Introduction

Teun Hoekstra

Harry van der Hulst

Michael Moortgat

1. The Origin of Lexicalism

1.1. *The Lexicalist Hypothesis*. The lexicalist research program is the outgrowth of Chomsky's *Remarks on Nominalization*.[1] The full impact of this article can only be appreciated in the light of the subsequent developments of which it was the cause. Compared with the far-reaching consequences of the lexicalist position, the seminal paper itself strikes one as rather dated: it is heavily involved in the Generative Semantics polemic, which was at its height in 1970; as a result of its programmatic nature, the proposals on the organization of the lexicon remain sketchy. This discrepancy between its actual contents and the role it has played in the development of generative theory ranks *Remarks* among the revolution-making rather than revolutionary texts: the significance of *Remarks* lies less in what it says itself than in what it caused others to say.[2]

Remarks introduces a major problem shift in the generative tradition. This tradition, up till then, might be characterized by its transformational holism. Since transformations were considered to be the only means to express relatedness, virtually the entire computational burden of relating meaning to surface form had been borne by the transformational component.[3] In *Remarks*, the holistic approach is abandoned. It is suggested, for the first time, that the expressive power of a generative grammar is not concentrated primarily in its transformations.[4] Rather, the apparent complexity and heterogeneity of the phenomena can be factored into the contribution

[1]Chomsky (1970). *Remarks* is the written version of a series of lectures given at MIT in the fall of 1967.

[2]Kuhn (1957:135) [3]Bresnan (1978:14) [4]Jackendoff (1977:xi)

GLOT-LEXICAL GRAMMAR (1980), 1-48.

of interacting systems, each based on simple principles.[5] The methodological shift from holism to modularity bore witness to the flexibility of the generative research program: it required the abandonment of a hard core assumption: the dominant role of transformations in relating the underlying level of meaning to the observable level of sound. As a move in the current controversy, *Remarks* was effective indeed: it dealt the deathblow to Generative Semantics, whose consistent application of the *Aspects* methodology was felt as a perversion of the original aims, but could not be countered on any principled grounds within the Standard framework itself.

One has to go back to the broad organization of generative theory as originally laid down in *The Logical Structure of Linguistic Theory (LSLT)*[6] to understand the motivation behind Chomsky's radical move. The base component of the theory outlined in LSLT is a context sensitive phrase structure grammar which accomplishes two functions: the generation of hierarchical constituent structure and the selection of lexical items. In order to avoid repetition of identical information with respect to the selectional requirements of related expressions, Chomsky proposes to enrich the theory with an abstract level of syntactic description, deep structure. Within such a syntactic two-level approach, related expressions could be given one representation at the underlying level where selection is handled; their differences could be attributed to the second, phonologically relevant syntactic level, surface structure. Transformations are introduced as the mediating device between the hidden and the observable syntactic levels. The transformational level, then, enables one to avoid redundant repetition of identical information indicative of a failure to capture significant generalizations.

Distributional similarities are shared by sentences (e.g. active vs. passive) as well as by words (e.g. nominalizations). Transformational holism, therefore, effaced the traditional distinction between morphology and syntax: within the Standard Theory, derivation and compounding[7]

[5]Chomsky (1977b, 5) [6]Chomsky (1955)

[7]Cf. Lees (1960) and Botha (1968). For the Standard treatment of inflection see Bierwisch (1967), Wurzel (1970), Kiefer (1970, 1973). In most of these works inflectional rules spell out syntactic features. These features have different sources but since some are introduced by transformations the spell out rules are ordered after the syntax in the so-called readjustment component. Cf. fn. 62.

were handled by transformational rules. On the other hand, since non-transformational interpretive mechanisms were lacking, anaphoric phenomena[8] and control of subjectless complements[9] were treated transformationally too. Because *all* distributional similarities in surface structure had to be captured by deriving them from common deep structures, no adequate constraints on the abstractness of deep structure could be formulated. Due to the necessary remoteness of deep structure and to the variety of relations that had to be expressed, the transformational component soon appeared to be unlimited in its power. Investigation into the mathematical properties of transformational systems[10] contributed to the general pessimism concerning the explanatory value of a powerful transformational system. The quest for constraining the power of the transformational component as far as possible could begin.

Restricting the power of the transformational component can in principle be effected in either of two ways: one can enrich the other components in such a way that the new division of labour gives the transformational component less work to do, or one can impose limits on the expressive power of the transformational rules themselves. In *Remarks*, Chomsky concentrates on the first approach, i.e., on the demarcation problem. He presents the proper balance between the various components of the grammar as entirely an empirical issue.[11] Yet, subsequent practice has shown a strong tendency to reduce the content of the transformational component in favour of enrichment of the other components. As illustrated

[8]For a transformational treatment of pronominalization cf. Lees & Klima (1963) and Langacker (1969). An interpretive theory of anaphora was introduced by Jackendoff (1969). It constitutes an important step towards eliminating the monopoly position of the transformation. Moreover, it made clear that surface structure was much more crucial to semantic interpretation. This was a first step leading to surface structure interpretation and hence to the reduction of the importance of the level of deep structure in general.

[9]Cf. Rosenbaum's (1967) rule of Equi-NP-Deletion.

[10]Kimball (1967), Ginsburg & Partee (1969), Peters & Ritchie (1971, 1973). These investigations established the equivalence of *Aspects*-type transformational grammars and Turing automata. For an illuminating discussion of linguistic consequences, especially on learnability, see Levelt (1976).

[11]Chomsky (1970:194)

above, the imperialistic attitude of early transformational work resulted in the annexation of word-formation processes and semantic-interpretive phenomena. Chomsky's own work after *Remarks* deals with unburdening the transformational component by appealing to interpretive rules and principles: this is the effect of the identification, within trace theory, of empty nodes as bound anaphors.

In *Remarks*, Chomsky suggests to attack the content of the transformational component from another angle: the base[12] and the lexicon. He proposes to establish the relation between verbs and the corresponding derived nominals by means of lexical rules relating entries, rather than by transformations relating the full phrase markers in which these entries occur. The empirical arguments adduced in favour of a lexical treatment center on the limited productivity, and the occurrence of semantic non-compositionality in nominalizations; transformations are supposed to handle the fully regular processes, not processes governed by lexical exceptions. We see here the emergence of a set of empirical criteria used to decide on lexical or transformational treatment. *Remarks* is only concerned with nominalizations; it is clear, however, that they are presented as a test case for the validity of the distinction, and that the methodology introduced here was supposed to apply, at least, to derivational processes in general.[13] Together with the interpretive theory of anaphora where pronouns are not transformationally derived but inserted in deep structure instead, this position resulted in the Strong Lexicalist Hypothesis, which states that no lexical material may be introduced by transformations.[14]

[12]Base-generating NP's with a derived nominal as head required a richer theory of phrase structure rules than that assumed in *Aspects*. Jackendoff (1977) is the most detailed account of an X-bar base along the lines of Chomsky's suggestions in *Remarks*.

[13]Others have extended the lexicalist hypothesis to other or all word formation processes. E.g. Wasow and Roeper (1972), Schachter (1976), Jackendoff (1972, 1975) and Halle (1973). Cf. fn. 14 and 19.

[14]This formulation is due to Jackendoff (1972) where it is referred to as Extended Lexicalist Hypothesis. This position still allows for base generation of bound inflexional affixes that are transformationally attached to verbal elements. A stronger position, not allowing for any free generation of affixes, is taken by Brame (1979) Lapointe (this volume) and Bresnan (1979). In the literature one can find a variety of terms referring to this stronger position (e.g. strong, extended, generalized lexicalism) or to crucial principles involved (e.g. lexical integrity, spel-

The introduction of lexical rules distinct from transformations did not appear out of the blue: it was the logical outcome of Chomsky's introduction of the lexicon, as sketched in *Aspects*.[15] With respect to the lexicon, the *Aspects* theory differs crucially from the theory outlined in *LSLT*. In *LSLT* the tension between hierarchical structure and context-sensitive selectional properties of lexical items was attributed to a context-sensitive set of phrase structure rules; in *Aspects*, the phrase structure component is considered to be a context-free rewriting system accounting for the hierarchical constituent structure. The context-sensitivity of selectional requirements is transferred in its totality to the lexicon: lexical items are associated with a subcategorization frame which functions as the structural description of the insertion transformation. It has been stressed by Heny that the introduction of a lexicon consisting of items associated with a transformational subcategorization condition, and lexical rules relating entries, "actually undercut the basic arguments *for* a transformational component as these had originally been proposed".[16] In *Remarks*, however, "the lexical rules are presented as an *addition* to a theory which as it were by nature includes a set of distinct transformational rules".[17] Instead of *assuming* the existence of a transformational level, the question might have been raised as to which processes traditionally handled by transformations did *not* fall within the scope of the enriched lexicon.

As a result of Chomsky's sketchy proposals on the organization of the lexicon, the revolutionary potential of *Remarks* remained untapped for a long time. The research in the wake of *Remarks* did not attempt to exploit the full resources of lexical rules by fleshing out in detail the organization of the lexicon; instead, it concentrated on the aspect of lexical rules already elucidated in *Remarks*, i.e. on morphology proper.

(rule 14 cont'd)
ling prohibition). In this paper we single out *Strong Lexicalism* and the *Principle of Lexical Integrity* to refer to this position.

[15]See Chomsky (1965), Chapter 4, par. 2 *The Structure of the Lexicon*, and especially the section on derivational processes, where it is suggested that 'it may be necessary to extend the theory of the lexicon to permit some "internal computation"' (p. 187).

[16]Heny (1979b:318)

[17]Heny (1979b:319)

1.2. *The Rediscovery of Morphology*.[18] A paper which played a dominant role in the establishment of a morphological component within the generative framework was Halle's *Prolegomena to a Theory of Word Formation*. His proposals were elaborated in the work of Siegel and Aronoff. The suggestion of *Remarks* to introduce lexical rules was independently taken up by Jackendoff.[19] In this section we will shortly characterize both lines of work.

In general, rules dealing with word structure have two functions. First, they account for the formation of new words and secondly, they account for the internal structure of existing words. In short, then, these rules make explicit the notion 'possible word of language L'. We could characterize these two functions as *dynamic* and *static*, respectively. Jackendoff's view of the lexicon focuses largely on the static function of lexical rules. He assumes that all words are stored in the lexicon, fully specified, and that lexical rules function to analyze these words in order to compute their amount of independent information. This view came to be known as the full entry theory of the lexicon. The fact that some lexical rules can also be used to create new words is considered as a secondary phenomenon. Once a new word is created it directly falls within the scope of the redundancy function of lexical rules. Siegel's account of word structure, on the other hand, is purely dynamic.[20] All non-atomic words are generated each time anew. Such an approach conforms to the historicist approach characteristic of standard generative grammar: diachronic developments are recapitulated in the synchronic grammar. A positive result of Siegel's view on the lexicon is that much attention is paid to the restrictions that affixes can impose on their base and, as a special case, to co-occurence restrictions that hold between affixes.[21]

[18]Cf. Lipka (1975)

[19]Cf. Halle (1973, presented in 1972), Siegel (1974, 1978), Aronoff (1976), Jackendoff (1975).

[20]Aronoff (1976, published version of a 1974 dissertation) takes a position which is intermediate between Jackendoff's and Siegel's.

[21]Co-occurence restrictions can be handled in a number of ways (for a survey cf. Schultink (1975)). Siegel has claimed that all such restrictions can be explained in the following manner. She observes that the English affixes can be classified in two groups. Independent evidence for this classification comes from the behaviour of phonological rules with respect to derived words. All co-occurence restrictions follow at once if we assume that the rules that introduce these classes of affixes

In Aronoff's and Siegel's work it has been demonstrated that restrictions can hold at any level of representation of lexical items, provided that the properties which are referred to are available in the adjacent cycle.[22] Examples are given of *phonological* restrictions (i.e. sensitivity to certain segmental or accentual properties of the base) and *morphological* restrictions (several suffixes require a [+latinate] base; affixes may also be sensitive to specific other affixes, negatively or positively). Other restrictions involve the categorial or subcategorial value of the base or selectional and more general semantic properties.[23]

(note 21 cont'd)
are extrinsically ordered with respect to each other. The behaviour of the phonological rules can now be explained in two ways. As for the word stress rules, Siegel assumes that these are ordered between the two classes of word formation rules. This explains why certain affixes, e.g. those introduced after the stress rules, never affect the stress pattern of their base. With respect to other phonological rules it is necessary to assume that the class of WFR's which is ordered last introduces affixes associated with a strong grammatical boundary ('#'), whereas the first class does not have this property; their application results in a weak '+' boundary, which has no phonological impact. It should be noted, however, that the stress facts would also fall out correctly if we ordered the word stress rule after all WFR's, given the presence of the independently motivated strong boundary. It seems then that Siegel's hypothesis suffers from a certain redundancy (cf. Booij (1977)). Booij also pointed out that the ordering restrictions could be accounted for in terms of a general condition on the application of WFR's, viz. the condition that no stronger boundary affix may be peripheral to a weaker boundary affix. In the light of attested counterexamples to the ordering hypothesis (Aronoff 1976) it may be necessary to interpret this condition in a relative way. Obviously, such a relative interpretation of Siegel's version of the ordering hypothesis is impossible. Allen (1978) argues on independent grounds that the ordering facts must be accounted for in terms of a condition on boundary combinations. It will be clear that within a full entry theory of the lexicon such an explanation is the only possible one.

[22]Siegel (1978) and Allen (1978) have proposed a *meta*-restriction on the application of WFR's, known as the *Adjacency Condition*. This condition states that "no WFR can involve X and Y, unless Y is uniquely contained in the cycle adjacent to X". The condition is based on the assumption that each application of a WFR creates a new cyclic domain. A WFR, then, can only refer to material introduced by the previous rule (or to properties of the underived base, if there is no previous rule).

[23]Examples of these restrictions can be found in Siegel (1974, 1978), Aronoff (1976), Allen (1978).

It is a typical feature of work in generative morphology that the non-morphological aspects of lexical rules are viewed in function of the restrictions they impose on affixation. Instead of exploring the full potential of the newly introduced device of lexical rules the work in generative morphology is heavily biased toward the morphophonological component of these rules. Consequently, far less attention was paid to the fact that the other levels of representation not only impose restrictions on word formation rules but also are subject to changes themselves. In this respect, the work of Jackendoff appeared to be more fruitful for the development of lexicalism.

1.3. *Strong Lexicalism*. It was not until the mid-seventies that one became aware of the full impact of the methodological changes introduced in *Remarks*. The work in generative morphology had, in the mean time, firmly established word formation rules as a part of grammatical theory; it was generally recognized, by 1975, that the lexicon had to be more than a stored list of basic irregularities: the regular aspects of word structure required an amount of internal computation within the lexicon.[24]

Jackendoff's[25] study of morphological and semantic regularities in the lexicon constitutes an important step in the gradual process that resulted in a full awareness of the implications of the *Remarks* position. Jackendoff is no longer exclusively concerned with the morphophonological aspect of lexical rules; he also takes into account the changes that can be effected on the level of subcategorization and semantic representation within the lexical item. Moreover, he separates the morphophonological aspect and the syntactic-semantic aspect of lexical operations and makes them independent subparts of lexical rules.[26] The most important extension of the *Remarks* framework in this article is Jackendoff's lexical treatment of causatives. The relation between transitive causative verbs and the homophonous intransitive counterparts (*Bill opens the door* vs. *the door opens*)

[24]The recognition of lexical rules can be found in the familiar reference to a "lexicon along the lines of Aronoff" in sketches of background assumptions. Cf. Chomsky (1977:71).

[25]Jackendoff (1975)

[26]Cf. Jackendoff (1975:650) on the distinction between separate morphological (MR) and semantic (SR) subrules.

is handled by a lexical rule, and not transformationally, because of the limited generality and lexical governance of the process. Notice that in the *open* pair, no morphological operation is involved: the morphological subpart of the causative rule establishes the identity relation. The absence of a morphological operation makes it impossible to appeal to the basic principle of *Remarks* - 'derivational morphology in the lexicon' - to decide on a lexical treatment of causatives; yet, the power of lexical rules that was independently motivated for the treatment of derivational processes did not need to be extended in any sense to cover the non-derivational causative relation as well. Ironically, Fiengo's dissertation[27] proposes to subsume the causative relation, together with the equally unproductive governed process of middle formation, under the general rule of NP Movement. In the enthusiasm for the newly developed Trace Theory, the *Remarks* moral of eliminating exception ridden unproductive processes from the transformational component might occasionally be forgotten.

Oehrle's approach toward the English dative alternation[28] provides another illustration of the growing awareness that lexical rules could handle syntactic phenomena. The example is particularly telling, since nobody would propose a rule of *word formation* to treat the relation between *give NP to NP* and *give NP NP*: this relation, again, involves no morphological operation on the governing V's, but is restricted to a relation between two alternative subcategorizations for three-place predicates. Oehrle decides on a lexical treatment of the dative alternation because of the limited generality of the process. He is the first to notice the important point that the existence of a transformational alternative constitutes an unattractive indeterminacy in the grammar: one would prefer the theory of grammar to be constrained in such a way that only one alternative would remain open. The questions Oehrle raises in this respect[29] at the end of his dissertation prefigure the imminent breakthrough of the strong lexicalist position.

> (i) The domain of lexical rules (...) is smaller than the domain of transformations, since the domain of lexical rules is only material contained in a lexical entry. Is it possible to exclude transfor-

[27]Fiengo (1974) [28]Oehrle (1976) [29]Oehrle (1976:282-3)

mations from operating in this domain ?[30]

(ii) All operations performed by lexical rules must be structure-preserving (due to the definition of this term). Is it possible to treat all structure-preserving rules as lexical rules ?

(iii) Is there any connection between the notion lexical rule and the distinction between rules that have lexical exceptions and those that do not ?

Bresnan's *Toward a Realistic Model of Transformational Grammar*[31] can be considered as the major step in the development of the lexical research program since *Remarks*. This paper expounds the full consequences of the *Remarks* position with respect to the balance between the base and the transformational component; these consequences, with their devastating effect on the status of deep structure, had not been drawn before, least of all by Chomsky himself. Bresnan proposes to treat *all* lexically governed, bounded, structure-preserving processes lexically.[32] If one adopts this strong version of the lexicalist hypothesis - and one *can* adopt it without stretching the concept of lexical rules beyond what was always assumed to be their power - the properties of governance, boundedness and structure-preservation can be derived from the fact that lexical rules relate entries associated with finitely specified subcategorization features that must be satisfied by base-generated structures. If, on the contrary, one holds on to a transformational approach the properties of governance, boundedness and structure-preservation will remain unexplained; they will need explicit stipulation. The structure of the theory is loose enough to permit both alternatives; the *methodological* superiority of the lexical approach suggests that one should solve this indeterminacy by a further reduction of the role attributed to the transformational component. In Bresnan's (1976) model, the role of transformations is limited to the

[30]Cf. Williams' (1974) ordering theory which classifies transformations according to their *maximal* domain of application. A positive answer to (i) would imply that there are no transformations whose domain of application never exceeds the lexical domain.

[31]Bresnan (1976), published in a revised form as (1978).

[32]Bresnan (1976:25).

treatment of long-distance processes that fall outside the local scope of lexical processing. In later developments of her position, the transformational treatment of long-distance processes is abandoned in favour of an interpretive approach.[33]

As a result of the far-reaching reduction of the role of the transformational component deep structure and surface structure can coincide: *one* level of *syntactic* representation suffices to establish the relation between meaning and sound. In the lexicon, subcategorization frames are immediately associated with the corresponding semantic representation; lexical rules are defined as operations on these semantic representations and the corresponding syntactic frames. The elimination of structure-preserving movement rules concerns the processes that, within transformational theories, are handled by NP Movement. It is important to notice, that a surface treatment of *control* phenomena also forms an integral part of Bresnan's lexicalism. This positon toward control is known as the VP Hypothesis; it had been defended by Bresnan since 1971.[34] Under the VP Hypothesis, subjectless infinitival complements are represented syntactically as VP's (not as full S's); their syntactically missing subject is filled in at the level of semantic representation associated with the V's that select such subjectless infinitival complements.

As remarked above, Chomsky never returned to the demarcation between lexicon and syntax after his introduction of the lexicalist hypothesis. His work in the tradition of *Conditions*[35] has instead been concentrated on the development of the trace theory of movement rules. Within trace theory, cyclic movement rules are reduced to the very general format Move α, which generalizes over NP Movement and Wh Movement. The movement rules are considered to be instances of structure-preserving substitution transformations replacing a base-generated empty node by a lexically filled one, and leaving a coindexed trace on the original site of the displaced constituent. The class of possible movements is severely reduced by the interpretation of traces as bound anaphors: only the proper antecedent-anaphor relations are allowed as outputs of the general rule Move α. Conditions on anaphors are independently motivated in the semantic-

[33]Bresnan (1979), Kaplan & Bresnan (1979).

[34]Bresnan (1971). [35]Cf. Chomsky (1973).

interpretive component: Chomsky's suggestion to let the semantic conditions on anaphors govern a transformational component consisting solely of the structure-preserving Move α illustrate the strategy of reducing the content of the transformational component by appealing to the interpretive rules. This semantic attack on the transformational component *complements* the lexical attack based on *Remarks*.

By the identification of traces and bound anaphors, the semantic interpretation can be based totally on the level of surface structure: the role of syntactic deep structure is completely trivialized. Trace Theory, in other words, makes it possible to construct a non-transformational grammar, consisting of a unique level of *syntactic* representation and *interpretive* rules.[36] Chomsky avoids to take this radical position, because he argues that there is a specific property distinguishing structure-preserving movement rules from semantic-interpretive rules, viz. Subjacency. Koster[37] has claimed that pushing trace theory to its utmost consequences, the alleged distinction between movement rules and interpretive rules can be eliminated. The theory of grammar he proposes is fully interpretive: one level of syntactic structure, containing base-generated empty nodes, forms the input to the interpretive rules, which require the empty nodes to conform to the semantic conditions on anaphora. The empty nodes in Koster's framework concern bounded phenomena, formerly treated in terms of NP Movement and control, as well as unbounded processes, generally subsumed under Wh Movement. As stated above, the abstractness in syntactic structures associated with the class of bounded processes can be further eliminated by treating these phenomena lexically, i.e. without the intervention of traces. The double front attack on the transformational component seems to converge, then, in the total eclipse of transformational processing: bounded phenomena fall within the scope of lexical rules, processes with unbounded effect can be handled in an interpretive way.

1.4. *Further Developments*. Wasow's *Transformations and the Lexicon*[38] has been the most serious attempt, within the lexical research program, to define a compromise position between Bresnan's strong lexicalism and trace theory. Wasow brings together the various criteria that have been proposed

[36]Chomsky (1973:sect. 71), (1977:416)

[37]Koster (1978: 31)

[38]Wasow (1977)

since *Remarks* as bearing on the distinction between lexical and transformational rules; he organizes these criteria into a framework with a rich deductive structure, which leads him to a rejection, for example, of Fiengo's NP Movement analysis of causatives and middle formation. The crucial point of Bresnan's strong lexicalist position is the interpretation of structure preservation in bounded contexts as a necessary and sufficient condition for lexical treatment. Wasow weakens this bold hypothesis: he allows for a class of bounded structure-preserving transformations. This class appears to be very small: as stated above, causatives and middle formation are considered to be lexical. On close inspection, the bounded structure preserving transformations within Wasow's theory are verbal passives and raising cases. It is clear that this position is, methodologically, less attractive than Bresnan's strong lexicalism: allowing for a class of structure-preserving transformations prevents the elimination of Emonds' Hypothesis as a constraint on syntactic movement rules. Let us investigate, therefore, what has forced Wasow to adopt this 'less radical departure from the usual version of EST'.[39]

Lexical treatment of verbal passives and raising is excluded within Wasow's framework by his interpretation of the boundedness criterion. In its general form, this criterion states that, since the word is the maximal processing unit of the lexicon, lexical rules can only refer to the information contained in the subcategorization frame of the items they relate. Under this interpretation, the locality of lexical rules follows from the organization of the grammar. Wasow, however, gives a more narrow interpretation for the boundedness criterion. He stipulates that lexical rules are formulated in terms of particular grammatical relations (in the sense of thematic functions). Wasow's functional interpretation of boundedness accounts for a *proper subset* of the bounded rules under the structural interpretation, which simply limits the scope of lexical rules to the subcategorization domain.[40]

[39]Wasow (1977:328)

[40]The lexical passive, under Wasow's interpretation, only applies to logical objects ("themes"). Dative passive (*John was given a book*) cannot be lexical, then, since an indirect object is affected. Under the general interpretation of boundedness, lexical treatment is not excluded: the indirect object in a V NP NP structure is part of the subcategorization frame. For alleged raising-to-object cases, Wasow assumes a V S sub-

The functional interpretation of the boundedness criterion does not follow by necessity from the definition of lexical rules, contrary to the scope limitation to the subcategorization domain. In later elaborations of the (1977) position[41], Wasow has given up the attempt to combine function-dependency with boundedness; as a result, all bounded structure-preserving processes can now be regarded as lexical, whereas the property of function-dependency differentiates between minor and major *lexical* rules.

The latter part of the seventies shows a growing amount of literature in the lexicalist tradition. We single out for special reference the work by Brame and Gazdar, as these have proposed well-defined theories of their own.

Brame[42] has developed a brand of lexicalism based on his early work on the VP Hypothesis. His approach toward bounded function-dependent processes is similar to Bresnan's: these processes are handled in terms of operations on the functional representation associated with lexical items. Within Brame's framework, the level of syntactic phrase structure is totally dispensed with: it is redundant given the functional representations associated with lexical items. All interpretation is carried out on the basis of these functional structures and a set of rules of composition mapping strings of lexical items onto their functional structures directly, i.e. without a mediating level of syntactic representation. Brame's main concern, however, lies not with bounded processes, but with long distance phenomena. Long distance processes are treated by means of the interpretive mechanism of *operator binding*, an operation that plugs displaced constituents into vacant argument slots associated with lexical items. These vacant argument slots can be separated from the binding operator in an unbounded way. Binding is possible, as long as the *accessible scope property* holds. This property has the effect that no domain is accessible

(note 40 cont'd)
categorization; passives of the type *John is expected to win* cannot be handled lexically because the fronted NP bears no grammatical relation to the passivized V (it is no part of the subcategorization frame either). As stated above, the VP Hypothesis assigns *expect* verbs a frame V NP $\overline{VP}$: passivization of the NP in this frame is bounded in the general sense.

[41]Wasow (1978) and this volume.

[42]Brame (1976a), (1978a), (1979b)

for more than one long distance process.

Reducing the role of the transformational component has resulted in a renewed interest in the possibilities of phrase structure grammar. There is no longer any ground for the prejudice that phrase structure grammar would be insufficient for linguistic description. In a number of papers[43], Gazdar has demonstrated within a very explicit framework how relation-changing and long-distance processes of a transformational grammar can be base-generated. The theory presented by Gazdar is compatible with lexical-interpretive theories, but the motivation behind it is different. Gazdar is primarily interested in the generative capacity of his base-generated syntax. In (1979a) he investigates the interesting class of context-sensitive accepting systems that induce context-free languages. In (1979b), he approaches long-distance processes within a non-transformational framework that exploits the interpretation of categorial nodes as complex symbols. Eliminating the transformational component has the effect of reducing the generative capacity of the resulting class of grammars to that of the class of context-free phrase structure grammars.[44]

The elimination of the transformational component, a consequence of the full exploration of the lexicon and interpretive mechanisms, is a characteristic feature of the lexical approach. The resulting organization of the grammar leads to a more explanatory account of linguistic phenomena handled by transformations elsewhere. The appeal of the lexical research program will be apparent from its stimulating influence on interdisciplinary research in the fields of language acquisition, computational models of cognitive processing and formal semantics.[45]

[43]Gazdar (1979a), (1979b)

[44]Gazdar (1979b: 1)

[45]For language acquisition, cf. Maratsos (1978) and Roeper e.a. (1979), for work on language processing cf. Wanner and Maratsos (1978), Kaplan and Bresnan (1979). As for work in formal semantics see Dowty (1978). Thomason (1976), Keenan & Faltz (1978).

2. Principles of Lexical Organization

In the present chapter, we attempt to sketch how the grammar as a whole could be organized in accordance with the thematic commitments of lexicalism. The foregoing will have shown that the general program of lexical grammar covers quite different approaches; the lack of an accepted formalism tends to give undue prominence to the differences among individual work. We will not concentrate on differences here, but stress the underlying similarities in the ideas guiding the lexical research. The resulting sketch is heavily influenced by the authors cited in the references; it might well be the case, however, that neither of them would fully agree with the total picture we have distilled from their work.

2.1. *Formal Properties of Lexical Items.* The rules of a transformational component apply to *simple* objects: they relate phrase markers. The class of formal objects which forms the input to lexical rules is not simple but *composite*. The entries related by lexical rules contain several distinct types of information, viz., (i) morphophonological information associated with a categorial label, (ii) contextual syntactic properties represented in the form of a subcategorization frame, (iii) the translation into the language of semantic representation associated with the particular syntactic frame. The properties of the operations which make up a lexical rule are determined by the formal characteristics of these different types of information.

First, with regard to the morphophonological operations Jackendoff[46] has rightly observed that the canonical derivational rules exhibit *structure-building* power. This makes them the lexical analogues to phrase structure rules. On the morphological level, the rule of *able* affixation can be regarded as an expansion of the category A into a concatenation of a verbal stem and the affix *able*, together satisfying the "is a" relation with respect to the expanded symbol. The application of additional derivational rules would result in further hierarchical structure.

(1) A → V able

[46]Jackendoff (1975:668)

In the second place, the syntactic frame, or subcategorizational feature, serves a double function. First, it defines the context for lexical insertion. Secondly, when a lexical rule relates two entries, the subcategorizational frames form the syntactic input and output of the mapping defined by the rule. Chomsky was the first to suggest that subcategorizational features should not be regarded as atomic predicates (as features generally are), but as strings with internal structure.[47] Lexical insertion, on this view, is a transformation with the syntactic frame as its structural description. In the same vein, Vergnaud[48] describes the mappings between syntactic frames of related entries in terms of *lexical transformations*. One will notice that the structural change associated with lexical insertion is trivial: it consists in the replacement of the terminal element Δ by a lexical item. The structural effect of the mapping between frames of related entries is trivial in another way; it is by necessity structure-preserving, in the sense that both of the related entries must be inserted into base-generated structures.

Third, the relation between a syntactic frame and the corresponding translation is governed by a principle of compositionality: subcategorizing an item for a particular frame entails that its meaning is built up in a compositional way as a function with the translation of the phrases mentioned in the frame as arguments. Lexical rules manipulate the input function associated with a frame. They rearrange the way the arguments plug in into the original function, and change the degree of a predicate by increasing or decreasing the number of arguments. The device of *lambda abstraction* allows one to represent these manipulations formally. The *able* rule might again serve as an illustration.

(2) ABLE (i)=j for
$\langle i, [\Delta]_V$, [V NP], V'(NP')$\rangle$ ·
$\langle j, [[\Delta]_V$ *able*$]_A$, [A], $\lambda x \Diamond \exists y [V'(x)(y)]\rangle$

The output of (2) is a complex intransitive expression, which will map into a sentence denotation when combined with the NP denotation corresponding to the grammatical subject. The lambda operator in the transla-

[47] Chomsky (1965:122) [48] Vergnaud (1973)

tion of the *able* adjective abstracts on the variable corresponding to the logical object of the input verb. The effect of combining the lambda expression with the translation of the grammatical subject will be, then, that the subject is understood as the logical object of the V to which *able* was affixed. The device of property abstraction used here to realize the proper mapping between syntactic frame and logical argument position, is not restricted to lexical rules. It will equally be used in the case of long-distance processes, e.g. topicalization, where one will abstract on the hole variable corresponding to the gap in the construction, and subsequently apply the resulting function to the translation of the topicalized constituent.[49] Yet, in the lexicon, the λ operator is restricted in its scope to abstract only on arguments of the subcategorizing predicate, whereas no such restriction holds for long-distance cases.

This brief discussion will suffice to show that each type of information within the lexical item forms the input to a particular type of formal operation. Consequently, we will not view lexical rules as monolithic wholes, but as *compounds of elementary operations*, i.e., as n-tuples taken from the sets of morphological, syntactic and semantic operations defined over the lexicon. This analytic approach toward lexical rules brings about the following problem shift with regard to the demarcation problem.

In his original discussion of the balance between lexicon and syntax, Chomsky suggested to approach the demarcation problem in terms of *rule typology*; a set of distinctive properties of lexical vs. transformational rules would have to provide the tool for drawing the boundary. The ensuing quest for criteria based on rule properties, epitomized in Wasow (1977), has led to the situation sketched in the historical reconstruction: they are too weak to draw a sharp distinction between lexical and transformational rules.

[49]Cf. Gazdar's rule scheme for Topicalization (1979b:13). In the translation, *h* is an unindexied variable ranging over denotations of type α (i.e., NP denotations if α=NP, PP denotations if α=PP, etc.). S/α is a *derived constituent*, c.q. a S containing a gap of type α (cf. also Gazdar (this volume) for the notion 'derived constituents'].

(i) $< 44, [_S\ \alpha\ \ S/\alpha], \lambda h[(S/\alpha)'](\alpha') >$

We suggest that the difficulties with the rule-typological approach are inevitable, and intimately linked with the atomic concept of lexical rules. We abandon this view in favour of a theory which analyzes lexical rules as compounds of elementary operations. Since the types of information contained in a lexical item form a cross-section of the levels of representation in the grammar as a whole, it will not come as a surprise that from a formal point of view the inventory of rule types needed in the grammar is quite small and that, in fact, the available set of rule types suffices to characterize the components of a lexical rule.

Given the concept of lexical rules as compounds of elementary operations, the demarcation problem can be approached from a different angle. The frontier which sets apart the lexicon from other components will no longer be mapped out in terms of rule types, but in terms of rule *scope*: each component fixes its specific parameters on a shared set of formal rule types. The lexical parameters reflect the basic principle that rules operating in this component have access exclusively to information associated with particular lexical items, and not to the sum total of information contained in the P-markers in which these items appear. The limited, local scope of lexical rules, then, need not be postulated; it rather follows from an uncontroversial principle of lexical organization.

The following sections are elaborations on the strategy of translating the demarcation issue as a scope problem. The paragraph on word-structure sets off the lexicon from the phrase structure rules and the phonological component. The paragraph on lexical syntax discusses the scope distinctions between lexical rules and transformational or semantic-interpretive rules.

2.2. *Aspects of Word Structure.*

2.2.1. *Lexical Morphology.* In this section we will deal with the formal aspects of lexical rules one of whose constituent parts is an operation at the morphophonological level. We will propose a theory that accounts for the formal relation that exists between the first member of the ordered pairs that constitute these rules (cf. the *able* rule in (2)). It has been suggested above that rules dealing with the syntax of morphemes are not typologically distinct from rules dealing with the syntax of words (Phrase Structure Rules). In recent literature, this observation is echoed in a number of implicit and explicit suggestions to the effect

that word structure could be accounted for in terms of an X-bar system below the word level.[50] In what follows, we will explore the possibilities of a system of *Word Structure Rules* (WSR'S) analogous to the phrase structure rules of the syntactic base.

A proviso has to be made from the start. The theory of word structure presented here should be interpreted in a relativistic spirit. The system of WSR's captures in an elegant way the formal properties of a subset of the known word formation processes which might be characterized as *concatenative word based morphology*. According to Aronoff the class of items that can be input to a word formation rule is in fact limited to *words*.[51] His position implies that regular word formation processes do not derive words from less than an inflectional stem. Clearly this implication cannot be a universal claim since there are languages exhibiting a root-based morphology. This makes the theory of word based morphology into a relative claim subject to typological differences.[52]

We will assume that the hierarchical structure of complex words results from a contex-free system of word structure rules, generating an infinite

[50]Cf. Lapointe (1978)

[51]Aronoff means by *words* inflectional stems in the sense of Matthews (1974). Aronoff's claim implies then that we will not find internal inflexion, i.e. derivational affixes which are peripheral to inflexional affixes. Cf. Nida (1946) and for some counterexamples to this 'uninflected base hypothesis' Moody (1978).

[52]Even for English the claim may prove to be too strong. There seem to be derived words (e.g. *nominee*) which contain less than an inflectional stem. The claim that *nominee* is derived from *nominate*, though true from a semantic point of view, does not hold at the formal level here. However, Aronoff persists in a word based operation at the formal level using readjustment rules, (truncation rules), to delete the morpheme *ate*. This approach seems justified in this case especially because truncation involves a meaningless string of segments, rather than a true morpheme (Rardin 1975). We consider this to be a borderline case. An alternative would be to use suffix substitution rules, which may be more adequate when we are dealing with true morphemes on both sides. Clearly outside the scope of our system are words like *submit, predict*, etc., which are neither formally nor semantically word based. Moreover, a transformational account will have to be invoked for word formation processes like infixation, reduplication and perhaps category-switch (cf. Aronoff 1976, Allen 1978).

It seems then that the morphological component as a whole consists of various subsystems, only one of which is described here. However, by describing this subsystem we characterize the core of the word formation component of many languages.

set of well-formed word frames. The context-sensitivity governing the distribution of affixes is captured in the form of a finite set of insertion conditions associated with the affixes; their function is identical to that of subcategorization frames in syntax, i.e. they restrict the possibility to occur in word frames. The separation of context-sensitive distributional properties from context-free hierarchical structures is based on the same division of labour commonly adopted for the syntax of words (base rules vs. lexical insertion conditions).

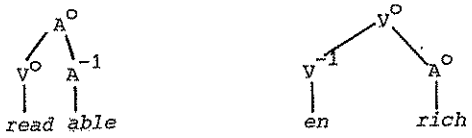
The complex symbols of the WSR component are defined as ordered pairs consisting of an integer *i* and a syntactic feature matrix. The features are taken from the set {±N, ±V, ±m, ...}; they define the categorial status of the morpheme. The integer defines the level of the category; it will become clear that for the purpose of describing the structure of complex words a distinction between two levels is sufficient. We will assume, then, that the value of the integer *i* is either 0 or -1. Let us discuss now the interpretation of the complex symbols defined by the WSR component.

The central claim of an X-bar system is that complex structures (phrases in syntax, derived words in morphology) are projected from *heads*; the categorial information of the head is carried over in the projection; only the level changes. Accordingly, the canonical rule schemes for the two-level WSR system would be (3) and (4), for suffixation and prefixation respectively. Sample structures generated by these rules are given in (5).

(3) $X^{0} \rightarrow Y^{0} \quad X^{-1}$

(4) $X^{0} \rightarrow X^{-1} \quad Y^{0}$

(5)



The above structures exhibit the following properties. The *affixes* are considered to be the heads of the construction: affixes receive a set of syntactic features; the features of the affix determine the category of the structure they project. The system of WSR's thus expresses

the traditional view that affixes determine the category of derived words.[53] Affixes themselves are of level -1, i.e. they are *bound morphemes*. They maximally project structures of level 0, which are free morphemes. The level 0, then, is considered to be the recursive cyclic domain of word formation, i.e. the maximal processing unit of the WSR's. Affixes select complements of a certain categorial type; these complements, or *bases* as they are generally called in the morphological literature, are themselves of the maximal level 0. In this way, the WSR system captures the characteristic property of word based morphology.

It has been frequently observed that prefixes usually do not have the power to change the categorial value of their bases. This claim is not valid in an absolute sense, as the existence of category-determining prefixes shows. It is clear, however, that among prefixes the elements with head-like properties form the exception. Therfore, we consider (3) to represent the unmarked orientation of head vs. base in a WSR, the case of a head preceding its base constituting the marked rewrite option. Consequently, we distinguish two classes of prefixes: those that function as heads in the WSR system, introduced by (4), and the much larger class of category-neutral prefixes.

To account for the properties of the latter class, we introduce the feature [±major] ([±m]), which is used here with the same interpretation as it has in syntax.[54] The feature distinguishes between projecting and non-projecting members of the same categorial type: [+m] X's are heads projecting an expansion of the maximal type; [-m] X's are non-heads and do not project higher structure. Category-neutral prefixes, then, will be assigned the feature [-m]; they are introduced by the following non-canonical rule:

$$(6) \quad X^{0} \rightarrow \underset{[-m]}{Y^{-1}} \quad X^{0}$$

In rule (6), it is not the prefix, but the base which functions as the head of the construction: the derived word has the same feature specification as the base.

The observation that in the unmarked case the rightmost member of an expansion functions as the head has traditionally been extended beyond

[53] Cf. Marchand (1969) [54] Cf. section 2.3.2.

the domain of derivational morphology: it is claimed that in compound formation too the second member is the categorial head. Marchand generalizes this observation in the scheme AB=B, which is held to apply to prefixation and compounding.[55] Allen (1978) captures this insight in her 'IS A' Condition which is explicitly meant to generalize over *all* word-formation processes.[56] Williams (1978) also maintains that it is always the second member of a derived word which is the head. In the system proposed here Marchand's scheme would be relevant for compounds, which we have not dealt with. It is redundantly relevant for the unmarked core of the WSR's.

So far, the following types of lexical entries have been distinguished.

(7) a. *read*, $\begin{bmatrix}+V\\-N\\+m\end{bmatrix}^{0}$, ...

b. *-able*, $\begin{bmatrix}+V\\+N\\+m\end{bmatrix}^{-1}$, $< \begin{bmatrix}+V\\-N\\+m\end{bmatrix}^{0} \text{ ---} >$, ...

c. *en-*, $\begin{bmatrix}+V\\-N\\+m\end{bmatrix}^{-1}$, $< \text{---} \begin{bmatrix}+V\\+N\\+m\end{bmatrix}^{0} >$, ...

d. *pre-*, $\begin{bmatrix}-V\\-N\\-m\end{bmatrix}^{-1}$, $< \text{---} \begin{bmatrix}+V\\-N\\+m\end{bmatrix}^{0} >$, ...

The previous paragraphs concern the context-free aspect of our system

[55] Cf. Marchand (1969:11,129)

[56] In Marchand (1969) and Allen (1978) the AB=B scheme has a semantic impact as well. Cf. the fact that in compounds generally the first member denotes the *kind of* thing denoted by the second member. With respect to affixes the semantic impact is also relevant. E.g. from a semantic point of view *-able* is the head: in its semantic translation the complement is represented by means of a variable. In a certain sense, it seems counter-intuitive to attribute semantic head status to affixes as opposed to what is traditionally called the base. This impression is based on a confusion of semantic *head* and the difference in the meaning of free and bound morphemes.

of WSR's. As indicated above, the context-sensitive distributional properties of bound morphemes are formalized as insertion conditions. These function as structural descriptions for the insertion of affixes; they factorize the structure generated by the WSR's and check whether this structure satisfies the insertion condition. Examples of insertion conditions associated with particular affixes can be found in (7b, c, d): they specify the categorial status of the base, and indicate whether the base follows or preceeds the affix.

Because insertion conditions involve transformational power, we would like to impose tight constraints on their capacity to analyze word structure. The analytical concept of lexical rules allows us to restrict the responsibility of an insertion condition to *morphological* well-formedness. All types of ill-formedness which are not of a purely morphological kind can be accounted for in the appropriate subparts of a lexical rule.[57] In this sense, the fact that *un-* cannot be attached to a base with negative content will not follow from its morphological insertion condition: a *semantic* constraint rules out the morphologically well-formed *unbad*. Similarly, the fact that *-able* requires a transitive base will appear from the syntactic subpart of the *-able* rule, not from its morphological insertion condition $\langle V^{O} \text{-} \rangle$.[58]

With the above restriction to morphological well-formedness in mind, we describe insertion frames in terms of (8). The categorial frame is separated from the morphophonological frame because the former, which specifies the category of the base, is obligatory while the latter is not. Morphophonological frames may be added in the form of Boolean conditions on the basic categorial frame, if an affixation rule requires such extra conditions.

[57]The suggestion to account for the ill-formedness of derived words in a 'modular' way is also found in Siegel (1978).

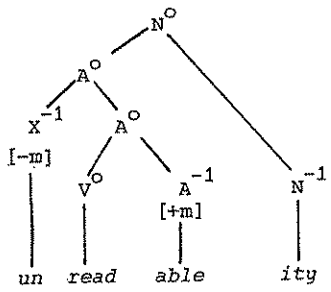
[58]Aronoff's *Blocking Principle* is concerned with a type of ill-formedness which falls totally outside the scope of lexical rules as conceived here, and might rather be accounted for in terms of a theory of economy. The principle states that the presence of one word in the lexicon can block the entering of a second one if both are derived from the same stem and have the same meaning.

(8) a. *Categorial* frames are of the form <α___ β>
where either α or β is empty and α and β belong to
the set defined by the syntactic features {±N, ±V, ±m, ...}

b. *Morphophonological* conditions on a categorial frame
are Boolean conditions on a., referring to
(a) the phonological string
(b) a specified morpheme
(c) a morphological feature[59]

Constraints on the power of insertion conditions take the form of restrictions on the material that can be analyzed by the variables in the formulation of (8). The *Adjacency Condition* (cf. note 22) is an attempt in this direction. Within a framework very similar to the WSR system presented here, Williams[60] has refined the Adjacency Condition as a *Head Condition*, which can be interpreted as a locality restriction on the maximal *depth* of structure that can be analysed by an insertion frame. In terms of the Head Condition, an insertion frame for an affix A can refer to the head of the base B to which A is attached, but not, e.g. to the "complement" of that head. Within the theory presented here, two different types of prefixes are distinguished: [+m] prefixes counting as heads, and [-m] non-head prefixes. For this theory, then, the Adjacency Condition and the Head Condition make different predictions with regard to structures like (9).

(9)



[59]In Siegel (1974, 1978), Aronoff (1976) and Allen (1978), several such restrictions are discussed. The suffix *-al* requires that "that if the verb ends in a consonant, the consonant must be [+ant]" (Siegel 1974: 166). Siegel (1978) argues that *un* cannot be attached adjacent to *dis*. The suffix *ity* can only be attached to a [+latinate] base.

[60]Williams (1978)

A condition on the insertion frame for *-ity* requires the presence of the feature [+lat] in its base. Given the Adjacency Condition, only the material added in the cycle preceding *-ity* affixation can be analyzed, i.e. the prefix *un-* which is [-lat]. When the Head Constraint factorizes the base for the *-ity* rule, it can descend to the head of that base, i.e. to *-able*, which is [+lat]. Moreover, it is necessary to go down to *-able*, since features such as [±lat] do not percolate upwards: adjunction of a [+lat] affix does not make the derived word [+lat]. It seems, then, that the locality required by the Head Condition lends support to the distinction between [+m] and [-m] prefixes.

Up till now, the discussion has been restricted to derivation. How would inflection be treated within the framework of lexical grammar ? On the one hand, the grammar will have to be organized in such a way that the fundamental differences between inflection and derivation are properly expressed: for example, inflectional rules are fully productive and transparant in meaning, they never change categorial labels. On the other hand, one would like to stick to the principle of lexical integrity, which treats words (inflection included) as *atomic* with respect to syntactic rules. This makes it impossible to generate inflectional morphemes on abstract base positions, and to adjoin them later to the appropriate stems by means of transformations. Let us therefore distinguish, within the lexicon, between a *store* and a *process* component. The lexical store takes care of all non-inflectional aspects of word formation; it is organized as a full-entry component augmented by a set of lexical rules factoring out redundant information. Inflected forms are not stored, but processed within the lexicon: on the basis of the output of the store, the lexicon processes the members of the inflectional paradigms, which can then be inserted, in their fully specified inflectional form, in a base-generated syntactic structure.[61]

[61]The occurence of internal inflection (cf. fn. 51) though a marked phenomenon, can be adduced as support for the decision to extract all inflection from the syntax: if inflection would be dealt with by syntactic rules, word internal inflection would require multiple interaction between the lexicon and the syntax, whereas Lexical Grammar wants to restrict the interaction between modules to one point of contact. The fact that inflected words are not stored accounts for the rarity of internal flection. Word formation on inflected bases could be compared to word

Formally, inflectional rules will be considered as 'spell out' rules of the general type (10).[62]

(10) $$/X/ \rightarrow /X+a/ \quad \begin{bmatrix} +F \\ -G \end{bmatrix}$$

The features spelled out by these rules are of two kinds: they are either *morphosyntactic* or *morpholexical*. Morpholexical features are local in the sense that they are only present at the X^{o} level; they do not project up to higher levels. The features dealing with verbal aspectual morphology can serve as an example here. Morphosyntactic features are projected from the head up to the phrasel X^{n} level; they are involved in phrase-internal and interphrasal agreement phenomena. Examples are number or person features in English. As stated above, the transformational treatment of agreement phenomena typical of older EST versions is excluded by the principle of lexical integrity. Non-transformational alternatives would treat agreement by means of filtering devices, or, by making full use of complex categorial symbols, directly in the base.[63]

2.2.2. *Lexical Phonology*. The more restrictive versions of Generative Phonology differ from standard versions[64] in two respects. First, strong constraints are placed on the abstractness of lexical representations and second, it is explicitly acknowledged that different manifestations of morphemes can be related by different types of rules.[65]

With respect to the abstractness of lexical representations (henceforth

(note 61 cont'd)
formation on the basis of potential though non-occuring words which is possible but, just like internal inflexion, not very common.

[62]Cf. Anderson (1977a) for a description of some of the properties of these rules. Anderson's framework is comparable to the traditional treatments of inflection in generative grammar. Cf. fn. 7.

[63]Cf. Gazdar (1979a) for a discussion of this possibility.

[64]Chomsky and Halle (1968).

[65]For a more detailed account of developments in phonology since Chomsky and Halle (1968) cf. Anderson (1979), Van der Hulst (1979).

LR's) the main differences depend on the acceptance or rejection of some version of Kiparsky's *Strong Alternation Condition*, i.e. the acceptance or rejection of properties of lexical items which never show up in the surface. Though this is not the place to weigh all the arguments, it will be evident that allowing such abstract properties clashes with the surface oriented approach which is characteristic of Lexical Grammar.[66] Consequently we will assume here that:

> a lexical representation is ... concrete in the sense that it is a *surface plan*, a plan to produce a certain kind of observable behaviour. It can hardly contain features specifications which would *always* be contradicted when the behaviour is actually carried out. (Linell 1979:245).

This statement does not yet settle the question as to how much phonetic detail is included in LR's. Positions ranging from 'non-distinct archisegmental', via 'roughly phonemic' to 'fully specified' have been defended.[67] It seems that only the latter position is consistent with the full entry theory of the lexicon. Therefore we will furthermore assume that LR's are fully specified representations of the careful pronounciation of words.[68]

A morpheme can have different manifestations depending on the environment it occurs in. This phenomenon is called *allomorphy*. In a framework in which LR's are as concrete as is assumed here the rules which account for allomorphy will appear to be of two different types, grossly speaking. Some rules will refer in their SD to nothing but phonological information (i.e. phonological *units* and phonological *domains*).[69] Other rules, however, will have to refer to specific lexical items (or diacritic features associated with these items) or to specific (classes of) affixes. The latter rules have been called *morpholexical* rules while the former are referred to as *phonological* rules.[70] Furthermore, it has been noted that the charac-

[66]Cf. Kiparsky (1968). For a clear survey of various positions cf. Kenstowicz and Kisseberth (1977).

[67]Cf. Hooper (1976), Linell (1979), Vennemann (1974) respectively.

[68]Comments on the notion 'careful pronounciation' can be found in Hooper (1976) and Linell (1979).

[69]Among this class are the segment structure and syllable structure rules. However we will only discuss here rules that deal with allomorphy.

[70]Cf. Anderson (1975). Aronoff's (1976) allomorphy rules are a subset

teristic properties of the morpholexical rules are very much like the properties that are attributed to the bounded syntactic rules, which are reanalyzed as *lexical* in the framework adopted here:

> Interesting parallels between word phonology and single-sentence syntax begin to emerge, which could be understood if both were viewed as lexical processes: the fact that they and only they can be lexically governed (i.e. apply to partially arbitrary classes of words), that they are structure preserving (which in phonology is nothing but the familiar property of being neutralizing), the fact that they never apply to the output of unbounded rules (i.e., to the outputs of sentence phonology and of root transformations like Q-movement, respectively). Kiparsky 1978, 42.

One can find the suggestion to place at least the morpholexical rules in the lexicon also independently motivated at other places.[71] However, there is a problematic aspect to all these proposals, when treatment in the lexicon is confined to *morpholexical rules*. Various publications on rule typology in phonology have shown that it is impossible to establish a clear-cut typological demarcation line between morpholexical and phonological rules: there are too many borderline cases.[72]
The dichotomy between lexical and non-lexical phonological rules can then be more profitably approached from the point of view of rule scope (in fact this is implicit in Kiparsky's remark cited earlier). The decisive criterion will then be formulated as (11):

(11) All phonological rules whose SD and SC are restricted to the domain of lexical entries are lexical

(note 70 cont'd)
of the morpholexical rules.
It should be noted that the acknowledgement of at least two types of rules dealing with allomorphy marks the abondonment of holism in phonology, where until then one class of rules mapped the underlying phonological representations onto surface phonetic representations. Similarities in the make up of the syntactic and the phonological component which existed under the holistic approach, also exist then under the 'modular' approach.

[71]Cf. Dell and Selkirk (1978:29): "And suppose furthermore that it turned out that all rules mentioning morphological features (call these *morpholexical* rules) applied either in the lexicon or as a component at surface structure prior to the application of the phonological rules".

[72]Cf. Dressler (1977), Linell (1979).

Given (11), *lexical phonology* contrasts with *sentence phonology*. It is important to notice that rules belong to one of these categories according to their *maximal* domain. Rules of sentence phonology affect lexical entries, when these have been inserted in base structures, but their triggering environment crucially extends beyond the word domain. Intonational rules and rules of external sandhi are the major classes of sentence phonology.

How do the rules of lexical phonology apply to lexical entries ? Since we have assumed that LR's are fully specified phonetic representations all rules function as redundancy statements: the rules capture the predictable aspects of the sound shape of lexical items. Rules that are dealing with allomorphy, however, have additional functions. First they play a role in computing morphological redundancy. In a structure [[X] + affix] the morphological redundancy can only be established if the X in the environment of the affix can be identified as the same morpheme as a X minus the affix. E.g. the rule of *velar softening* in English accounts for the fact that the sound *k* can be identified as the sound *s* in the environment of *ity*. Second, they function as distributional statements. We have proposed not to store inflected forms. But sometimes different inflectional categories select different allomorphs (e.g. Spanish *contar, cuento*). In such cases the lexical entry will have to contain more than one stem and the rules must make clear which stems occurs in which inflectional category.[73]

[73]One would like to restrict this 'multiple listing of stems' (cf. Hudson 1974, Hooper 1976, Lieber 1979b, Tiersma 1978, Orešnik 1979), to cases in which the allomorphy is governed by a non-productive rule. This would imply then that only in inflectional categories automatic phonological rules apply 'right-side up'. The role phonological rules play in computing morphological redundancy should not be identified with "upside down" rules in the framework of Leben and Robinson (1977), Leben (1979). This framework suffers from similar defects as the standard model. The main difference lies in the fact that abstract underlying representations are not stored in the lexicon, but 'created' by the phonological rules (cf. Janda (1978) for critical comments). In the framework proposed here the lexical rules relate surface segments *directly*. Tiersma (1978) outlines a very similar framework. The main difference with ours is his assumption that automatic rules always apply right-side up and that, consequently, lexical representations abstract from the effect of these rules.

2.3. *Lexical Syntax.* The analytic concept of lexical rules implies that a rule *can* but *need not* involve an operation at each level of representation within the lexical item: the only a priori constraint on possible combinations requires that the items linked by a lexical rule show morphological relatedness.[74] Since the limiting case of relatedness is identity, one can expect lexical rules that do not contain a morphological operation among their constituent parts. The debate on the trade relation between lexicon and syntax focuses primarily on these putative lexical rules showing no sign of derivational morphology. The master criterion of *Remarks* - transformations do not perform derivational morphology - cannot be invoked to classify this type of rule: it characterizes derivational morphology as a *sufficient*, not as a *necessary* criterion for lexical treatment. In order to decide on lexical or syntactic derivation, other distinguishing properties will have to be adduced. Following the approach toward the demarcation problem adopted above, we will seek these lexical properties in *scope* restrictions.

When the maximal processing unit of the lexicon is the word, the syntactic part of a lexical rule will have access to the information contained in the subcategorization frames of the items it relates. Scope limitations on the power of lexical rules, then, can be derived from the conditions on the structure that syntactic frames are allowed to factorize. The information contained in a syntactic frame has to be entirely motivated in terms of the contextual information necessary for the process of lexical insertion; lexical rules are given access to this information motivated by the insertion condition, and to nothing more.[75]

2.3.1. *The Translation Principle.* The arbitrarity of the syntactic environment mentioned in the insertion condition for the head of a phrase is severely limited by a semantic constraint governing the relation between syntactic subcategorization and the associated translation into the language of semantic representation. We will refer to this constraint as the Translation Principle. It might be formulated as follows:

[74] Jackendoff (1975:651)

[75] Cf. Oehrle (1976:283): "The problem of characterizing the possible structural relations amenable to treatment by lexical redundancy rules revives the problem of the depth and width of subcategorization frames."

The Translation Principle

Subcategorizing a X^{o} for a syntactic frame amounts to the claim that its meaning is a complex function with the various set-theoretic objects corresponding to the subcategorized phrases playing the role of arguments.[76]

By the Translation Principle, each syntactic frame is explicitly associated with a corresponding expression in the language of semantic representation; the meaning of the latter is built up compositionally from the meaning of the parts mentioned in the syntactic frame.

The adoption of the Translation Principle has important consequences for the status of syntactic deep structure. If the lexicon explicitly associates a translation with each syntactic frame, there is no longer any need for a *syntactic* level of deep structure where logical and syntactic functions coincide in a one-to-one way. The adoption of the Translation Principle, in other words, creates the possibility of a *surface syntax*: the traditional function of deep structure, viz. realizing the mapping between syntactic and logical arguments, is taken over by the translation.[77]

We will illustrate some salient properties of a lexical theory based on a sole level of syntactic representation with a fragment from Gazdar. The rules in (12) represent the syntactic frames and the associated translations for *believe*, *persuade*, and *promise* verbs.[78]

[76]This formulation is based on Dowty (1978:418). See also Jackendoff (1977:57): "Those lexical items which strictly subcategorize phrases in their environment can be thought of as semantic functions which take as their arguments the interpretations of the strictly subcategorized phrases".

[77]We saw above that this conception was originally due to Bresnan (1976) who proposed to extend the lexical representation with a level of functional structure, defined independently from the syntactic frame. In our exposition here, we shall follow Gazdar (1979a) who uses the lambda calculus as language for semantic representation. We shall not go into the question of which formal semantics is most preferable.

[78]Cf. Gazdar (1979a:18). These rules are triples of which the first member is a unique arbitrary integer (the number of the rule), the second member is a syntactic frame, and the third is a semantic rule showing how the semantic representation of the expression in the syntactic frame is built up from the semantic representation of its parts (ibid., p. 10). V^{*}_{n} represents the class of lexical items introduced by rule *n*.

(12) a. $\langle 18, {}_{VP}[V\ NP\ \overline{VP}],\ \lambda x[V'(\overline{VP}'(x))]\ (NP')\rangle$

$V^*_{18} = \{believe, expect, \ldots\}$

b. $\langle 19, {}_{VP}[V\ NP\ \overline{VP}],\ \lambda x[V'(\overline{VP}'(x))(x)]\ (NP')\rangle$

$V^*_{19} = \{persuade, force, ask, \ldots\}$

c. $\langle 20, {}_{VP}[V\ NP\ \overline{VP}],\ \lambda x[V'(\overline{VP}'(x))(NP')(x)]\rangle$

$V^*_{20} = \{promise, \ldots\}$

From a categorial point of view, the frames introduced by these three rules are identical. They represent the surface environment for the lexical items in question without introducing any syntactic abstractness. The difference between the three sets of lexical items lies in the *translations* induced by these identical syntactic frames. The one-to-many relation between a single frame and three distinct translations does not threaten compositionality: the translations are uniquely associated with each type of verb by the rule integer, which can be thought of as a feature on the lexical items in the set of V's it introduces.

The verb classes of (12) can be distinguished on the basis of their control properties and argument structure. As to the latter, a distinction has to be made between arguments and non-arguments in the sense of Freidin (1976). A constituent counts as a non-argument with respect to a given verb when it occurs in the syntactic frame of that verb (i.e., when it is a subcategorized syntactic argument) but does not function as a *logical* argument of the predicate in the translation.[79] An example of a non-argument is the syntactic object of *believe*-type verbs. The translation associated with this type shows that *believe* has only one argument, the propositional object $(\overline{VP}'(x))$, in which the translation of the syntactic object, NP', plays the role of subject. This syntactic object has no logical function with respect to *believe* itself, then. This state of affairs can be contrasted with *persuade* cases: for this type of verb, the syntactic object functions both as logical object and as controller of the complement subject. *Believe* in other words, is a two-place predicate, *persuade*

[79]Non-arguments in the sense of Freidin (1976) can be identified with the landing sites of NP Movement rules in transformational theories. Cf. Koster (1978:24).

a three-place one.

Control properties differentiate *persuade* from *promise* cases. For the latter, the understood subject of the complement is the matrix subject, whereas the former have their complements controlled by the matrix object. Within lexical grammar, the distinction is again captured on the level of semantic representation; syntactically, the *persuade* and *promise* frames do not differ. The translation of a *persuade* VP consists of a complex transitive function combined with an NP argument. The argument corresponds to the translation of the syntactic object, the complex function is the expression within the scope of the lambda operator. One must interpret this function-argument structure in such a way that the NP denotation is plugged in for the occurrences of the variable which the lambda operator abstracts on: the translation thus establishes the fact that the syntactic object NP is the controler of the subject of the complement. Compare this with *promise*. The translation of the *promise* VP consists only of the complex intransitive function formed by the lambda operator; the translation of the syntactic object of *promise* is within the scope of this operator here. Within the translation of a *promise* VP, then, there is no argument satisfying the lambda function. This function will only be satisfied when the intransitive VP translation is combined with the NP denotation corresponding to the subject of *promise*. Plugging in the NP denotation of the subject for the occurrences of the variable *x*, we see that the syntactic subject of *promise* controls the complement.

The interpretive treatment of control illustrated above is known as the VP-Hypothesis. Tenseless complements are not generated as full S's with empty PRO subjects, but as simple VP's, which is their observable structure. These VP's are *translated* as propositional arguments, of course; their syntactically unexpressed subject is filled in in the translation associated with the matrix predicates. The motivation for an interpretive treatment is derived from the fact that grammatical control is not an exclusive property of tenseless complements: XP complements of all kinds can be grammatically controlled.[80] In the following two pairs from Bach (1979:520), (13) is accounted for in terms of an abstract syntactic element PRO and the Specified Subject Constraint, within trace theory; (14)

[80]Cf. Bresnan's treatment of XCOMP's in Bresnan (1979), and Bach (1979).

shows the same interaction of control properties with anaphoric antecedent selection.

(13) a. Mary persuaded the men to kill themselves/*herself
b. Mary promised the men to kill herself/*themselves

(14) a. I regard John as proud of himself/*myself
b. I strike John as proud of myself/*himself

If (14) requires the grammar to provide an interpretive procedure to establish the underlying subject-predicate structure of object and predicative complement - and surely nobody would nowadays propose an abstract *syntactic* source for (14) as Generative Semantics would have done - the same interpretive procedure will make abstract syntactic structure in the case of (13) superfluous: the abstract syntactic source for controlled subjects in tenseless complements is an artificial limitation of the domain of grammatical control to one of the major phrases, VP.[81]

The lambda expressions appearing in the above translations are equivalent to the notion "phrasal" as distinguished from "basic" (i.e. lexical) verbs used within the Montague tradition. The adoption of phrasal categories alongside lexical ones enables theories based on Montague Grammar to characterize a relation changing rule like Passive in a uniform way for the different instances of this rule in English: Passive takes as its input a transitive verb (i.e. a lexical *or* a phrasal transitive verb), and affects the object argument of this categorial type. The notion "object argument" is uniquely defined as the argument which, combined with a tran-

[81]Chomsky (1978) recognizes that (13) and (14) should be accounted for in a generalizing way. He proposes to extend the opacity condition (i.e. SSC) to examples like (14): 'the opacity condition applies without modification if we assume that among the rules of interpretation there are "structure-building" rules that assign to [(14)], representations such as [(i)] in LF.

(i) a. they regard me as $[_{\overline{S}}$PRO be very much like each other/(them)$]$
b. I impress them as $[_{\overline{S}}$PRO be very much like each other/(them)$]$

Chomsky notices that these structure-building rules are "reminiscent of earlier work that relied on transformational rules for lexical decomposition."

sitive verb (lexical *or* phrasal), results in an intransitive verb denotation.[82] Because Passive, within Lexical Grammar, would be characterized as an operation defined on the semantic representation of lexical items, and because the semantic representations within the language of the lambda calculus incorporate the lexical/phrasal verb distinction, the same characterization of Passive is possible here[83], as Gazdar's theory effectively shows. Passive can be defined on the structures represented in (15), it has (16) as output, i.e. it turns transitive VPs into intransitive ones, where the syntactic subject of the passive construction is mapped on the logical object of the original function. The important thing to notice is that the variable F in (15) ranges over lexical *or* complex, phrasal functions. Rules <18> and <19> of (12) satisfy the input for Passive (with F equal to the lambda expression); <20> does not, since in the VP translation, there is no NP denotation satisfying the lambda function.

(15) <m [V NP X], F (NP') >

(16) <n [V X], λx[∃y V'(x)(y)] >[84]

2.3.2. *Conditions on Subcategorization.* The conditions on subcategorization frames proposed in the literature are attempts to couch the Translation Principle into syntactic terms. The oldest example is Chomsky's principle of strictly local subcategorization. As a general condition on the structural description of the lexical insertion transformation, the Strict Locality Principle states that the syntactic domain of a frame [α _ β] for the insertion of a lexical item X^0 does not extend beyond the structure dominated by the category symbol that introduces X^0, i.e. X^1 within $\overline{X}$ theory.[85] The Strict Locality Principle is based on the assumption that the X^1 level defines the necessary and the sufficient domain for the semantically relevant constituents in the sense of the Translation Principle. This assumption is incorporated, e.g., in Jackendoff's

[82]Thomason (1976), Dowty (1978), Bach (1979).

[83]Keenan (this volume) gives a substantial criticism of a purely lexical rule of passive, taking only basic, i.e. lexical predicates as its input.

[84]This is a gross simplification of Gazdar's (1979a:28) passive metarule.

[85]Chomsky (1965:99)

(1977) base theory, where the X^1 level is supposed to contain "semantic arguments", as opposed to the higher levels where restrictive and non-restrictive modifiers are generated.[86]

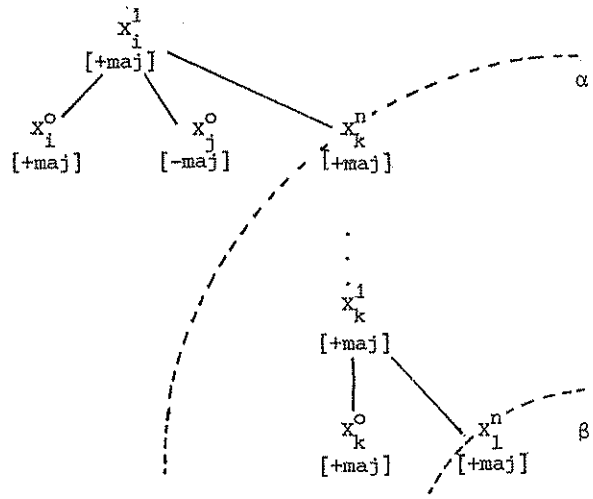
The semantic constraint on the form of subcategorization laid down in the Translation Principle is only a first step towards restricting possible frames. The formulation of the principle refers to subcategorized *phrases* (i.e. constituents of type X^n) and their role with regard to the translation into the language of semantic representation. Apart from the semantically relevant phrases, an insertion condition for a lexical item will have to refer also to details of the syntactic fine-structure, i.e. to purely syntactic information such as the presence of specified formatives (particles, complementizers, governed prepositions etc.), or to the presence of morphological features on certain subcategorized items (case features on governed NP's, inflectional form of an embedded V under a higher matrix V, etc.). It is sometimes suggested in the literature that the form of the subcategorization frame of a head would be fully predictable on the basis of its semantic representation, and hence superfluous.[87] We would claim that the sensitivity of a head for the types of syntactic detail mentioned above pleads for the recognition of an independent level of syntactic subcategorization, governed but not completely determined by the translation principle.

It is clear that Chomsky's Strict Locality Principle has little to say with respect to the syntactic fine-structure: it only proposes *horizontal* restrictions on the structures that can be factorized by a subcategorization frame. Constraints on the *depth*, i.e. the vertical dimension of a frame, will have to be added to the Strict Locality Principle. Consider

[86]There may be doubts as to whether the $\bar{X}$ level contains all *and only* the semantically relevant arguments, in a base-generated syntax. The subcategorization domain of a V can be invaded by adverbial phrases (including sentence adverbs) which, when generated in their surface position, cannot be interpreted as arguments of the head in whose subcategorization domain they occur (they rather have this whole phrase as an argument themselves). For a solution concerning one type of adverbial (instrumental phrases), see Bresnan (forthcoming b): it is suggested that instrumentalization is a lexical process, changing a *n*-place predicate into a *n*+1 place predicate. If the instrumental phrase is added to the subcategorization frame by lexical rule, it can subsequently be affected by other lexical rules (e.g. compound formation: *snow-covered* cf. Roeper & Siegel (1978)).

[87]Cf. Lapointe (this volume).

the following tree structure:



The tree drawn above is the output of a base rule in accordance with Jackendoff's theory. It complies with Jackendoff's restriction that categories outside the main projection line are either maximal phrases projected from [+maj] heads, or X^0 constituents with the feature [-maj].[88] The question addressed here is the following: given that this structure is factorized by the structural description for the insertion of the encircled head-of phrase X^0, what will be the maximal depth for this insertion condition ? It appears that the structural description for the insertion of the encircled X^0 can maximally descend to the head of a subcategorized phrase: we will refer to this restriction on factorization as the *Head Constraint*. An insertion condition will have to exhibit this maximal type of depth in the case of, for example, selection of verbal morphology by a matrix verb taking a VP complement: the features accounting for the verbal morphology of the complement head are classified as morpholexical in the theory adopted here, i.e. they would only be present at the X^0 level of the complement head and not be projected to the X^n level induced by this head.

[88] Cf. Jackendoff (1977: 36). For the [±major] distinction and its role in morphology, see p.22.

Lapointe[89] rightly associates the Head Constraint with the primitive notions of "cyclic domain" (where all major X^n phrases count as cyclic) and the derived property of Subjacency. When a frame could analyze the complement structure of a subcategorized phrase, it would have to pass two cyclic boundaries as the tree structure above shows. We already refered to the structurally parallel restriction on factorization in morphology: it appears, then, that the notions of cyclic domain and subjacency represent a purely structural type of limitation on the maximal computational space of rules in the grammar as a whole, the cyclic parameters being set to the characteristic values for the specific components.

The theory of subcategorization outlined so far has a certain amount of predictive power: it prevents certain types of operations from being formulated as lexical rules, thereby limiting the class of possible lexical rules. We can illustrate this with an example borrowed from Dowty (1978). Given a verb of the *persuade* type, subcategorized as [V NP $\overline{VP}$], it would be impossible to write a lexical rule to the effect that it would interchange the *object* of the complement $\overline{VP}$ with the object or subject of the main verb, so that, e.g. *John ! persuaded Bill kiss Mary* would mean the same as *John persuaded Mary to kiss Bill*. This type of operation is ruled out because the constituent make-up of the subcategorized $\overline{VP}$ is not defined by the contextual frame of a *persuade*-type verb. Lexical rules affecting a V of the type [V NP $\overline{VP}$] do not have information at their disposal as to the internal structure of the VP; therefore, no constituent internal to that phrase can be affected. In this way, the theory of lexical rules adopted here accounts for a well-known asymmetry in the accessibility of subjects vs. non-subjects of subcategorized tenseless complements.[90]

[89]Lapointe (1978: 17)

[90]Case marking PP's (*by* phrase, *to* phrase) and governed PP's (*approve of NP, depend on NP to VP)* seem problematic in this respect: the NP's in these phrases freely participate in lexical rules applying to V (e.g. pseudo-passives), and control can be effected from out of the PP's. In Moortgat (1980) it is suggested that because of their semantic emptiness these P's do not count as heads in the sense of the Translation Princi-

It will be clear from the foregoing that the syntactic part of lexical rules obeys different restrictions than the corresponding rules in a transformational component: lexical rules are both weaker and stronger than transformations under the standard restrictions (cf. Lasnik & Kupin (1976)). They are stronger in that the SD part of a lexical rule allows for *richer expressive devices* than those tolerated in a transformational component: in order to be able to give sufficient information concerning the insertion conditions for a lexical item, there is empirical need for the inclusion of more than the three terms (two affected terms, one catalyst) allowed by the Lasnik & Kupin formalization.[91] The use of richer expressive devices is counterbalanced by the more *limited scope*, as far as the SD part of lexical rules is concerned (i.e. their syntactic domain does not extend beyond the X^1 level), and by the extremely trivialized possibilities with regard to the structural changes permitted (i.e. insertion of a terminal element, or mapping of frames with structure preserving effect).

2.3.3. *Lexical Properties*. In the preceding paragraphs, we sketched a general strategy to limit the power of lexical rules: lexical rules are given access to the syntactic information included in the insertion condition of the items they relate, and to nothing more; the information contained in the syntactic frame of a lexical item has to be entirely motivated in terms of the contextual information necessary for the process of lexical insertion. With regard to the demarcation problem lexicon-syntax, the essence of the lexical approach appeared to be the following: old-style NP-Movement processes are reinterpreted in terms of properties of individual lexical items related by lexical rules, rather than as transformational mappings between the entire phrase markers in which these items can occur. Let us now discuss the consequences of this shift from

(note 90 cont'd)
ple. Consequently, they are treated as [-major] phrase introducers. On the level of semantic representation, the NP's in minor PP's are direct arguments of the verbs that subcategorize these minor PP's; this accounts for their behaviour with respect to lexical rules.

[91]Notice that the property of string-adjacency is not expressible in the Lasnik & Kupin framework. Within this framework, rules requiring string-adjacency (such as Dative and Passive) could not be formulated as transformations in the first place.

entire phrase markers to the more limited syntactic scope characteristic of the lexicon. The question we want to answer is: what properties of lexical rules are derivable from the organization of the lexicon sketched above ? These lexical properties have been listed in various works within the lexicalist tradition[92]; we will particularly stress the fact that these properties have the status of *theorems* in the theory sketched here: they follow from the way the lexicon is organized, and need not be explicitly postulated.

Within Jackendoff's full entry theory of the lexicon, all the information concerning individual items is stored; the evaluation measure for the lexicon is characterized in terms of independent information content; the redundant information is factored out of the independent information content of the lexicon by reference to lexical rules. This theory allows for a natural account of *lexical governance*: given a lexical function, say Passive or *-able*, and a class of suitable inputs for this function, the fact that there is no output for a subclass of the input is captured by there simply being no reference to the redundancy function, since there is no predictable information to be factored out. In this way, lexical rules can give expression to the property of *rule governance* characteristic of processes such as Passive, Dative, Causative, etc. Transformations are not suited to express the governed nature of these processes: they express fully regular, exceptionless relationships. Treating phenomena like those mentioned above by transformation, then, inevitably results in the introduction of "exception features" to code the behaviour of particular lexical items with regard to allegedly regular transformations. The desire to obviate an excessively powerful device such as the exception feature[93] has as a matter of fact led to the general abandonment of a transformational treatment for Dative, Causative, and at least a subpart of Passive, which were all formerly derived by means of the core rule Move NP. The history of NP Movement shows that the actual range of this "very general" rule is effectively limited now to the "syntactic"

[92]E.g. Bresnan (1976), Wasow (1977), Roeper & Siegel (1977), etc.

[93]In the Lasnik & Kupin (1976) framework, reference to exception features is excluded.

part of Passive and to subject raising.[94] We will return to this issue in the next section.

It has often been remarked that the grammar is characterized by principles of *locality* which limit the "computational space" of a given rule to a well-defined small domain. As far as the type of locality associated with Move NP in a transformational theory is concerned, this can be interpreted as the direct result of the fact that lexical rules relate entries, and that the syntactic information included in an entry must be characterizable in the form of a finitely specifiable subcategorization frame.[95]

Note furthermore that lexical rules, as mappings between subcategorization frames, can aptly be described as *relation changing rules* (cf. Dowty (1978)). This follows from the Translation Principle governing the relation between a contextual frame and the associated translation. The subcategorized phrases in a frame play the role of arguments to the head; lexical rules, then, can only be operations manipulating the way in which subcategorized syntactic phrases map into the semantic argument positions (e.g. suppressing syntactic expression of a logically necessary argument, inverting the order of syntactic constituents vis-à-vis logical arguments, etc.). Long distance processes, generated by means of Move Wh in a transformational grammar, are not re-interpretable in terms of properties of individual lexical items; consequently, they do not show this function dependency: they apply blindly, and are not restricted to subcategorized phrases, let alone to specific functions among the subcategorized phrases (object, ...).

Finally, input and output of a lexical rule are both entries which must be inserted in structures resulting from the application of the base rules, i.e. they are both contextual features satisfied by some base-generated structure. The property of *structure preservation* is, therefore, a necessary characteristic of lexical rules. There is no need for a general constraint on the power of the transformational component to this effect.

[94]Cf. Chomsky (1979b:160-162) for a discussion of his view on the difference between the "lexical" and the "transformational" passive.

[95]Within the framework of lexical grammar, then, the local bounded nature of processes accounted for in terms of Move NP within transformational theories does not require the addition of postulates such as Koster's (1978) Locality Principle or Bounding Condition to the grammar.

In the next section we will compare the deductive structure of the lexical theory with competing EST versions.

2.3.4. *The Movement Alternative*. Within the framework of Trace Theory, phrase structure rules and lexicalization are considered to be optional: a category α which is not expanded into terminal elements results in an *empty node* $[_{\alpha}e]$. Base generated empty nodes function as landing sites for the structure preserving substitution rules, i.e. the cyclic movement transformations generalized to the core rule Move α. Movement of a category α leaves behind an empty node $[_{\alpha}e]$; the overgenerating effect of the rule Move α is filtered but by the interpretation of traces as bound anaphors that have to stand in the proper antecedent-anaphor configuration with respect to the moved category.[96]

The identification of NP Movement and Wh Movement as particular instances of one rule Move α, which is not construction specific, is excluded on principled grounds within Lexical Grammar. Only the effect of NP Movement rules can be captured in terms of local, bounded subcategorization properties of lexical items: Wh Movement, even in its most simple application, exceeds the domain of lexical rules by analyzing the COMP. We can expect, then, that within the approach that collapses these distinct processes, measures will have to be taken to account for their differences. First, the possibility of reducing the SD part of transformations to the general format "Move α" crucially depends on the requirement of structure preservation. Notice that there is a difference in the status of the landing sites of NP movements and Wh movements. Emonds (1976:5) makes clear that in order for the structure preserving hypothesis to count as an explanatory concept, restrictions on the generation of empty nodes are needed so as to allow for empty nodes only in positions that can be motivated independently from the movement operation at hand. The fact that NP movements are relation changing is therefore predicted as their landing sites are, in accordance with this requirement, always argument positions. It will be clear that the phrase structure expansion in COMP does not obey this restriction on the distribution of base generated empty nodes. The structure-preservingness of Wh Movement is *stipulated* by

[96]For the other type of base generated empty node, PRO, cf. p. 34.

assuming an XP landing site in COMP. Second, the core rule "Move α" is considered to be unbounded. The bounded effect of NP movements is analyzed as the result of a number of general conditions on rule application as proposed in Chomsky (1973).[97] The unbounded effect of processes analyzed as involving the rule "Move Wh" in Chomsky (1977a) again requires the postulation of auxiliary hypotheses in the form of *unless*-statements attached to these conditions. Attributing the conditions to Universal Grammar, i.e. to the biological endowment of the child, does not alter the methodological fact that, within Trace Theory, the conditions accounting for the boundedness of NP Movement have the status of postulates, whereas within lexical grammar the local nature of lexical rules follows as a theorem from the finite contextual information associated with lexical items.

Instead of concentrating on the differences between NP Movement and Wh Movement with respect to their application, we will focus on the differences concerning their results: within Trace Theory, the empty nodes left by NP Movement and Wh Movement are in principle undistinguishable; within lexical grammar, NP Movement processes count as lexical rules, which leave no traces behind. The question is, then, whether NP Movement traces play any role in Trace Theory, and, if so, whether they are necessary for this function.

Empty nodes in surface structure fulfill three distinct functions. First, on the *phonological* level, they serve as blocking devices for the English rule of complementizer contraction. Second, on the level of *logical form*, they are interpreted as bound variables. Third, on the level of *shallow structure*, they serve to identify the underlying thematic function of displaced constituents directly from surface structure.

With respect to the contraction facts, it is well known that only Wh Movement traces function as blocking elements: the empty nodes resulting from NP Movement (in the raising case *ought*) nor the empty PRO subjects (in a control case like *want*) prevent contraction to take place.[98] The

[97]These conditions are reformulated in terms of conditions on the distribution of empty nodes in Chomsky (1978) and as conditions on case marking and binding in Chomsky (1979a).

[98]Cf. Postal & Pullum (1978).

observable phonological effect in the contraction cases is the typical kind of strong evidence for the underlying reality of an abstract construct such as the empty node. The absence of this phonological effect in NP Movement and control cases points to a problem for the transformational theory, which postulates abstract syntactic elements in these two cases as well as in the Wh Movement case. Within lexical grammar, the asymmetry between Wh Movement traces and other empty nodes falls out naturally; the syntactic structures of NP Movement and control sentences contain no abstract empty nodes to block contraction. Compare the following underlying structures.

(17) a. I wanna become president

$[_S I_i$ want $[_{\overline{S}}\ [_{NP} e_i]\ [_{VP}$ to become president]]]

(Trace Theory; e = PRO)

$[_S$ I want $[_{\overline{VP}}$ to become president]]

(Lexical Grammar)

b. John oughta know better

$[_S$ John$_i$ ought $[_{\overline{S}}\ [_{NP} e_i]\ [_{VP}$ to know better]]]

(Trace Theory; e = trace left by Move NP)

$[_S$ John ought $[_{\overline{VP}}$ to know better]]

(Lexical Grammar)

Recent elaborations of Trace Theory have exposed another area where the traces left by Wh Movement differ from the other empty nodes. On the level of logical form, only Wh Movement traces are translated as variables bound by the displaced *wh*-constituent which acts like a quantifier.[99] NP Movement traces and PRO are not treated as bound variables; they play no role at the level of logical form, then.

It appears that the traces left by NP Movement are only used at the level of shallow structure for the identification of underlying relations. If this is their sole function, there is no motivation to allow them in the theory, since mechanisms to identify underlying relations *without*

[99]This only concerns traces left behind on the original position of the Wh-constituent. Other traces left behind by successive application in intermediate complementizers are not spelled out in logical form.

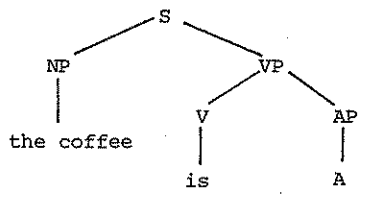
the intervention of traces coding derivational history are already needed on independent grounds in the lexicon. Consider again the *able* rule and its effect on the syntactic and semantic level.

(18) a. ABLE(i)=j for

<i, $[\Delta]_V$,[V NP], V'(NP') >

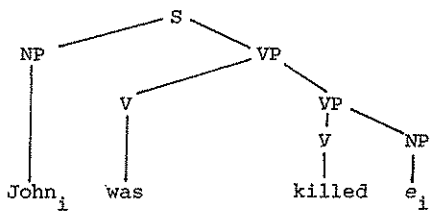
<j, $[[\Delta]_V able]_A$,[A], $\lambda x \Diamond \exists y[V'(x)(y)]$ >

b.



drinkable

On the subcategorizational level, the *able* rule transformes a transitive VP into an intransitive AP; the associated operation on the level of semantic representation performs lambda-abstraction over the property-denotation of the AP to the effect that when this AP is combined with a grammatical subject-NP, the NP is interpreted as the logical *object* of the verb corresponding to the *able* adjective. The logical subject of this verb remains unexpressed syntactically. Apart from the specific meaning of the *able* morpheme, the manipulation of functions, and the identification of underlying functional status directly from surface structure in the *able* case is identical to the necessary manipulations in the case of passive (grammatical subject interpreted as logical object, logical subject unexpressed). Yet, in this latter case, Trace Theory would make use of an empty node, coding the transformational history.



From the foregoing discussion it appears that undisputably lexical rules, such as the *able* rule, independently need the power to identify logical functions without the use of traces. This obviates the only kind of motivation we were able to find within Trace Theory for NP Movement traces. Notice again that Occam's Razor can only be applied to the *local* traces: Wh Movement phenomena still need an abstract gap of some type for the identification of the logical function of displaced constituents, since the rule goes beyond the domain of traceless lexical manipulation.

The fact that traces left by NP Movement play no role whatsoever in the grammar indicates that there is no empirical motivation for the full range of traces allowed by the theory of freely generated empty nodes. The question arises, then, of how to limit the generation of empty nodes to the desired proportions, i.e. to the traces associated with long-distance processes.[100] Interpretive treatments of Wh Movement, such as Koster's (1978), are crucially based on the *full* range of freely generated empty nodes, which as we stated above, does not exclude local structure preserving movements in principle.

The only theory which properly recognizes the apart status of long-distance gaps, without having recourse to an overgenerating theory of empty nodes, is Gazdar's (1979b): antecedents to long-distance gaps (Wh-phrases, topicalized constituents, etc.) are introduced by PSR's, concatenated with a "derived" constituent characterized by means of a feature as containing a gap of the appropriate type.[101]

The foregoing comparison of the trace-theoretic framework with the approach of lexical grammar is specifically meant to forestall tedious debates over notational variants. We wanted to stress the fact that trans-

[100]One could think of allowing empty nodes only in the expansion of COMP, thereby creating the necessary landing sites for Wh movement. This approach would ensure that only at the original position of Wh-constituents an empty node would occur. However, only the empty nodes would be used that are not in accordance with the requirement discussed above. Moreover, once the role of the transformational component has been reduced to deal with long distance processes only, it seems unwarranted to invoke the transformational mechanism, especially if alternatives are available (cf. Bresnan's (1978) position, which used this mechanistic transformational approach, and the abandonment of this position in more recent work).

[101]Cf. Gazdar (1979b) and this volume for an exposition of the theory of derived nodes.

fering relation dependent and lexically governed phenomena to the lexicon provides a methodologically superior position in that properties of the phenomena which require the postulation of primitive concepts within trace theory follow as theorems from the organization of the grammar according to the thematic commitments of lexical grammar. Moreover, the power of the independently needed device of lexical rules does not have to be enlarged in any sense for the treatment of these phenomena.