theory of Syntax*, MIT Press, Cambridge, Massachusetts.

Chomsky, Noam: 1981, *Lectures on Government and Binding*, Foris, Dordrecht.

Chomsky, Noam: 1982, *Some Concepts and Consequences of the Theory of Government and Binding*, Linguistic Inquiry Monograph 6, MIT Press, Cambridge, Massachusetts.

Church, Alonzo: 1941, *The Calculi of Lambda Conversion*, Annals of Mathematical Studies 6, Princeton University Press, Princeton.

Comrie, Bernard: 1981, *Language Universals and Linguistic Typology*, University of Chicago Press, Chicago.

Correa, Nelson: 1988, *Syntactic Analysis of English with Respect to Government-Binding Grammar*, Syracuse University Ph.D. dissertation, Syracuse, New York.

Fillmore, Charles: 1968, 'The Case for Case', in E. Bach and T. Harms (eds.), *Universals in Linguistic Theory*, Holt, Rinehart, and Winston, New York, pp. 1–88.

Fillmore, Charles: 1977, 'The Case for Case Reopened', in P. Cole and J. M. Sadock (eds.), *Grammatical Relations, Syntax and Semantics, Vol. 8*, Academic Press, New York, pp. 59–81.

Gazdar, Gerald, Ewan Klein, Geoffrey K. Pullum and Ivan Sag: 1985, *Generalized Phrase Structure Grammar*, Harvard University Press, Cambridge, Massachusetts.

Gruber, Jeffrey: 1965, *Studies in Lexical Relations*, MIT Ph.D. dissertation, Cambridge, Massachusetts; also in *Lexical Structures in Syntax and Semantics*, 1976, North-Holland, Amsterdam.

Guerssel, Mohamed, Kenneth Hale, Mary Laughren and Josie W. Eagle: 1985, 'A Cross-linguistic Study of Transitivity Alternations', Parassesion on Causatives and Agentivity, Chicago Linguistics Society, 21st Regional Meeting, Chicago.

Hale, Kenneth and Samuel J. Keyser: 1987, 'A View From the Middle', *Lexicon Project Working Papers 10*, Center for Cognitive Science, MIT, Cambridge, Massachusetts.

Hale, Kenneth, and Mary Laughren: 1983, 'The Structure of Verbal Entries', Walpiri Lexicon Project, MIT, Cambridge, Massachusetts.

Hobbs, Jerry: 1985, 'Ontological Promiscuity', *Proceedings of the 23rd Annual Meeting of the Association for Computational Linguistics*, Chicago, Illinois, pp. 61–69.

Hobbs, Jerry, William Croft, Todd Davis, Douglas Edwards and Kenneth Laws: 1987, 'Commonsense Metaphysics and Lexical Semantics', *Computational Linguistics* **13**(3–4), 241–250.

Jackendoff, Ray: 1972, *Semantic Interpretation in Generative Grammar*, MIT Press, Cambridge, Massachusetts.

Jackendoff, Ray: 1975, 'Morphological and Semantic Regularities in the Lexicon', *Language* **51**, 639–671.

Jackendoff, Ray: 1976, 'Toward an Explanatory Semantic Representation', *Linguistic Inquiry* **7**, 89–150.

Jackendoff, Ray: 1983, *Semantics and Cognition*, MIT Press, Cambridge, Massachusetts.

Jackendoff, Ray: 1987, 'The Status of Thematic Relations in Linguistic Theory', *Linguistic Inquiry* **18**(3), 369–411.

Knuth, Donald: 1973, *Fundamental Algorithms*, The Art of Computer Programming, Vol. 1, Addison-Wesley, New York.

Kornfilt, Jaklin: 1989, 'Turkish Psych-Verbs and the Unaccusative Hypothesis', Talk given at I.B.M. Thomas J. Watson Research Center, February 20, 1989, Yorktown Heights, New York.

Kornfilt, Jaklin and Nelson Correa: 1990, *Deriving the Structure of Lexical Entries from Conceptual Structure*, Syracuse University and Universidad de Los Andes; Syracuse University, CASE Center, Technical Report No. 9021.

Koster, Jan: 1987, *Domains and Dynasties: The Radical Autonomy of Syntax*, Foris, Dordrecht.

Levin, Beth: 1985, 'Lexical Semantics in Review: An Introduction', in Beth Levin (ed.), *Lexicon Project Working Papers 1*, Center for Cognitive Science, MIT, Cambridge, Massachusetts, pp. 1—62.

Levin, Beth and Malka Rappaport: 1986, 'The Formation of Adjectival Passives', *Linguistic Inquiry* **17**(4), 623—661.

Marantz, Alec: 1984, *On the Nature of Grammatical Relations*, MIT Press, Cambridge, Massachusetts.

Olson, Davis and Ellen Bialystok: 1982, 'Spatial Cognition: The Mental Representation of Objects and Form', in B. de Gelder (ed.) *Knowledge and Representation*, Routledge & Kegan Paul, London, pp. 121—138.

Perlmutter, David: 1978, 'Impersonal Passives and the Unaccusative Hypothesis', *Proceedings of the Berkeley Linguistics Society 4*, University of California, Berkeley, pp. 157—189.

Pylyshyn, Zenon: 1984, *Computation and Cognition: Toward a Foundation for Cognitive Science*, MIT Press, Cambridge, Massachusetts.

Rappaport, Malka and Beth Levin: 1986, 'A Case Study in Lexical Analysis: The Locative Alternation', manuscript, MIT, Cambridge, Massachusetts.

Rappaport, Malka and Beth Levin: 1988, 'What to Do with Theta-Roles', in Wendy Wilkins (ed.), *Syntax and Semantics: Thematic Relations Vol. 21*, Academic Press, New York, pp. 7—36.

Schank, Roger: 1973, 'Identification of Conceptualizations Underlying Natural Language', in R. Schank and E. Colby (eds.) *Computer Models of Thought and Language*, Freeman Press, San Francisco, pp. 187—247.

Sproat, Richard: 1985, 'Review of Schank's 'Identification of Conceptualizations Underlying Natural Language', in Beth Levin (ed.), *Lexical Semantics in Review: An Introduction*, MIT, Cambridge, Massachusetts, pp. 115—124.

Stowell, Timothy: 1981, *Origins of Phrase Structure*, MIT Ph.D. dissertation, Cambridge, Massachusetts.

Talmy, Leonard: 1985, 'Lexicalization Patterns: Semantic Structure in Lexical Forms', in *Language Typology and Syntactic Description 3*, Cambridge University Press, Cambridge.

Travis, Lisa: 1984, *Parameters and Effects of Word Order Variation*, MIT Ph.D. dissertation, Cambridge, Massachusetts.

Wilks, Yorick: 1977, 'Methodological Questions About Artificial Intelligence: Approaches to Understanding Natural Language', *Journal of Pragmatics* **1**, 69—84.

Williams, Edwin: 1980, 'Predication', *Linguistic Inquiry* **11**(2), 203—238.

Williams, Edwin: 1981, 'Argument Structure and Morphology', *The Linguistic Review* **1**, 81—114.

Williams, Edwin: 1986, 'A Reassignment of the Functions of LF', *Linguistic Inquiry* **17**(2), 265—299.

Zubizarreta, Maria Luisa: 1987, *Levels of Representation in the Lexicon and in the Syntax*, Foris, Dordrecht.