UNIVERSITY OF CALIFORNIA

Los Angeles

Both Lexicons

A dissertation submitted in partial satisfaction of the

requirements for the degree Doctor of Philosophy

in Linguistics

by

Chris Golston

1991

The dissertation of Chris Golston is approved.

Jaan Puhvel

Donka Minkova

Bruce Hayes

Donca Steriade, Committee Co-Chair

Victoria A. Fromkin, Committee Co-Chair

University of California, Los Angeles
1991

...

for Alec Michael Chan-Golston

TABLE OF CONTENTS

ACKNOWLEDGMENTS

I would like to thank the UCLA Department of Linguistics for the five years of generous financial support I received as a graduate student.

I gratefully acknowledge the support of my doctoral committee: my co-chairs Vicki Fromkin and Donca Steriade, for their invaluable expertise, care and encouragement; Bruce Hayes, for introducing me to prosodic phonology and keeping me from going down many a useless path; Donka Minkova, for her help with English, her hospitality and her patience; Jaan Puhvel, for his aid with Greek and Latin as well as with the Hittite that never appeared in the final version of this dissertation. I could not have had a better committee and hope the final product here does not disappoint these exceptional scholars.

I would especially like to thank the (unofficial) sixth member of my doctoral commitee, Tom Cornell. I cannot imagine what this thesis would have looked like if it had not been for his many timely suggestions (e.g., introducing me to the work of Merrill Garrett and Joe Emonds) and thoughtful critiques.

I won't try to list all the faculty, staff and fellow students in the UCLA Linguistics department for their friendship and support. Sue Banner Inouye and Yasushi Zenno stand out especially as other members of the *DSG*.

Special thanks to Cheryl Chan, for going over the data and arguments and helping me sort things out. Without her insight, support and understanding I could not function, much less write a dissertation.

VITA

February 3, 1961	Born, Fort Hood, Texas
1983	B.A., Philosophy and Classical Languages University of California, Berkeley
1983-1984	Teacher, Kindergarten and Grade 1 San Francisco Chinese Parents Committee School San Franciscio, California
1984-1986	Teacher, Grades 9-12 Maybeck High School Berkeley, California
1986-1988	M.A., Linguistics University of California, Los Angeles
1987-1991	Teaching Assistant Department of Linguistics University of California, Los Angeles

PUBLICATIONS AND PRESENTATIONS

Shaul, D. L., R. Albert, C. Golston and R. Satory (1987). "The Hopi Coyote Story as Narrative: the Problem of Evaluation", *Journal of Pragmatics*, 11, 3-25.

Golston, C., (1989a). "Clitics in Homeric Greek: less Evidence that PIE was Head-Final". Proceedings of the Fifteenth Annual Meeting of the Berkeley Linguistics Society.

Golston, C., (1989b). "What a Phrasal Affix Looks Like". Proceedings of the Western Conference On Linguistics, Vol. 2.

Golston, C., (1989c). "Floating High Tones in Ancient Greek". Proceedings of the Arizona Phonology Conference, Vol. 3.

Golston, C., (1990a). "Level-Ordered Lexical Insertion". Proceedings of the Western Conference on Linguistics, Vol. 3.

Golston, C., (1990b). "Minimal Word, Minimal Affix". Proceedings of the 21st North-Eastern Linguistics Society Meeting.

Golston, C., (1991a)."Level-Ordered Lexical Insertion in Ancient Greek". Paper presented at the Meeting of the Linguistics Society of America, 1991.

Golston, C., (1991b). "Level-Ordered Lexical Insertion: Evidence from Speech Errors". Proceedings of the Arizona Phonology Conference, Vol. 4.

ABSTRACT OF THE DISSERTATION

Both Lexicons

by

Chris Golston

Doctor of Philosophy in Linguistics

University of California, Los Angeles, 1991

Professors Victoria A. Fromkin and Donca Steriade, Co-Chairs

The dissertation proposes that the words and affixes of a language are stored and processed in two lexicons: the "Lexicon" contains content words and derivational affixes, the "Phrasicon" contains function words and inflectional affixes. The insertion of lexical material in grammatical and on-line processing is taken to be a two-stage, level-ordered process. The first stage is insertion from the Lexicon, the second is insertion from the Phrasicon.

Chapter 1 reviews grammatical and production models and how they model lexical storage and processing. A grammatical model and a production model are proposed that incorporate two lexicons, as well as level-ordered lexical insertion.

Chapter 2 draws on minimal prosodic weight requirements imposed on content words and derivational affixes in English, Ancient Greek and Latin. Function words and inflectional affixes are shown not to be subject to these requirements, supporting the

claim that they are stored in a different component of the grammar in which these weight requirements do not hold.

Chapter 3 investigates phonological aspects of level-ordered lexical insertion. Prosodic constituents above the word are claimed to be formed *after* the first stage of lexical insertion but *before* the second. Analyses of phrasal stress and of reduced forms of function words are given for English, Ancient Greek and Latin.

Chapter 4 concerns word-formation processes. It is argued that types of affixation, compounding, and other word-formation processes apply either in the Lexicon or in the Phrasicon, but not in both.

Chapter 5 presents support for level-ordered lexical insertion from the study of speech errors. Those types of errors that commonly occur, as well as those that do not, support the claim that lexical insertion occurs in two distinct stages.

Chapter 6 is concerned with different types of aphasia. Aphasic patients show differential impairments that affect content words and derivational affixes but spare function words and inflectional affixes or vice versa, supporting the claim that there are two lexicons.

1. Introduction

How many lexicons does a grammar have? And how many lexicons does a speaker use in producing an utterance? The usual assumption is one. In this dissertation, however, I will present evidence that points to the existence of *two* lexicons in both formal grammars of natural languages and in formal models of speech production. The first lexicon, for which I use the term "Lexicon" is to contain content words (lexical items, open class words) and derivational affixes such as *-ity* and *-ness*. The second lexicon, the "Phrasicon", is to contain function words (non-lexical items, closed class words) and inflectional affixes such as plural *-s* and past tense *-ed*. This hypothesis about the modular organization of lexical storage I will call *The 2 Lexicon hypothesis.*

A second hypothesis pursued here is that the selection and insertion of the phonological material that consitutes words and affixes is level-ordered. That is, the phonological forms of words and affixes are not inserted all at once in speech production or in the grammatical derivation of a sentence, but in two distinct and separate stages, each of which corresponds to one of the two lexicons mentioned above. In the first stage of lexical insertion, the phonological forms of content words and derivational affixes are inserted into the syntactic and semantic representation of an utterance; only at a later stage are function words and inflectional affixes inserted. Thus the first stage of lexical insertion inserts phonological forms stroed in the "Lexicon", whereas the second stage of lexical insertion inserts phonological forms stored in the "Phrasicon". This hypothesis I will call *Level-Ordered Lexical Insertion.*

Evidence for these two hypotheses will be drawn from two major areas. First, grammatical evidence will be presented for the 2 Lexicon Hypothesis (Chapters 2 and 4) and for Level-Ordered Lexical Insertion (Chapter 3). Second, speech error and aphasic

evidence will be presented for Level-Ordered Lexical Insetion (Chapter 5) and the 2 Lexicon Hypothesis (Chapter 6).

This chapter will serve as a review of some of the relevant literature concerning lexical representation and storage in grammatical and production models. The grammatical models (1.1.1) range from those having no lexicon at all to those having one or two lexicons. I propose a grammatical model of my own which incorporates the 2 Lexicon Hypothesis and Level-Ordered Lexical Insertion. The production models (1.1.2), include those with one lexicon and those with two. Again, I propose a production model which incorporates the 2 Lexicon Hypothesis and Level-Ordered Lexical Insertion. The resulting picture is a single model which can serve both as the outline of a grammatical model and as a the outline of a production model. Finally (1.2), I review some terminology central to the discussion: content words, function words, derivation and inflection. After reviewing some earlier proposals as to the nature of these items, I tentatively sketch a new way of characterizing these types of words and affixes.

This review will thus serve three purposes. First, it will provide a discussion of some of the issues involved in modeling lexical storage and access. Second, it will provide a review of the major proposals extant for modeling the lexicon. Third, it will begin to articulate a model of lexical storage and access that can serve *both* in a grammar *and* in a production model.

1.1 How many lexicons?

Before going further, I should state what it is that a lexicon is meant to be. Emmorey & Fromkin's (1988) definition seems well-suited to both grammatical models and production models:

> The mental lexicon is that component of the grammar that contains all the information--phonological, morphological, semantic, and syntactic--that speakers know about individual words and/or morphemes.

That is what a lexicon is. What remains to be discussed is i) How many lexicons does a speaker have? and ii) How is it (are they) organized?

1.1.1 Grammatical Models[1]

Zero Lexicons

Early generative conceptions of the grammar (e.g., Chomsky 1957; Lees 1960) had nothing we would today call a lexicon: all word-formation, both affixation and compounding, was accomplished by essentially syntactic transformations. Lexical items were inserted into kernel-sentences by rules; crucially, the rules that inserted lexical items were of the same type as phrase structure rules--rewrite rules of the form *a* -> *b*, 'a is rewritten as b':

Phrase Structure Rules	S -> NP + VP
	NP -> T + N
	VP -> V + NP
	V -> Aux + V
Lexical Rules	T -> the, a
	N -> man, ball, etc.
	V -> hit, took, etc.
	Aux -> (have + en) (be + ing) (be + en)

Chomsky later argued against such lexical rules on grounds of simplicity: since many morphological properties (declensional classes, strong or weak verbs, etc.) "are entirely irrelevant to the functioning of the rules of the base and are, furthermore, highly idiosyncratic, the grammar can be significantly simplified if they are excluded from the rewriting rules and listed in lexical entries, where they most naturally belong" (1965:87).

---

[1] In this section I am indebted to Hammond and Noonan's (1989) overview of generative morphology.

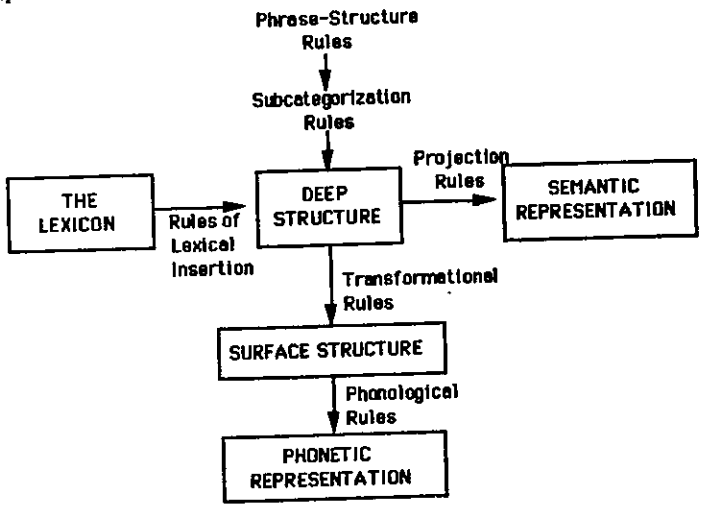
One Lexicon

With the publication of *Aspects of the Theory of Syntax*, the grammar was to contain

> no rules...that introduce the formatives belonging to lexical categories. Instead, the bac of the grammar will contain a *lexicon*, which is simply an unordered list of all lexical formatives. More precisely, the lexicon is a set of *lexical entries*, each lexical entry being a pair (D, C), where D is a phonological distinctive feature matrix "spelling" a certain lexical formative and C is a collection of specified syntactic features (a complex symbol). (Chomsky 1965:84)

A number of features of this 'lexicon' stand out. First, it was taken to be part of the syntactic *base component*, rather than a separate component (but see Newmeyer's 1986 characterizaion below). Second, it was not a word-formation component; affixation and compounding still were done in the syntax. Third, it had no structure: it was taken to be an *unordered list of lexical formatives*. Fourth, it was the repository of irregularity: "In general, all properties of a formative that are essentially idiosyncratic will be specified in the lexicon" (1965:87). Finally, lexical insertion *preceded* transformational rules. Newmeyer (1986:74) illustrates the *Aspects* model as follows:

(1) *Aspects* model



It was this rather restricted conception of the lexicon that was taken up in the *Sound Pattern of English* (Chomsky & Halle 1968). At this point, it was still possible to give an overview of the grammar of a language without mentioning the lexicon:

> [A] grammar contains a syntactic component which is a finite system of rules generating an infinite number of syntactic descriptions of sentences. Each such syntactic description contains a deep structure and a surface structure that is partially determined by the deep structure that underlies it. The semantic component of the grammar is a system of rules that assigns a semantic interpretation to each syntactic description, making essential reference to the deep structure and possibly taking into account certain aspects of surface structure as well. The phonological component of the grammar assigns a phonetic interpretation to the syntactic description, as well as the associated semantic and phonetic representations. (Chomsky & Halle 1968:6-7)

The grammar took on a tripartite organization into syntactic, semantic and phonological components. The lexicon was still merely part of the syntactic component: an unordered

list of formatives that contained no rules of word-formation. Affixation and compounding were still done in the syntax.

By the beginning of the 1970s, problems began to emerge with doing word-formation in the syntax. Problems arose both with compounding and with affixation. Syntactic derivations of compounds (e.g., Lees 1960) violated a number of otherwise valid syntactic principles such as recoverability. And the syntactic derivations of certain nominals such as *revolution* and *referral* failed to capture idiosyncracies in their meanings (Chomsky 1970). As Hammond & Noonan point out "the problem is much more general than Chomsky indicated.... Just to take one example, the fact that there now exist *white blackboards, green blackboards,* and so on clearly shows that the meanings of compounds like blackboard cannot be derived from the meanings of their component parts" (1988:4).

Work by Jackendoff (1972) and Halle (1973) led to the creation of a word-formation component within the grammar. Halle's proposal might be diagrammed as follows:

(2) *Prolegomena*-type lexicon

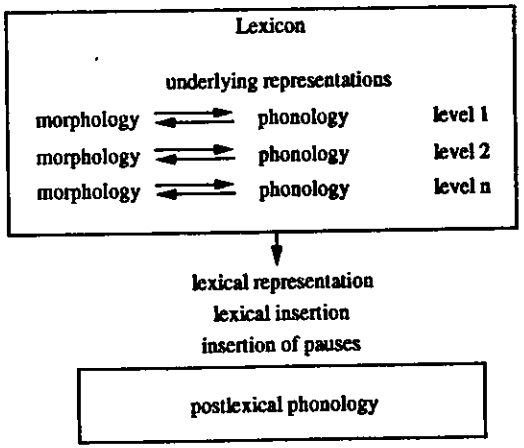
Lexicon
Dictionary: a, be, man, manly, the, transmit...
List of morphemes: a, be, man, -mit, -ly, the, trans-...
Word-formation rules
Filter

This word-formation component consisted of a *dictionary* contained all occurring words (simple or complex) in the language; a separate *list of morphemes* of the language; a set of *rules* for creating words from those morphemes; and a filter that supplies idiosyncratic information about specific words.

Aronoff (1976) modified this model by omitting bound morphemes such as *-mit* (*transmit, omit*) and thus collapsing the dictionary and the list of morphemes. Morphology was to be word-based, not morpheme-based. He also restricted word-formation rules from accessing syntactic or phonological information, yielding a truly *modular* and essentially *autonomous* lexicon.

This notion of a modular and autonomous lexicon was developed into a highly articulated model of the lexicon in Lexical Phonology (Mohanan 1982, Kiparsky 1982, 1986). Work by Siegel (1974) and Pesetsky (1979) had established the notion of levels of morphology with concomittant levels of phonology. Kiparsky and Mohanan developed a stratified and highly organized lexicon that greatly restricted possible morphological and phonological operations. The model looked something like this (from Kaisse & Shaw 1985:9):

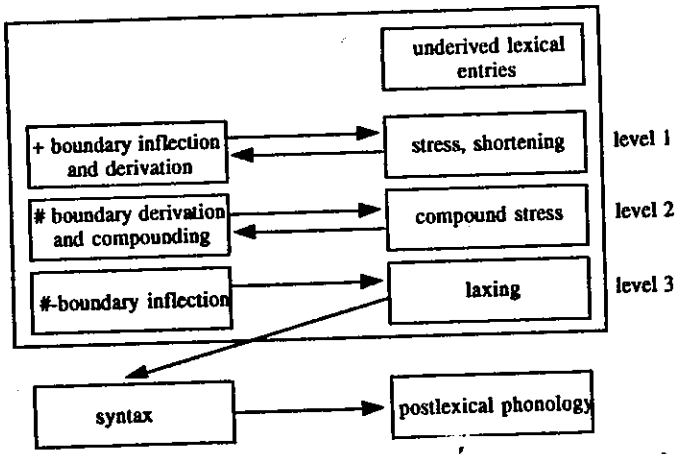
(3) Lexical Phonology



Lexical Phonology (LP) is an elaborate morphological and phonological theory and I will not attempt to discuss it thoroughly here (see Kaisse & Shaw 1985 for a short review and

references). An interesting aspect of LP, for our purposes, are the levels of affixation and compounding. Consider Kiparsky's (1982) model for English:

(4) Lexical Phonology, words and affixes



'+ boundary' inflection includes irregular forms like *teeth, children, ran, and went*; + boundary derivation includes affixes that attract stress such as *-al* (*philósophy, philosóphical*) and *-ous* (*advántage, advantágeous*). '# boundary' derivation includes stress-neutral affixes such as *-ness* (*advantágeous, advantágeousness*) and *-ness* (*philosóphical, philosóphicalness*). '# boundary' inflection consists of regular inflectional affixes like *-s* (*dogs*) and *-ed* (*baked*), all of which are stress neutral. Part of the appeal of LP is that the SPE symbols "+", "#" are replaced by rule domains defined in terms of levels.

The model in (4) explicitly partitions affixes into different levels, with (regular) inflection on its own level after the derivational levels. Justification for this comes from the fact that inflectional affixes in English always 'come outside of' any derivational affixes. Thus plural *-s* may not occur 'within' a derivational affix like *-er*: cf. *cook-er-s* vs. **cook-s-er*. Inflectional morphology also may not occur 'within' a compound: cf. *head-hunter-s* vs. **head-s-hunter*. Level-ordering rules of affixation guarantees that some classes of affixes (e.g., regular inflection) will always occur outside of others (e.g., derivational affixes).

Most work in LP is not clear about where function words such as *if, the,* and *should* are stored in a model such as (4). One would assume that they would go in the top box in (4) with other underived lexical entries like *dog* and *eat*. But as Kaisse & Shaw (1985:9) note,

> From various remarks of Kiparsky and Mohanan, we gather that the store of underlying lexical representations, at least for English, contains only roots belonging to the major categories N, V, and A. The representations of affixes are contained with the word formation rules (WFRs) of the morphology [on levels 1, 2 and 3 above--C.G.], while the representations for words belonging to minor categories are stored elsewhere, and thus do not receive lexically assigned stress, only postlexical (sentence) stress.

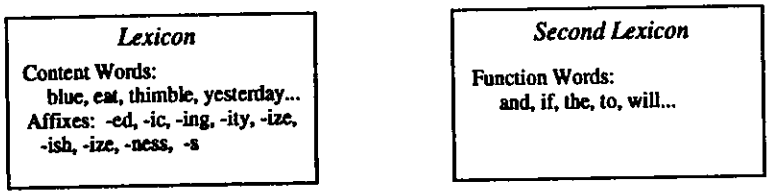
One problem with storing function words with other underived words is that one would then expect them to undergo the same sorts of morphological processes (derivational affixation and compounding) that these other words undergo. (This will be discussed in greater detail in a later chapter.)

LP is the most complete morphological and phonological model of the lexicon to date that uses a single lexicon. Let us now turn to models of lexical storage and word formation that utilize two lexicons.

Two Lexicons

The earliest explicit 2-Lexicon model I am aware of is found in Stockwell, Schachter & Partee (1972) (hereafter *MSSE*). If I understand them correctly, they claim that the phonological forms of words and affixes of English are stored as follows:

(5) *Major Syntactic Stuctures of English*: the Second Lexicon



The phonological forms of content words and affixes are stored in 'the Lexicon', while those of (at least some) function words are stored in 'the Second Lexicon'. The need for two lexicons comes from having two stages of lexical insertion:

> The present grammar utilizes a second lexical inertion procedure which follows the last rule of the transformational component. The function of the second insertion process is to attach phonological matrices to clusters of semantic-syntactic features that have resulted from operations of the transformational component. Such an operation has been made informally many times before. In particular, Fillmore (1966) proposed that pronouns were to be viewed as feature clusters whose phonological realizations were not interestingly related and therefore ought to be inserted following the transformational operations. (1972:793)

Note that function words may have *syntactic* entries in the (first) Lexicon, but that they have no *phonological* entries in that lexicon.

> Deep structure articles, pronouns and prepositions which will later be given their appropriate phonological representation in the Second Lexical Lookup are listed in the first lexicon under identifying labels in lower case letters between quotation marks, e.g. "the", "much/many". These labels are identificatory only since such items have no phonological representation until the Second Lexical Lookup. (1972:730)

An example makes the two stage insertion device clear. Suppose the phonological forms of all words and affixes are inserted prior to transformations:

(6)	Deep Structure	[NP [V NP]VP]S	[NP[V NP]VP]S
	Lexical Lookup	Rex bit Fido	I bit him
	Passive	Fido was bit by Rex	* him was bit by I

Most English pronouns have completely suppletive forms for different cases: e.g., *I, me, mine*. If their phonological forms are inserted before a transformation such as passive, aberrant output like **him bit I* is produced. This is not a problem for proper names like *Tim* or *Joe*, or for noun phrases like *the dog*: *Joe bit the dog, The dog was bit by Joe*. Stockwell *et al.* propose that the phonological forms of (at least certain) function words be inserted after all transformations have taken place:

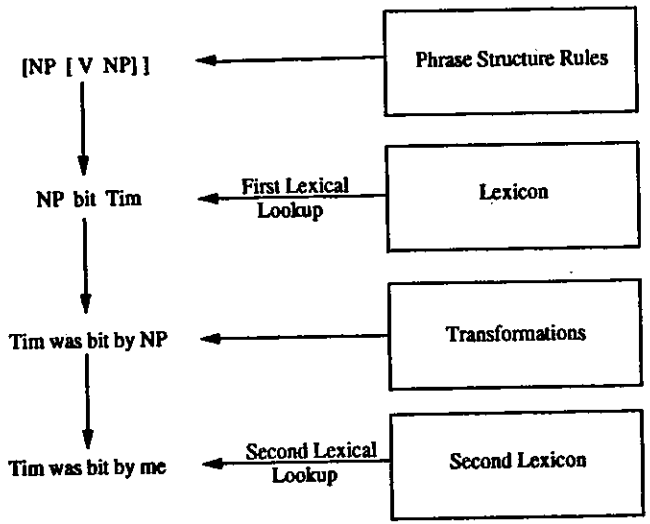
(7)	Deep Structure	[NP [V NP]VP]S	[NP[V NP]VP]S
	First lexical lookup	Rex bit Fido	NP_1 bit NP_2
	Passive	Fido was bit by Rex	NP_2 was bit by NP_1
	Second lexical lookup	----------	He was bit by me

The second lexical lookup is used for the following types of words (1972:797):

(8) Items involved in the second lexical lookup
1. Determiners;
2. Pronouns--both independent and relative
3. Negative adverbials, particles, quantifiers and determiners;
4. Prepositions'
5. Conjunctions;
6. Quantifiers resulting from conjunction reduction

The second lexicon, then, provides the phonological forms for all words involved in the Second Lexical Lookup. Though Stockwell *et al.* do not diagram their grammar, their discussion suggests something like the following:

(9) *Major Syntactic Stuctures of English:* the grammar



On this approach, then, the lexicon contains phonological listings for all content words, all affixes and (perhaps) some function words, and the Second Lexicon contains only function words (at least those listed in (8)).

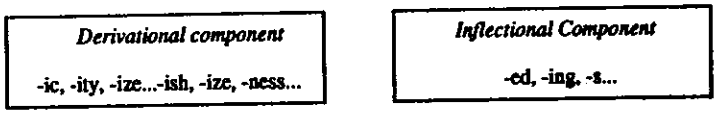
Paul Schachter (personal communication) points out that given the logic of the proposal in *MSSE*, agreement morphology and therefore inflectional morphology in general, has to be in the second lexicon as well. The reason is that agreement has to be post-transformational to account for the possible effect of passivization:

> The girl sees the boys.
> The boys are seen by the girl.
> The boys see the girl.
> The girl is seen by the boys.

I am, however, unable to find any mention of inflectional morphology in this regard in *MSSE*.

A very different approach to lexical storage can be found in Anderson's Extended Word and Paradigm framework (1977, 1982, 1988). Arguing mainly from his claims i) that inflection is 'what is relevant to the syntax' and ii) that (regular) inflection always comes outside of derivation, Anderson proposes a model with what Perlmutter (1988) has called the "split morphology hypothesis": derivational morphology forms one component, inflectional morphology forms another. Anderson's model may be diagrammed as follows:

(10) Extended Word and Paradigm: morphological components



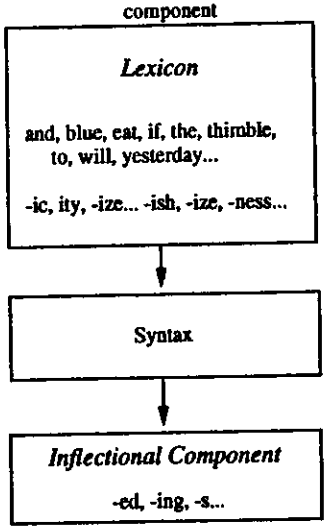
Like the two lexicons in *Major Syntactic Stuctures of English*, Anderson's two morphological components occupy different positions within the grammar. The derivational component feeds the syntax, the inflectional component is fed by it. Anderson (1988:25; cf. also 1982) cites a number of ways in which inflectional morphology is dependent upon the syntax:

> a. CONFIGURATIONAL PROPERTIES: assigned on the basis of the larger structure a word appears in (e.g., case in NP; special forms of verbs in relative clauses)
> b. AGREEMENT PROPERTIES: aspects of the exact form of a word which are determined by reference to the properties of some other word in the same structure.
> c. INHERENT PROPERTIES: properties of a word which must be accessible to whatever rule assigns agreement properties to other words in agreement with it (e.g., gender of noun)
> d. PHRASAL PROPERTIES: properties of phrasal domains which determine the way these domains behave syntactically but which are realized on particular words within the structure (e.g., the effect of tense in defining the scope of binding relations).

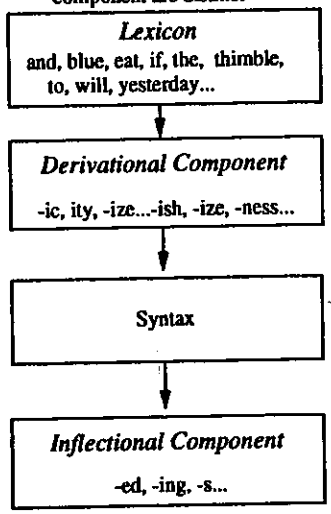
Anderson does not say where words or roots are kept in his model. II assume that they are either stored in the same component (which I'll call the Lexicon) as derivational morphology (11a) or in a component that precedes the derivational compenent (11b):

(11) Extended Word and Paradigm model: possible grammars

a. Lexicon subsumes derivational component



b. Lexicon and derivationl component are distinct

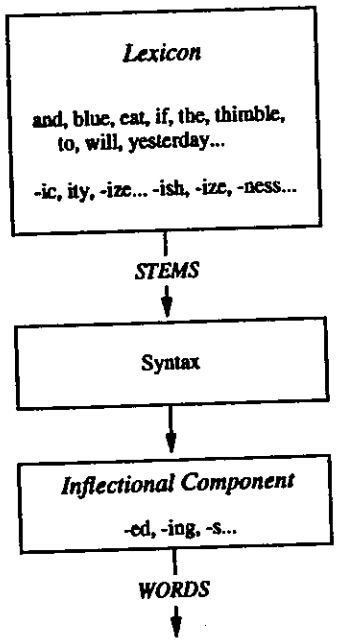


In either case, derivation and inflection are stored in separate components.

Note that Anderson's form of the grammar and his separation of derivation and inflection into two separate morphological components makes sense of his claims that inflection is what is relevant to the syntax: derivation precedes the syntax and thus cannot refer to features (such as agreement) that are only statable in terms of the syntax. Also, the fact that inflection occurs outside of derivation is derived from the position that the derivational and inflectional components occupy in the grammar. This is similar to the level ordering of inflectional affixes in LP, but with one important difference: in LP, nothing forces (regular) inflection to occur on a later level than derivation since neither is defined with respect to any other part of the grammar that could independently determine its level. The Extended Word and Paradigm model, on the other hand, forces inflection to occur at a later level (and thus to occur outside of derivation) precisely because syntax must intervene between derivation and inflection. Anderson derives the fact that inflection occurs outside derivation from transitivity: derivation must precede the syntax (*category, categorical* and *categorize* have different subcategorization requirements, different selectional restrictions, etc.) and inflection must follow it (*categorizes* is only allowed if c-commanded by a 3rd person singular subject)--consequently, inflection must follow derivation.

Anderson notes that in a highly inflected language such as Latin, the output of the derivational component will not be a *word* (since all words in Latin are inflected) but a *stem* (i.e., a word minus its inflectional affixes). He thus modifies Aronoff's claim that morphology is word-based: "it is not words but stems that function as the base of word-formation rules" (1988:28). Using (11a) as a model, the status of stems and words in Extended Word and Paradigm model may be given as follows:

(12) Extended Word and Paradigm model: stems and words



The syntax, then, manipulates stems, not words. As Anderson points out, this is in direct conflict with the usual interpretation of Chomsky's (1970) Lexicalist Hypothesis, which he (Anderson) gives as:

> LEXICALIST HYPOTHESISThe syntax neither manipulates not has access to the internal forms of words.

In the above models (11a, b; 12) I have included both content words and function words in the topmost boxes. As with LP, however, I am not aware of anything in the Extended Word and Paradigm model that dictates this. Where function word (or their stems or their roots) are kept is not discussed; indeed, as I have pointed out already, where words in general are kept is not discussed. For the sake of exposition, however, I

will assume something like (11a) when discussing the Extended Word and Paradigm model.

In a similar vein to the Extended Word and Paradigm model and to the *MSSE* framework, Emonds (1985) has proposed that all inflectional morphology and at least some function words are inserted after movement transformations. This he calls "Late Lexical Insertion" (LLI):

> Late Lexical Insertion: If a morpheme M inserted in a cyclic domain D has a contextual insertion feature that must be satisfied after (rather than before) transformations apply in D, then M is in a closed category. (1985:177)

Emonds' argument for LLI draws on the fact that many FWs and inflectional affixes satisfy insertion contexts which are only produced by transformation. Consider the distribution of the English coordinative morphemes *so* and *too*. Emonds argues that *so* and *too* are allomorphs of a single morpheme, call it 'K'. K appears as *so* when it has been moved to COMP (a), but as *too* when it appears in situ (b):

a. Mary will leave town, and so will John.
   *Mary will leave town, and too will John.
b. Mary will leave town, and John will too.
   *Mary will leave town, and John will so.

Emonds' claim is that the decision to insert *so* or *too* as the allomorph of K must be made *after* the transformation that fronts K has applied: if K is fronted, insert *so*; if not, insert *too*. According to LLI, many function words and all inflectional affixes are carried through the derivation as feature-complexes and "lexicalized" only late in the derivation.

Unlike *MSSE* and the Extended Word and Paradigm model, however, Emonds' account employs a single lexicon:

> If all non-definitional properties in inflectional morphology can be predicted on the basis of independently motivated principles of a universal

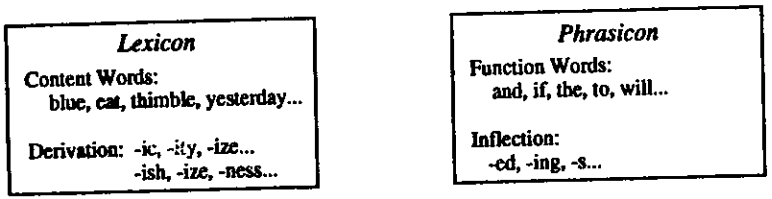
> syntax, there is no point in speaking about a separate component for inflection. A rule of inflectional morphology is simply a syntactic rule... While such rules are also constrained by principles of wider scope, this does not justify considering them to form a "component".... Thus, I maintain that inflection is not a component, but, more properly speaking, is the interface between phonology and the transformational subcomponent of syntax. (1985:246)

Despite the lack of a second lexicon, then, Emonds seems to be in agreement with Anderson's assertion that inflection is what is relevant to the syntax.

The 2 Lexicon Hypothesis

I turn now to the model argued for in this dissertation, which I call the 2 Lexicon Hypothesis, given in (13). It combines elements of the *MSSE* model, which has function words in a separate component, and the the Extended Word and Paradigm model, which has inflectional morphology in a separate component:

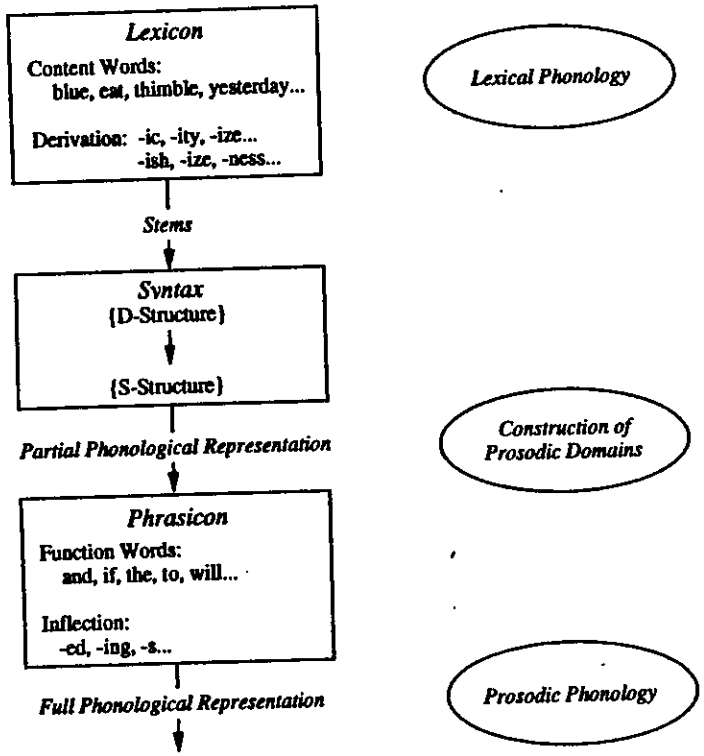
(13) the 2 Lexicon Hypothesis



According to this model, content words and derivational affixes are stored in one component, the 'Lexicon', function words and inflectional affixes are stored in another, the 'Phrasicon'. The Lexicon and the Phrasicon each define natural classes: the former defines a natural class {content words, derivational morphology}, the Phrasicon defines the natural class {function words, inflectional morphology}. The function of this dissertation will be to provide evidence that i) these are indeed natural classes and ii) that the model in (12) is an appropriate way to model these natural classes.

As with the *MSSE* and the Extended Word and Paradigm models, syntax is the mediating factor between these two components of the grammar. The Lexicon is concerned with words *qua* stems, the Phrasicon with words *qua* members of phrases. The Lexicon provides the stems that drive the syntax; the Phrasicon annotates and completes the phrasal representations generated in the syntax by "spelling out" features in the representation that have not yet been replaced by phonological strings. Some features are spelled out as inflectional affixes (agreement, case, tense/aspect), other are spelled out as function words (anaphors, modals, determiners).

(14) The 2 Lexicon Hypothesis and the grammar



The output of the syntax is thus a *partially phonologized S-structure representation*: content words and derivational affixes have been spelled out, but function words and inflectional affixes are still represented featurally. This partial phonological rerpesentation is then fully annotated by function words and inflection, yielding a fully phonologized representation.

The circles on the right of (14) indicate blocks of phonological processes that apply in tandem with the corresponding component or level on the left. Thus, by 'Lexical Phonology' I mean those rules and rule domains that include lexical roots and stems. By 'Construction of Prosodic Domains' I mean the construction of prosodic constituents larger than the word over partially phonologized S-structure representations: these prosodic constituents include phonological words, phonological phrases, intonational phrases, etc. Finally, by 'Prosodic Phonology' I mean those rules, such as the assignment of sentence-level stress, whose structural descriptions involve phonological words, phonological phrases, intonational phrases, etc.

1.1.2 Production Models

What exactly is a production model? At the very least it must be an explicit model of "the vocabularies and planning structures in terms of which sentences, (or more correctly, utterances) are constructed" (Garrett 1984). A production model *need* not correspond to a grammatical model–it is at least conceivable that we produce utterances without regard to what we (implicitly) know about their grammatical structure. Nevertheless, it is significant that even a production model that is not directly modeled on a grammatical model may yet come to resemble one. Garrett, for instance, tries

> to take natural language processing performance on its own merits, so to speak, and see what it suggests about the underlying computational systems. Surprisingly enough, we find that the outcome is near to what a straightforward interpretation of linguistic rule systems as processing systems would lead one to expect. (1976:24)

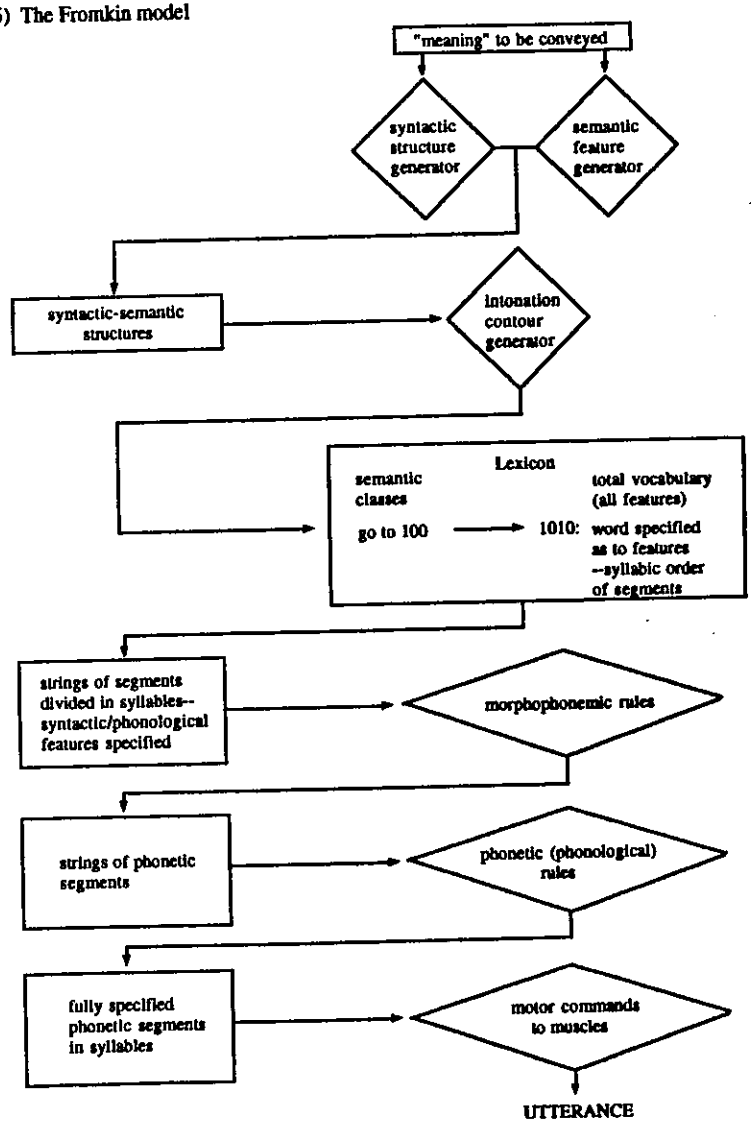
Similarly, Kean suggests that

> The language faculty is to be characterized in terms of a grammar and a processor. The grammar is an account of a person's knowledge of the structure of his language, and the processor provides an account of how that knowledge is exploited in use. (1980:26)

Indeed, the strongest hypothesis is that levels of processing correspond directly to levels of grammatical representation. Although we may *eventually* be driven to weaker alternative hypotheses, to prematurely abandon the strongest hypothesis consistent with the facts would surely be a mistake (Garrett and Kean 1981).

In the psycholinguistic literature, questions about the *place of the lexicon* in a production model received attention before questions about the *structure of the lexicon*. The first explicit linguistic model of sentence production is Fromkin's (1971) utterance generator, seen in (15). Rectangular boxes stand for levels of representation; diamonds stand for the processes that translate between levels of representation. Thus, a general meaning is converted into a syntactic/semantic representation by a syntactic structure generator and a semantic feature generator; an intonation contour is assigned to the utterance; lexical items are selected on the basis of appropriate semantic classes and inserted into the appropriate positions in the syntactic/semantic representation, yielding (roughly) a Surface Structure representation. Morphophonemic rules (e.g. *a ~ an* allomorphy, voicing assimilation in past *-ed* and plural *-s*) apply to yield strings of segments; these segments undergo phonological rules to yield a fully specified phonetic representation that is converted to motor commands to produce speech.

(15) The Fromkin model



Three features of Fromkin's model are especially important for the discussion at hand:

(i) First, notice that intonation contours that determine sentence-level stress relations are generated *before* lexical insertion applies. The motive for this is speech errors like those below:

(16) Utterance: We have a laboratory in our computer.
(Target: We have a computer in our laboratory.)

In the intended targer utterance, the first noun *computer* was to be stressed; the speech error involves exchanging *laboratory* for *computer*. But notice that the stress does not exchange with the word: in the actual utterance *laboratory* is stressed. Fromkin took this as evidence that stress was assigned to the syntactic position rather than to the word that occupied it. But if stress is assigned to the syntactic position, then it does not matter which word occupies that position: if the error involves exchanging two nouns, the one that ends up occupying the stressed position will receive stress regardless of whether stress is assigned before or after lexical insertion. Thus the diamond marked 'intonation contour generator' could just as well occur further down in the process, after lexical insertion. (The reason for this will become more clear later, when phonological rules for deriving sentence stress are considered in Chapter 3.)

(ii) Fromkin's model has only one lexicon. Having a single lexicon makes this production model compatible with single-lexicon grammatical models like LP, but not with grammatical models like those of *MSSE*, the Extended Word and Paradigm model or the 2 Lexicon Hypothesis put forth here.

(iii) Since the model generates from top to bottom without feedback loops and since only one pass is made through the lexicon, lexical insertion appears (at first glance) to occur all at once. Interestingly, however, Fromkin allows for the possibility that lexical

insertion is a two stage process (something usually taken as motivation for two lexicons). Consider the following type of error:

(17) Utterance: a money's [z] aunt
(Target: an aunt's [s] money)

Comparison of target and utterance shows that *aunt* has been exchanged with *money*, 'stranding' the plural suffix *-s*, which ends up attaching to *money* rather than *aunt*. The plural suffix would have been [s] in the target but surfaces as [z] in the speech error. Fromkin points out that this must be the result of a morphophonemic rule (plural morpheme *-s* surfaces as [z] after a vowel) rather than a phonological rule (since nothing prohibits [s] from following a vowel: *peace, space, Janice*. Crucially, the phonological form *money* must be established before the phonological form of the plural affix *-s*: if the form of the plural were established earlier, the speech error would result **in an aunt's [z] money*. As Garrett (1980) notes:

> ...Fromkin's account includes two aspects of lexical selection which are temporally ordered with respect to each other. The first is a semantically directed selection, the second determines the segmental specification of a semantically dictated element. Subsequent to this assignment of lexical forms to sites in the syntactic tree, the application of morpho-phonemic rules is assumed in order to spell out the phonetic shape of morphemes which are not fixes by the preceding processes.

Note that the same holds for the *a ~ an* allomorphy between target and utterance: *a money's aunt, an aunt's money*. Again, this cannot be due to purely phonological rules since the sequence schwa-vowel is perfectly licit in many dialects at least: *the aunt's money*. Fromkin's model, then, "implicitly but not explicitly accounts for the inclusion of words and stems and the exclusion of inflectional affixes in exchange errors, by positing that at the stage where words and stems exchange, the grammatical morphemes are not yet phonologized" (Fromkin (in prep)). (The late phonologization of grammatical

morphemes is first made explicit and discussed in greater detail in Garrett's (1975, 1984) model, to be discussed below.)

Fromkin's implicit two-stage insertion of the phonological forms of lexical items should not be confused with another two-stage lexical selection process developed by Butterworth in a series of articles, especially Butterworth 1983, 1989. Somewhat confusingly, Butterworth's two stage model of lexical access also involves 'two lexicons'. But these lexicons are divided not by the lexical *items* they contain (eg. content vs. function words) but by the lexical information they contain. Specifically, the 'semantic lexicon' contains semantic representations of words paired with phonological address; these phonological addresses correspond to phonological address in the 'phonological lexicon' which are paired there with phonological representations. Lexical access for any particular word is a two-stage process that involves two lexicons:

> In the first stage, the speaker accesses a "semantic lexicon". This in essence, is a transcoding device that takes as input a semantic code and delivers as output a phonological address. The second stage takes the address as input to another transcoding device--the "phonological lexicon"--and delivers a phonological word form as output.
>
> (Butterworth 1989:110)

Butterworth's two-stage lexical access model may thus be taken as a more explicit model of how the *first* stage of lexical insertion in Fromkin's model is processed. That is, Butterworth's semantic and phonological lexicons may be incorporated into Fromkin's model as subcomponents of her 'Lexicon'. Lexical insertion, then, may be taken as a two-stage process that involves translating "syntactic-semantic structures with primary stress and intonation specified" into "strings of segments divided in syllables with syntactic/phonological features specified" (see Fromkin's model above). The first stage translates the syntactic-semantic structures into a phonological address (in the semantic subcomponent of the lexicon), the second translates this phonological address into a fully

specified phonological representation (in the phonological subcomponent of the lexicon). In later work on the structure of the lexicon (Fromkin 1985), Fromkin explicitly divides the lexicon into three sub-lexicons: one semantic, one phonological and one orthographic. But this goes beyond the scope of the present study (see Emmorey and Fromkin 1988 for discussion).

As alluded to earlier, it was not until Garrett's model (1975, 1980) appeared that Fromkin's two-stage lexical insertion process (lexical, then grammatical items) was made explicit. Garrett also added additional speech error evidence for the claim, noting that lexical and grammatical items are involved in different types of speech errors. 'Exchanges', for instance, occur primarily with lexical rather than grammatical items (but see below, Chapter 5); 'stranding' commonly occurs with all sorts of inflectional affixes but rarely occurs with any derivational affixes; and 'shifts' occur almost exclusively with function words and inflectional affixes.

(18) Shifts

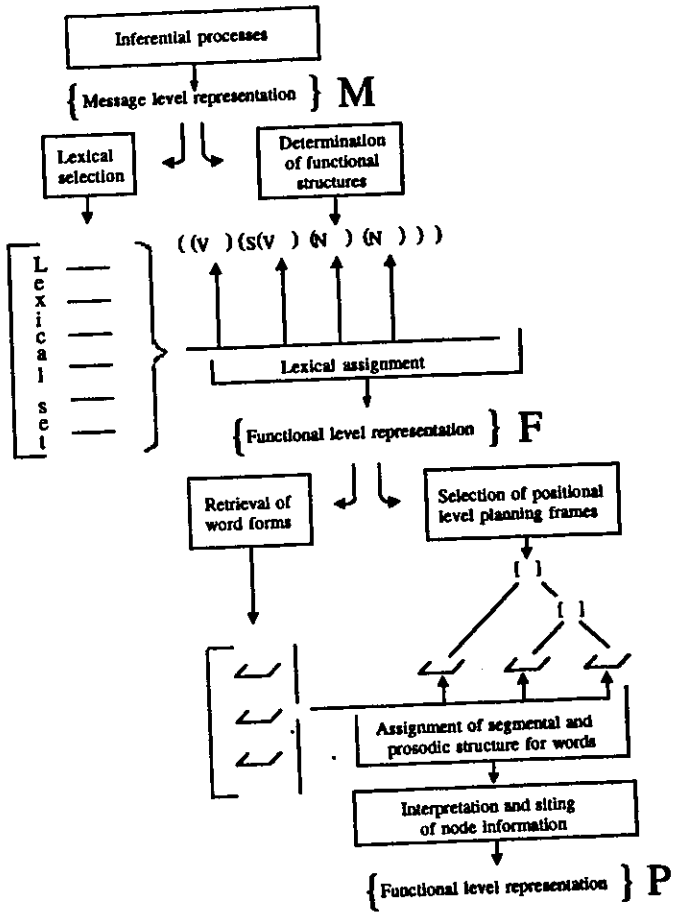
a. he has the spent most of ø time on his synthesis
b. the girl who ø taught I last year
c. what lies Joes tellø (*intended*: what lies Joe tells)
d. as I keeping suggestø (*intended:* as I keep suggesting)

(Underlined words and affixes are those that have 'shifted', "ø" indicates the position from which they have shifted.) Garrett later used the same model as a psycholinguistic model for a type of acquired speech disorder called 'agrammatism' (1982, 1984). The processor appears in (19) (from Garrett 1988).

Garrett's model is centered around three levels of representation, M(essage), F(unctional), and P(honological). The M level corresponds essentially to Fromkin's "meaning to be conveyed". Levels F and P correspond *very roughly* to traditional D-structure and S-Structure in generative grammar (Garrett 1980:216): Garrett characterizes these levels "as a *functional* level, in which phrasal membership and

grammatical functions of words are determined, and a *positional* level, in which the serial order of words and some apsects of their form are specified (1980:190). F and P thus correspond most closely to the GPSG relations of immediate dominance (ID) and linear precedence (LP), respectively (Gazdar, Klein, Pullum & Sag 1985), with the additional proviso that level P has some phonological representation while level F does not.

(19) The Garrett Model



Speech production, according to this model, proceeds as follows, with different types of speech errors marking the way:

> [W]hen the initial semantically directed access is executed...the output of that process is held, and used in conjunction with message level features to

> construct the functional level representation, in the course of which lexical elements are assigned phrasal membership (word exchanges occur here). Detailed phrasal environments are constructed (surface) clause by clause via planning frames, in the course of which (partially) phonetically interpreted forms for open class vocabulary are retrieved on the basis of the lexical representations in the functional level representations...and assigned to phrasal positions (sound exchanges and most stranding exchanges occur here). Features of planning frames are interpreted as bound or free forms and mapped onto positions in the lexical string (shifts occur here). Regular sound changes apply to yield a detailed phonetically interpreted string capable of supporting the direction of motor planning systems... (Garrett 1980:212)

The crucial aspects of the Garrett model, for our purposes here, are:

i) Lexical stems (i.e., 'open class vocabulary' minus inflectional affixes) are inserted before function words and inflectional affixes.
ii) The model employs a single lexicon.
iii) Lexical content is assigned successively to syntactic structure.

The stages in (iii) are as follows:

a. Lexical stems (not yet phonologized)--yields level F
b. Lexical stems (phonologized)
c. Function words, inflectional affixes phonologized--yields level P

(a) and (b) approximate Butterworth's *semantic* and *phonological* access, respectively. (b) and (c) approximate Fromkin's insertion of *lexical* and *grammatical* items, respectively.

iv) Function words and inflectional affixes are parts of 'planning frames' that interpret syntactic structures.

Lexical stems are *inserted into* positions in these planning frames, whereas grammatical items are parts of these frames *ab initio*. This part of the model is motivated by the 'stranding exchanges' discussed above:

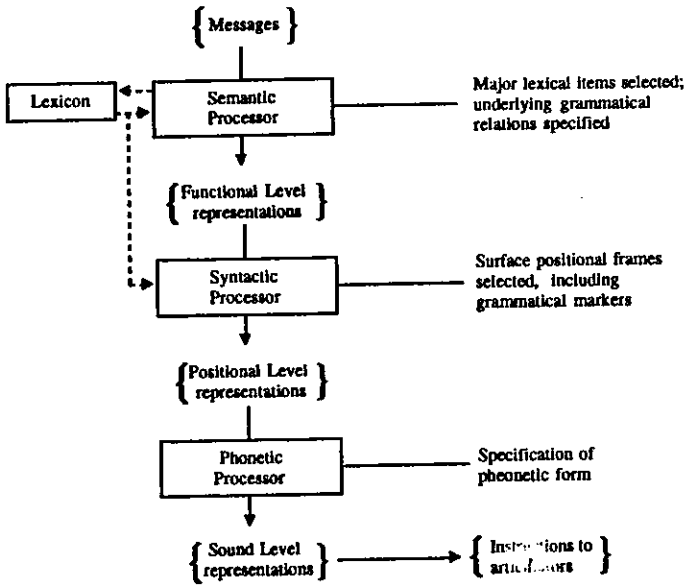
(20) Utterance: a money's [z] aunt
(Target: an aunt's [s] money)

The article and possessive affix are parts of a noun-phrase planning frame that we might characterize roughly as [INDEF ____ POSS]$_{NP}$. A stranding exchange results when a lexical stem (e.g., *money*) is assigned to a 'slot' in the wrong planning frame. In the case at hand, *aunt* and *money* are assigned to each other's planning frames with the result that the stranded grammatical items (*a* and *-'s*) end up on the wrong stem.

The Garrett production model, then, is compatible with the grammatical models espoused in *MSSE* and in the present work as regards successive levels of lexical insertion. Yet it is incompatible with these models (and with the Extended Word and Paradigm model) insofar as it has a single lexicon. In this respect it retains a more traditional view like that found in *Aspects* and in LP. Interestingly, *MSSE* takes two-stage lexical insertion as its main *evidence* for two lexicons: i.e., there is one lexicon for each stage of lexical insertion.

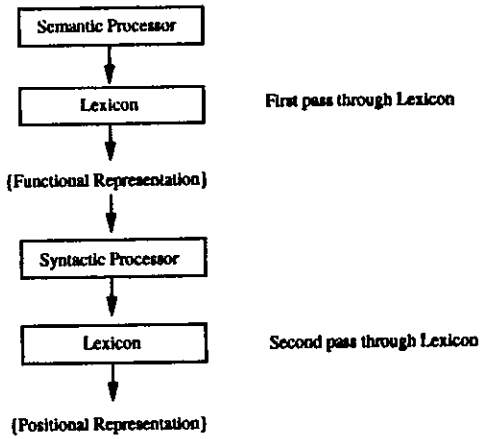
The anomaly of one lexicon with two stages of lexical insertion is clearly brought out in Lapointe & Dell's (1989) clarification of the Garrett model (21). Though Lapointe and Dell do not note this, their version of Garrett's model brings out clearly the dilemma inherent in having *two*-stage lexical insertion with *one* lexicon: the lexicon has to feed into the production process at two places--once to select the major lexical items and once, later, to select grammatical items.

(21) The Garrett Model (again)



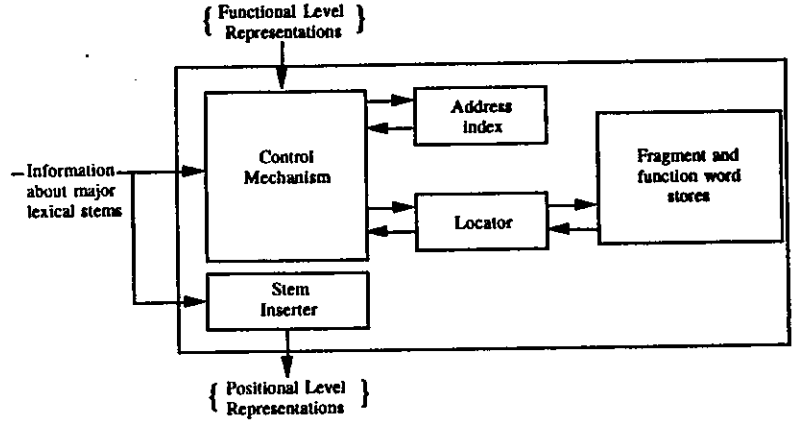
This essentially entails a *feedback loop* into the lexicon in an otherwise loopless top-down processor.

(22) The Loop



Lapointe and Dell's own model, a synthesis of their earlier work (Lapointe 1983, 1985a,1985b, 1986; Dell 1985, 1986), avoids this loop by splitting the lexicon in two.

(23) Syntactic Processor (Lapointe 1985)

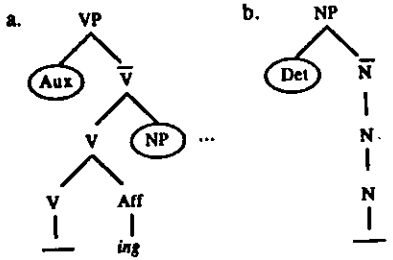


The function words and inflectional affixes that annotate planning frames are stored as subcomponents of the syntactic processor (the rightmost box in the diagram).

Information about major lexical items comes from the Lexicon, as in the Fromkin and Garrett models. But grammatical items are stored within the syntactic processor:

> We will assume that the information computed during syntactic processing is represented in two separate but interacting types of stores. The first type, the fragment store, contains phrase and function word fragments in an unordered set. The second type, the notion stores, involves sets of semantic notions of a kind typically associated with function words and inflections. (1989:114)

(24) Phrase fragments



(25) Function word fragments

The representations in (24) and (25) allow Lapointe and Dell to give explicit characterizations of the processes and levels of representation involved in deriving the Positional level from the Function level. Leaving many details aside, the stages in production proceed something like this: Given an F-level representation such as

(26) F-level representation
[indicative, active, durative, past, singular] ([KATE: x] ([def] (BOY): y] KISS (x, y)))

the syntactic parser selects phrase fragments that correspond to KATE, BOY, and KISS and the stem-inserter inserts the lexical stems *Kate, boy* and *kiss* into those phrase fragments. At this point the intermediate representation would be something like:

(27) Intermediate representation
[[Kate]$_{NP}$ [___ [kiss-ing]$_V$ [___ boy]$_{NP}$]$_{VP}$]$_S$

Meanwhile, fragment and function word stores are located that 'write out' the missing phrasal information, yielding a P-level representation:

(28) P-level representation
[[Kate]$_{NP}$ [was [kiss-ing]$_V$ [the boy]$_{NP}$]$_{VP}$]$_S$

(For the full account see Lapointe & Dell 1989).

An important property of Lapointe and Dell's EG model is that it maintains "a strong hypothesis about the relation between structural levels in the grammar and in the speech processor" (1989:146ff):

(29) Grammar and Processor: correspondences in levels of representation

Grammar	*Processor*
Semantic Representation	Message Level
D-Structure	Functional Level
S-Structure	} Positional Level
Phonological Representation	
Phonetic Representation	Sound Level

Lapointe & Dell claim that the relation between the grammar they assume and the processor they argue for obeys the 'Correspondence Hypothesis':

(30) *Correspondence Hypothesis*
Every representational level in the grammar corresponds to a level in the processing system.

The Correspondence Hypothesis (CH) is a weakened version of the Unique Correspondence Hypothesis (UCH):

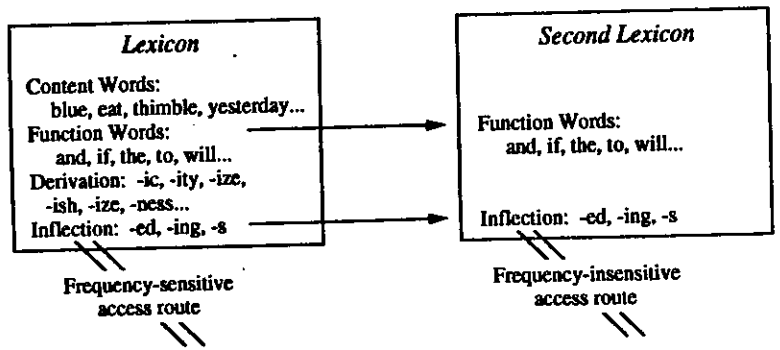
(31) Unique Correspondence Hypothesis
Every representational level in the grammar corresponds to a unique representational level in the processing system.

Their model fails to meet the UCH because more than one grammatical level (S-Structure and Phonological Representation) is associated with a single processing level (the Positional Level).

Lapointe & Dell's EG model is incompatible with a grammar that incorporates a single lexicon: no component of such a grammar corresponds to the EG fragment and function words stores. But the function word stores *do* correspond directly to the 2nd Lexicon in the *MSSE* and the fragment stores (insofar as they contain inflectioal affixes) *do* correspond directly to the Inflectional Component found in the Split Morphology Hypothesis. And, taken together, the fragment and function word stores correspond directly to the Phrasicon in the model outlined here.

Bradley, Garrett & Zurif (1980) proposed a psycholinguistic model in which grammatical items are *doubly-listed*: i.e., they are stored along with other words and affixes in the Lexicon and, in addition, are stored in a separate lexicon for rapid access in (normal) speech production:

(32) A Double-Listing Model



Evidence for such a model came from the agrammatic (telegraphic) speech that often accompanies Broca's aphasia. Very briefly, agrammatics tend to omit function words and inflectional affixes in their speech and to be less aware of these items in their comprehension. Bradley, Garrett & Zurif hypothesized that acquired damage to (what I have called here) the Second lexicon lies at the root of agrammatism.

Additional evidence came from word recognition experiments. Bradley (1978) and Bradley and Garrett (1979) performed word recognition tests on both normal and agrammatic speakers to see if content and function words had different accessing times for either population. In their experiments they found that they did, but only for normals. More precisely, the time that normal speakers took to recognize a content word was a function of the frequency of that word: more commonly occurring content words tended to be recognized much more quickly than ones that occur less commonly. But, for normals at least, the frequency of a function word did *not* affect the amount of time it took for a subject to recognize it as a word. This they took as evidence that normals have a special rapid-access route to grammatical items (i.e., to the Second lexicon). Agrammatic speakers, on the other hand, showed frequency affects both for content words *and* for function words. Bradley and Garrett took this to mean that, through impairment to the Second lexicon, agrammatics had lost the rapid access to grammatical elements that normals have; consequently, agrammatic subjects had to rely on the frequency-sensitive access route for both lexical and grammatical items.

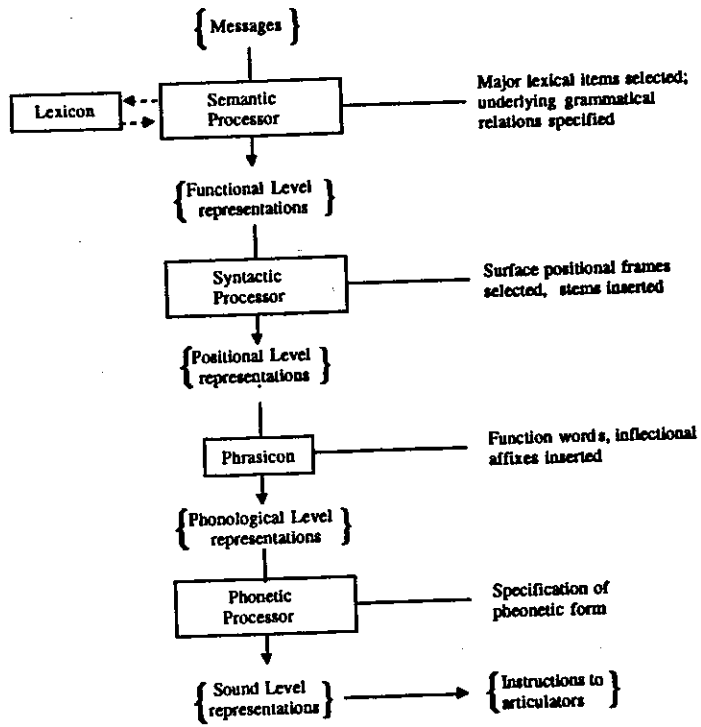
It must be admitted, however, that subsequent experimentation by a number of researchers has *not* reproduced these results. The recognition of both content and function words was shown to depend on frequency in English (Gordon and Caramazza 1982, Garnsey 1985), French (Segui, Mehler, Frauenfelder and Morton 1982) and Dutch (Kolk and Blomert 1982). As Friederici points out, "it seems, therefore, that the frequency reults of the Bradley experiments should be treated with extreme caution"

(1985:135). Subsequent work by Friederici (1985) has brought the issue back to life. This will be discussed below (Chapter 6, section 4.)

The 2 Lexicon Hypothesis and processing model: revising the Extended Garrett model.

The 2 Lexicon Hypothesis may be incorporated into the Extended Garrett model with only a few revisions. Like Lapointe & Dell's EG model, the 2 Lexicon Hypothesis model incorporates a second lexicon. Unlike the EG model, however, Stem insertion precedes the insetion of all grammatical items. Grammatical morphemes do not begin life as parts of planning frames but, like stems, are inserted into planning frames (only at a later stage than stems). Evidence for this will be discussed in later chapters (especially Chapter 5 on Speech Errors and the analysis of "shifts"). A preliminary version of the resulting model is given below:

(33) Revised Extended Garrett Model (preliminary formulation)



Note the addition above of another level of representation, "the Phonological level". This level represents the output of the Phrasicon and is the first level of representation to contain full phonological representation for all words and affixes. This additional level gives the processing model one level of representation for each level of representation in the grammar:

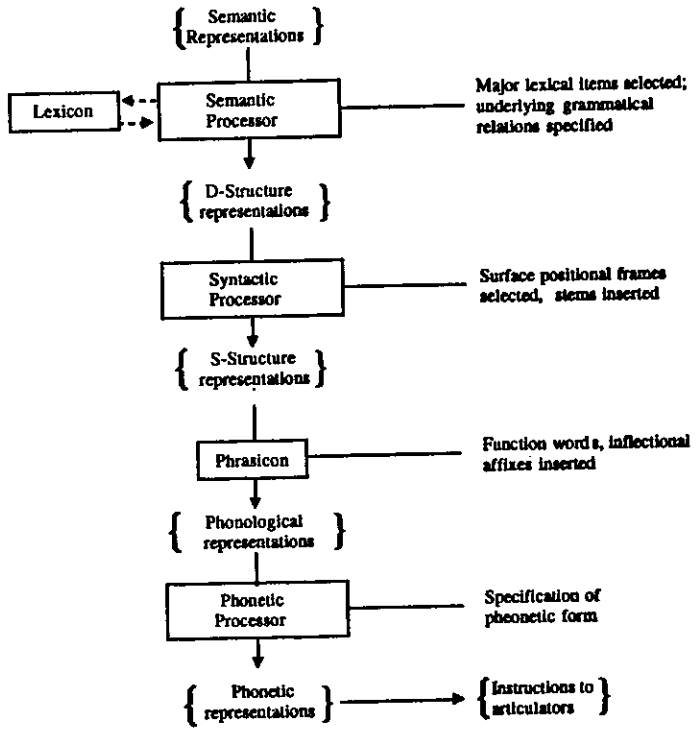
(34) Unique Correspondence of Grammatical and Processing Levels

Grammar	*Processor*
Semantic Representation	Message Level
D-Structure Representation	Functional Level
S-Structure Representation	Positional Level
Phonological Representation	Phonological Level
Phonetic Representation	Sound Level

Actually, the UCH must be revised somewhat, since not all of the grammar (presumably) is actually used in processing. For example, since Chomsky & Halle 1968, many generative phonologists have held that Velar Softening (*opaque* [k], *opacity* [s]) and Trisyllabic Shortening (*serene, serenity*) are synchronic phenomena of English (Halle 1977, Kiparsky 1982, Halle & Mohanan 1985). Despite claims that psycholinguistic evidence supports this (Cena 1978), the reality of such rules has been challenged by a compelling array of experimental research (Jaeger 1986, McCawley 1986, Wang & Derwing 1986 and references therein). There seems to be little evidence that such 'rules' are used in processing: no slips of the tongue seem to involve them, no brain damage seems to affect them, etc. Conceivably, then, there are levels of grammatical description (e.g., levels of derivational affixation) that are *not* matched by levels of processing. In this case, the UCH must read. "Every representational level in the processing system corresponds to a unique representational level in the grammar (but not vice versa)".

With that proviso, however, the Revised Extended Garrett model may be re-labeled so that processing levels are named directly by the corresponding level of representation in the grammar:

(35) Revised Extended Garrett Model (final version)



The Revised Extended Garrett model above corresponds directly to the 2 Lexicon Hypothesis model given above.

Conclusion

Although most researchers would agree that some kind of lexicon or lexicons are necessary both in grammars and in production models, the exact number and nature of lexicons is still under debate.

The number of lexicons used in current theories ranges from a low of one to a high of two. The nature of lexicons depends in part, of course, on the number of lexicons.

Grammars with one lexicon assume that all words and affixes in a language are represented in that lexicon (LP, Emonds model, Fromkin model, Garrett model). Grammars and production models with two lexicons, however, differ in what each lexicon holds: ***MSSE*** contains an explicit model in which one lexicon holds content words and all affixes, while the other holds only function words. The Extended Word and Paradigm model seems to have all words and derivational affixes in one lexicon, with inflectional affixes in the other. Lapointe & Dell's EG model has function words and inflectional affixes in a store separate from content words and derivation.

I have tried to outline a grammatical model and a production model that are essentially homomorphic. The grammatical model incorporates a 2 Lexicon Hypothesis which puts content words and derivational affixes in a "Lexicon", and function words and inflection in a "Phrasicon", with syntax mediating between the two. The production model is essentially equivalent to the EG model except that grammatical items do not come as part of planning frames but are inserted into planning frames, just like (but later than) lexical items. Also, the REG model includes an extra level of representation which corresponds to Phonologica Representation in traditional generative models, allowing for an extremely close fit between grammar and processor.

1.2 How many types of lexical item?

I have outlined a number of models of lexical storage and access, focusing primarily on where content words, function words, derivation and inflection are stored in a grammar and in a processing model. I now turn to these four types of lexical item to determine by what principles they may be defined more precisely.

Somewhat surprisingly, perhaps, what counts as e.g., a function word or an inflectional affix remains largely unchanged from *Syntactic Sturctures* to the present. What *does* change from one linguist to the next are the criteria that determine a lexical item as "content word" or "function word", "inflection" or "derivation". Along with changes in criteria go changes in terminology. Common terms for words include ***Lexical*** and ***nonLexical (Grammatical) words, Major*** and ***Minor Category words; Open*** and ***Closed Class words; Content*** and ***Function words,*** (1.2.1). Common terms for different types of inflection are more constrained: one finds primarily ***Derivational*** and ***Inflectional affixes,*** though the terms ***Lexical*** and ***Grammatical*** occur as well (1.2.2).

1.2.1 Types of words

Let us begin, then, with types of words. Regardless of nomenclature, the two classes of words generally distinguished are as follows:

(36) Types of words

A-type	*B-type*
Nouns	Prepositions
Verbs	Auxiliary Verbs
Adjectives	Modal Verbs
(Adverbs)	Pronouns
	Determiners
	Conjunctions
	Complementizers
	Particles
	(Adverbs)

I have parenthesized adverbs here because many discussions of word type do not mention them. De-adjectival adverbs like *quickly* and *slowly* pattern in most cases with the A type of word; non-derived adverbs like *not* and *yet* often pattern with the B type.

But given that these two main types of word are distinguishable, what distinguishes them? Proposals revolve around the following criteria:

i) open or closed class
ii) syntactic features
iii) semantic features

Open and Closed Class Items. A common way of distinguishing types A and B is to make a distinction between "open "and "closed "classes of words: the former classes are very large, typically numbering in the thousands, the latter are quite small, with (say) less than thirty or so members each. Speakers may, and often do, add new nouns, verbs and adjectives to their language (*palimony, smog, PC, UNESCO,* etc.) but rarely make up new prepositions, auxiliary verbs, etc. Similarly, borrowing is very common among the open classes; cases of borrowing determiners, modal and pronouns are much less common (e.g., the borrowing of *they, their* and *them* into English from Old Norse).

But inclusion in a closed class do not guarantee that a word falls into the B type (though the converse may be true). Kinship terms are a good example: English, for instance, has a limited number of them (*brother, sister, aunt, uncle, father, mother,* etc.) and one is hard pressed to think up new ones; yet all of these items function in the syntax, morphology and phonology just like any other noun. Any number of 'closed' classes behave this way: days of the week, names for months, the cardinal directions, and so on.

In any case, there is something unsatisfying about defining a class in terms of its size or in terms of limitations on its size. The question still arises: *why* are some classes closed while others are open? For this reason, I supsect, discussions of open and closed classes tend to include other defining properties, such as syntactic or semantic features.

Syntactic Features. The terms "Lexical" and "non-Lexical" (or "Grammatical") are often used for groups A and B. These terms have the advantage of being definable in terms of the binary syntactic features [±V] and [±N], as follows (Chomsky 1970, Bresnan 1976):

(37) Syntactic feature analysis of "Lexical"

	[+N]	[-N]
[+V]	A	V
[-V]	N	P

An obvious problem with this applying this analysis to classes A and B above is that it includes prepositions. The terms "major class" (N, A, V) and "minor class" (P, Aux, Modal, etc.) are often used, but these terms are not as straightforwardly defined in terms of syntactic features. Nespor & Vogel (1986:169) for instance, propose using the features in (2) to select items "with at least one positive specification according to the categorial feature system". This includes N, A, V but excludes P.

A related issue is the status of adverbs like *quickly* and *slowly*. They seem to be lexical insofar as they are built on adjectival stems but there is no 'room' in the chart in (3) for them. Following a suggestion by Emonds (1985:201), however, we can treat adverbs in *-ly* as inflected adjectives: this gives them a space on the chart in (3) and correctly categorizes them with the stems they are formed from. Emonds cites an additional advantage in classifying adverbial *-ly* as inflectional: "An interesting property of Enlgish inflection is that only one is allowed per word. Under the plausible assumption that the adverb-forming *-ly* suffix is onflectional on A, we then correctly predict that comparative and superlative suffixes cannot be added to it: *slowlier Vs. friendlier; *freeliest vs. holiest" (ibid). Further evidence that adverbial *-ly* is an inflectional affix will be given below in Chapters 4 (evidence from morphology) and 5 (evidence from speech errors).

Another problem with this type of analysis concerns pronouns. As pro-*nouns*, they surely carry the features [+N, -V], yet they are prototypically type B rather than type A. Emonds (1985) argues that this problem is quite widespread. He claims that all open class categories (N, A, V) have functional or closed class sub-categories which he calls "disguised lexical categories". Pronouns are a closed sub-class of nouns; auxiliary verbs (*have, be*) are a closed sub-class of verbs; modifiers such as *other, same, different, such, many, few, much, little* form a closed sub-class of adjectives. As was mentioned above, adverbs can also be open class (*daily, twice*) or closed class (*thus, then*).

It would seem, then, that the syntactic features in (2), though they may define natural classes in the syntax, do not define the natural classes lurking behind the terms "lexical" and "grammatical "or "open" and "closed" class.

Semantic Features. The terms "content words"and "function words" are often used to define the A and B classes. These terms refer straightforwardly to the *semantic* content of words and thus avoid some of the pitfalls of the syntactic feature analysis. Bradley, Garrett and Zurif's (1980) notion of interpretive burden is a case in point:

> The two (at present, not sharply defined) classes of words seem to have differeing roles in sentences, whether from a formal perspective...or from a computational perspective. In broad terms, the two classes diverge in terms of what might be called interpretive burden. The closed-class (grammatical morphemes, minor grammatical categories, nonphonological words) includes sentence elements that, by and large, are vehicles of phrasal construction rather than primary agents of reference, as is the case with open-class words (content words, major grammatical categories).
>
> (Bradley, Garrett, and Zurif 1980:277)

The most explicit treatment of this sort that I am aware of is that of Emonds (1985). He makes a basic distinction between semantic features, which distinguish one lexical item from another and syntactic features, which, among other things, distinguish one grammatical item from another. Emonds *derives* the fact that grammatical items form closed categories from two claims:

i) that grammatical items are distinguished solely by syntactic features
ii) that there is a limited and small number of syntactic features

Circularity of definition is avoided by defining semantic features in terms of syntactic rules: "A feature with semantic content not used in any syntactic rule is called a (purely) semantic feature" (1985:165).

Emonds' approach is certainly very close to the mark in capturing the essentail differences between the A and B words in (1). A strict interpretation of his definition of "semantic", however, is not without problems. Recall that a semantic feature is one "with semantic content not used in any syntactic rule". Conversely, "every feature/category that appears with morphemes in closed classes is syntactic, i.e., can be used to state syntactic rules" (1985:166). But Emonds himself note that words like *this* and *that* differ in terms of a feature that does not seem to be used in any syntactic rule. He claims, however, that "whatever feature distinguishes them surely distinguishes *here* and *there* also, and this feature is used in the syntax (i.e., *there* but not *here* appears in existentials, etc." (ibid) .

But "appearing in existentials" and "being used to state syntactic rules" are hardly equivalent notions. Consider some other cases: *come* and *go*, which Emonds claims to be function words as well, must differ in a purely syntactic feature, one that can be used to state syntactic rules. Similarly with pure deictics such as *this* and *that, here* and *there.* All of these pairs seem to involve the same feature, let us call it [proximate], yet this feature is not one that syntactic rules such as NP- or *wh*-movement refer to. Prepositions, which are also function words for Emonds, are another class of words that differ in terms of features that are not used in stating syntactic rules. Whatever the features are that differentiate *up, down, in, out, to, from,* etc., they are not syntactic features in the sense Emonds intends. Not to call these features semantic seems perverse, since they are ostensibly the same features that distinguish *rise* from *fall, enter* from *exit, arrive* from *leave*. Person features provide another difficult case for Emonds' approach: he notes that pronouns such as *you* and *we* differ in some feature and claims that this feature is syntactic since *you* but not *we* is allowed to delete in imperatives. But this is irrelevant since *we* (and *they, he, she* and *it*) ao not appear in imperatives in English to begin with (*We leave!). A final case is grammatical gender: *he* and *she* would seem to differ by

the same feature, say [± female] as *man* and *woman, boy* and *girl, stallion* and *mare, boar* and *sow*. Yet, on Emonds' account, the feature that distinguishes *he* from *she* must be "grammatical," while the one that distinguishes *man* from *woman* must be "semantic".

Discourse features. A slight shift in perspective seems to be necessary here. Suppose that content words and function words differ not so much in *syntactic* features, ones that are used to state syntactic rules, but in *discourse* features, ones that refer to the verbal interaction that the utterance is part of. Since discourse presupposes syntax, all of the cases Emonds' account covers will be covered here as well. But the problematic cases for Emonds' account fall into place if they are characterized in terms of discourse. Let us approach the four cases one at a time.

i) *Deictics.* Deictics refer straightforwardly to the location of the discourse participants. This is true not only of pure deictics like *here* and *there, this* and *that*, but of motional deictics such as *come* and *go* as well. *here, this* and *come* involve [proximity] *with respect to the speaker.* Features that differentiate tense and aspect are also deictic: Past, Present and Future refer to the time preceding, at or after the time of the discourse; Perfective and Imperfective involve the completion or non-completion of an act in terms of the time of discourse, and so on.

ii) *Prepositions.* Directional prepositions such as *up, down, in, out, to, from* are generally used with reference to the speaker's or hearer's position:

> Vertical and horizontal provide a perceptual frame of reference for path and directional descriptions. Indeed, one might think of them as internalized landmarks anchoring judgments of direction and orientation in much the same way primary colors anchor judgments of hue.
> (Miller & Johnson-Laird 1976:408)

Similarly for locative prepositions like *at, on, under, over*: given a configuration like "±" we must say the minus-sign is *under* the plus-sign. Turning the page upside down--i.e.,

altering *only* the relation of the "±" to the speaker or hearer--forces us to say that the minus-sign is now *over* the plus sign.

iii) *Person.* Person features are perhaps the clearest case of discourse oriented features: 1st person is defined in terms of the speaker, 2nd in terms of the hearer, 3rd in terms of the non-participants in the discourse. Although no syntactic rules refer directly to person, person features are inextricably part of the discourse.

iv) *Gender.* Gender is not a purely syntactic notion either, as the *he, she, boy, girl* examples show. But neither is it purely semantic: even in English the referents of *she* may be inanimate (boats, countries, etc.). And in languages with more developed gender systems, such as German, the difference between grammatical gender and actual gender is even more pronounced: things may be masculine (der Wagen 'the car') or feminine (die Sache 'the thing') as well as neuter (das Ding 'the thing'), feminine beings may be neuter (das Mädchen 'the girl'), etc. What grammatical gender does is help track reference from function words (pronouns, articles) to content words elsewhere in the discourse.

One final class of words should be discussed in this regard, so-called particles. Consider particles in Ancient Greek. These are for the most part hard-to-gloss evidentials, "expressing a mode of thought, considered either in isolation or in relation to another thought, or a mood of emotion" (Denniston 1934). A typical discussion of some of these helps convey their flavor:

> *toi* presses an idea upon the attention of the person addressed: 'I would have you know (or remember); *pou* conveys doubt, 'I suppose'.... Affirmation is expressed par excellence by *ée*, which ...afffects the thought as a whole: while *deé* and *ge* tend to cohere with the preceding word....
> Besides expressing modes of thought, these particles...also indicate moods of emotions, nuances. Thus pathos is often suggested by *deé*, irony or sarcasm by *deé* and *déethen* (sometimes by *ge*), interest and surprise by *ára* and *ge*, sympathy, encouragement, threatening hostility, and other

> attitudes by *toi*, sudden perception or apprehension by *kaì meén* and *kaì deé*. (Denniston 1934:xxxvii ff)

These particles form a closed class and are clearly *function* words on anyone's account. But it is equally clear that the features which distinguish them from one another are not features that are used to state syntactic rules. Words that express merely "modes of thought...moods of emotions, nuances" are surely best characterized in terms of the discourse.

Content words and function words, then, differ in *how* they refer. *Mary* and *she* may both refer to the same person, but the former refers to an unchanging entity while the latter refers to a constantly changing entity that is definable only in terms of the discourse: *she* is the female non-participant in the discourse that is visible to both speaker and hearer. Content words refer to their referents either directly (names) or by a specific sense/meaning (nouns, adjectives, verbs). Function words refer to their referents *only* via mediation by the discourse, its time and place and its participants.

1.2.2 Types of affixes. The traditional distinction between derivation and inflection is based partially on the observation that a set of affixes can be found which never occur "inside" of other affixes in the language which do not fall inside that same set. Bloomfield, for instance, notes that

> "the structure of a complex word reveals first, as to the more immediate constituents, an outer layer of *inflectional* constructions, and then an inner layer of constructions of *word-formation*.... The constructions of inflection usually cause closure or partial closure, so that a word which contains an inflectional construction (an *inflected* word) can figure in no morphologic constructions or else only in certain inflectional constructions. The English form *actresses*, for instance, can enter into only one morphologic construction, namely the derivation of the possessive *actresses'*... This latter form, in turn, cannot enter into any morphologic construction; it has complete closure. (1933:222)

Inflectional affixes include person, number, and case-marking on nouns and adjectives, tense, voice, mood and aspect marking on verbs. In languages such as Latin or English, these affixes always occur in layers outside of derivational affixation: the only affixes that occur outside of inflectional affixes are other inflectional affixes. The Latin passive marker *-ur*, for example, occurs as the outer layer of inflection:

	Active	Passive
Singular	fer-t	fer-t-ur
	carry-3sg	carry-3sg-PASS
Plural	fer-unt	fer-unt-ur
	carry-3pl	carry-3pl-PASS

The fact that it regularly occurs outside of other inflectional affixes and never occur inside of derivational affixes classifies the Latin passive as inflectional. Similarly with adverbial *-ly*. Not only is it never followed by an inflectional affix (see above), but it is never followed by a derivational affix either: ***quickliness* vs. *friendliness***. This layering of inflection outside of derivation provides the motivation for a late inflectional level of affixation in models such as Lexical Phonology or the Extended Word and Paradigm model.

A second traditional observation about inflection is that it comes in *paradigms*. Almost all nouns have singular and plural forms, almost all verbs have present and past tense forms, etc.

> Given one of these [forms], the speaker is usually capable of producing the other. Eaxh set of forms is called a paradigmatic set or paradigm, and each form in the set is called an inflected form or inflection. Some languages have large paradigms, which contain many inflctions. In Latin, for instance, the verb appears in some 125 inflectional forms, such as amaare 'to love', amoo 'I love', amaas 'thou lovest', amat 'he loves', amaamus 'we love', amem 'I may love', amor 'I am loved', and so on; athe

occurrence of one form usually guarantees the occurrence of all others.
(Bloomfield 1933:223)

Traditional inflectional paradigms were introduced into generative grammar primarily by Matthews (1972) in his Word and Paradigm (WP) model of inflectional morphology. Anderson later incorporated much of Matthews work into his Extended Word and Paradigm model of a post-syntactic inflectional component. One of Matthes' central observations is that inflectional endings often encode a number of features into a single portmanteau morph: thus the *-t* in Latin *fert* 'he carries' signals not only 3rd person and singular, but also present tense, active voice, and indicative mood. Portmanteaux are rare in derivational morphology but extremely common in inflection. As Matthews demonstrates, they are difficult to treat with concatenative morphology of the Item and Arrangement sort (e.g., Hockett 1954) and provide strong evidence for the traditional device of paradigms.

Many other differences between inflection and derivation have been proposed and I will not attempt to discuss them all here (see Siegel 1974, Aronoff 1976, Anderson 1982, Scalise 1984 for overviews). But the claim that there is a clear distinction at all between derivation and inflection has been challenged (e.g., Bybee 1985, Di Sciullo & Williams 1987). Much of the rest of the present work argues for a clear distinction between derivations and inflection, so I will leave the issue for now.

For those that do accept the distinction between derivation and inflection, the question remains: what is it that distinguishes inflection from derivation?

Open and Closed Class Items. An analysis based on open and closed classes will fail miserably in distinguishing derivation from inflection, since both classes are closed. Derivation may be "less closed" than inflection, since new derivational affixes can be created: e.g., *Dance-athon, Pink Floyd-athon,* etc. from *Marathon, work-aholic, worry-aholic,* etc. from *alcoholic.* But neither class of affixes is "open" in the way that nouns, verbs and adjectives are.

Syntactic features. Selkirk (1982) proposes that derivational affixes are marked for the features [±N, ±V]. This is meant to explain the observation that derivational affixes may change the part of speech of the word they attach to: derivational affixes carry nominal and verbal features and affixation produces right-headed words: $[[\text{obeac}]_{\text{Adj}}\ \text{ity}]_{\text{N}}$. On this view, inflectional affixes are not marked for the features [±N, ±V] and are therefore "transparent" in determining part of speech.

This is a useful way of characterizating inflectional affixes. But it does little in the way of telling us *which* affixes are derivational and which are inflectional. For that, we need to be able to specify *what it is* that makes affixes derivational or inflectional. Anderson's analysis of inflection, of course, is pertinent here: inflection is what is relevant to the syntax.

Semantic features. Emonds argues that Anderson's characterization is somewhat too broad. He proposes instead that "inflections are those bound morphemes which are relevant to transformational (as opposed to deep structure) syntax" (1985:195). This brings his criterion for inflection into line with his criterion for function words: inflectional affixes and function words lack purely semantic features.

Since the features that function words and inflectional affixes write out are generally drawn from the same set (tense, aspect, person, number, etc.), my arguments against Emond's analysis of function words may be applied to his characterization of inflection as well: they are not so much syntactic features as features relating to the discourse, its time, place and participants. Take third singular *-s*, for instance: no syntactic rule of English refers to the features [third person] or [singular], so in Emond's terms they should

not be *syntactic* features[2]. But they clearly are *discourse-related* features in that they encode a single non-participant in the discourse, i.e., [-speaker, -hearer, -plural]. Similarly, the features that distinguish tense morphemes, e.g., Present from Past, are not *syntactic* featuress that rules such as NP- or *wh*-movement refer to, but *semantic* features that place events on a time-line whose focal point is the time of the utterance.

Discourse Features. In general, then, it seems that both function words and inflectional affixes may be characterized as what is relevant to the discourse, narrowly construed. I say 'narrowly construed' becaue a great many things are relevant to the discourse. By *discourse* I mean simply the time, place and participants of the discourse. This is meant to preserve what is right about Anderson's and Emonds' analyses while including notions such as tense, aspect, location and person as well.

This *characterization* of function words and inflectional affixes is, of course, not yet a *definition.* To define function words and inflectional affixes in terms of discourse features would require a small but complete set of such features and a theory of why exactly *that* set of features is the relevant set. I have no such set of features at present, nor such a theory.

What this boils down to is that the discourse characterization of function words and inflectional affixes I have sketched here is not a definition that can predict *a priori* which words and affixes are stored in the Lexicon as opposed to the Phrasicon. Perhaps future research will provide such a definition.

But this lack of a precise definition need not pose an insuperable obstacle to the issue of lexical storage: many notions in linguistics theory and elsewhere are difficult to define

[2] Unless, of course, we take *agreement* to be a syntactic rule. But this is circular and adds no evidence for or against the claim that inflection is defined in terms of purely syntactic features.

but nevertheless important to use. The notion syllable is a case in point: as Keating points out, "the phonetic literature is quite inconclusive as to either what a syllable is, or how to define its properties, such as sonority, that might be exploited to define syllables, although phonologists have ben happty to assume that such questions are nearly settled" (1988:292). Despite any such definition of what a syllable actually *is*, however, syllables play a central role in phonological theory.

To abandon the syllable because we have no precise definition of it, is to throw out the baby with the bath-water. Likewise with content words, function words, derivation and inflection. All have defied explicit definition, but they continue to be invaluable in both grammatical and psycholinguistic research.

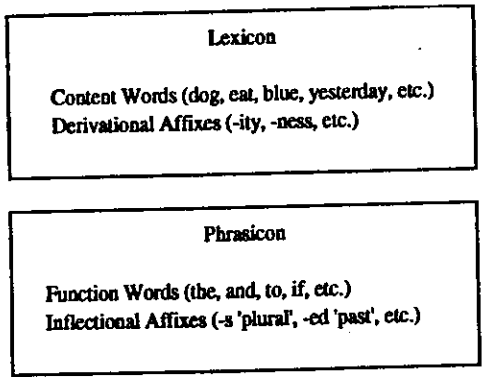
1.3 Conclusion

I have tried to motivate a model of lexical storage and processing that can be used *both* as a grammatical model and as a processing model. This 2 Lexicon Hypothesis model utilizes two lexicons that store fundamentally different types of item: the Lexicon stores words and affixes that refer directly to states, things and properties in the world; the Phrasicon stores those that refer times, places and participants in the discourse. The two lexicons are separated by syntax, which is fed by the Lexicon and which feeds the Phrasicon.

2. *Minimal Word, Minimal Affix*

Recent studies (Prince 1980, McCarthy and Prince 1986, 1990) have begun to articulate the notion of a 'minimal word': many languages place minimal prosodic restrictions on the size of well-formed words. In Estonian, for instance, content words must consist of at least a bimoraic foot (Prince 1990) and in Japanese, well-formed derived words must consist of at least a bimoraic foot (Poser 1990)[1]; words in Dyirbal must be at least disyllabic (Dixon 1972); and so on. As McCarthy & Prince note, however, *function* words in such languages often do not obey these minimality constraints, which hold exceptionlessly only for *content* words. English, for instance, has a strict minimal word requirement (CVX, i.e., CVV or CVC) which the articles *the* (CV) and *a* (V) defy; Diyari has a two-syllable minimal word requirement not met by the monosyllabic conjunction *ya* 'and' (McCarthy & Prince 1986, citing Austin 1981); and so on.

(1) The 2 Lexicon Hypothesis



[1] Underived words may consist of as little as a monomoraic syllable: Japanese has a number of words like *su* 'vinegar', *na* 'name', *ta* 'rice field' (Itô 1989).

Languages that impose minimal word requirements on content words (N, A, V, Adv) but do not impose those same requirements on function words (articles, conjunctions, prepositions, complementizers, etc.) provide an interesting type of support for the 2 Lexicon Hypothesis. Recall that on the 2LH content words and function words are in separate lexicons, as shown above in (1). (1) makes it possible to state minimal word restrictions that hold for content words but do not hold for function words without resorting to stipulations about category membership. A prohibition against CV content words may be stated simply as a condition on well-formed *words in the Lexicon*; *words in the Phrasicon* may or may not be subject to such conditions.

But the 2 Lexicon Hypothesis makes a further prediction in grouping derivational affixes with content words and inflectional affixes with function words: just as words-in-the-Lexicon may be subject to minimal-size requirements that words-in-the-Phrasicon are not subject to, affixes-in-the-Lexicon could be subject to minimal-size requirements that affixes-in-the-Phrasicon are not subject to. That is, all morphemes in the Lexicon may be subject to minimal requirements whereas morphemes in the Phrasicon may not be. Evidence from English and Attic Greek suggests just this: derivational affixes in these languages are subject to a Minimal Affix requirement that inflectional affixes are not subject to. Derivational affixes consisting of less than a syllable in these languages are prohibited; inflectional affixes consisting of less than a syllable are quite common.

In at least two languages, then, it seems that nouns, verbs, adjectives, adverbs and the affixes that create nouns, verbs, adjectives, and adverbs are subject to prosodic well-formedness conditions to which articles, conjunctions, pronouns, complementizers, etc. and inflectional affixes are not subject, as shown in (2).

(2)

	minimal requirement
Content Words	yes
Derivational Affixes	yes

Function Words	no
Inflectional Affixes	no

Again, this is modeled straightforwardly if we assume that the speakers of these languages store and process content words and derivational affixes in a different place than they store and process function words and inflectional affixes. On the 2 Lexicon Hypothesis, the table in (2) can be reduced to the statement in (3):

(3) Minimal prosodic requirements are met in the Lexicon.

Words and affixes *in the Phrasicon* are not subject to minimal prosodic requirements.

But an even stronger version of (3) is possible, one that holds not only for words and affixes but for roots as well--i.e., for the output of any level of word-formation in the Lexicon:

(3') Minimal prosodic requirements are met at all levels in the Lexicon.

The generalization in (2) may then be recast as in (2'):

(2')

	minimal requirement
Lexical Roots, Content Words	yes
Derivational Affixes	yes
non-Lexical Roots, Function Words	no
Inflectional Affixes	no

(2') and (3') essentially make the claim that minimal word requirements (in some languages) are epiphenomenal: they derive from minimal root and minimal affix requirements--ie, they derive from minimal requirements on *morphemes*. Morphemes-in-the-Lexicon are subject to minimal requirements to which morphemes-in-the-Phrasicon are not subject.

Evidence for (2') and (3') is presented for English (2.1), Attic Greek (2.2), and Latin (2.3) as follows. First, a minimal word requirement is established for each language and it is argued that this requirement follows from a minimal requirement on *roots*. It is then shown that function words and the roots from which they are built in these languages are not subject to the minimal word requirement--this reduces to the claim that the roots of function-words are not subject to minimal prosodic requirements. Finally, a minimal affix requirement of is

established for each language and it is shown that inflectional affixes in these languages are not subject to this minimal requirement.

2.1 English

The Minimal Word = μμ It has long been noted that English allows no (content) words that consist of a single light syllable. Lexical words are minimally bimoraic (4):

(4) English: Min Wd = μμ

This holds not only of lexical *words* but also of lexical *roots*.

Content words in English consist minimally of a bimoraic syllable. CWs may be heavy CVC (5a) or CVV (b), but not light CV (c)[2]:

(5) Minimal Word (English)

a. CVC		b. CVV		c. *CV	
pɛk	'peck'	see	'say'	*pɛ	*sɛ
pɪt	'pit'	sii	'see'	*pɪ	*sɪ
pʊt	'put'	suu	'sue'	*pʊ	*sʊ

Lexical Roots.

Determining what the lexical roots of English are is straightforward if only native English vocabulary is used--all of the words in (5) are roots as well as words and the minimal root requirement is the same as the minimal word requirement. The claim implicit in (2') is that *any* synchronically recoverable CW root in English is at least bimoraic: since English also contains a large number of non-native roots, mostly from French, Ancient Greek and Latin, these too must be checked against the minimal root requirement. We may divide non-native roots into free and bound roots.

Clearly, the free roots all meet the minimal word requirement (else they would not appear

[2] I follow most work in metrical phonology in replacing the SPE features [±tense] with a more universally applicable difference in length: [-tense] vowels are monomoraic [ɪ], [+tense] vowels are bimoraic [i]. (Liberman & Prince 1977:271)

as free roots) and thus are all minimally bimoraic: *just, form, cry, air, nymph, phone, rhyme, chrome, sex, mix, fuse*.

But the two mora requirement holds of bound roots as well. Sloat & Taylor (1975) provide a useful pedagogical list of classical roots in English. Of the approximately 1,150 root morphemes and allomorphs they list, 97% have two or more moras. The majority of these are CVC, *phil* 'love', *den* 'tooth', though some CVV show up as well, *flu* /fluu/ 'flow' (*fluid, fluent*), *my* /mai/ 'muscle' (*myology, myocardium*).

What about the remaining 3% that consist of less than two moras? These are listed below:

(6) Putative allomorphs of Classical roots in English (Sloat & Taylor 1975)

C_0V:	affable	clitic	epithet	plethora	pus
	analysis	crescent	flatulence	professor	quantity
	butter	dismal	pity	prophet	scilicet
C_0:	arrest	cognition	diarrhea	pregnant	surplice
apostle		cranium	enmity	problem	tmesis
	clandestine	cremate	multiple	remnant	

The roots in (6) would be counterexamples to the minimal root requirement if they were synchronicaly recoverable roots in English. Clearly, however, they are not: *fa* does not constitute a meaningful part of *affable* even for the most erudite of English speakers, nor is *st* a morpheme of *arrest*. Whatever pedagogical use the putative roots in (6) may have for the student of English etymology, they are not morphemes of the language and thus do not constitute counterexamples to the minimal root requirement in (2'). All synchronic lexical roots in English consist of at least two moras. This is the basis of the bimoraic minimal word requirement for the language: words are made from roots, so minimal restrictions on the latter entail minimal restrictions on the former.

English Function Words ≤ μμ

A number of English FWs occur as less than bimoraic in normal speech. These fall into two

classes: the articles *a* and *the*, which are underlyingly monomoraic, and reducible ("clitic") FWs such as *am* ('*m*) and *will* ('*ll*) which are underlyingly bimoraic but are reduced to non-moraic or monomoraic status in speech.

Except in emphatic or hesitating contexts, *a* and *the* occur as the light syllables [ə] and [ðə]. No comparable content words exist. Both words have heavier allomorphs that occur prevocalically, but, interestingly, neither of these is always bimoraic: Kenyon & Knott, for instance, transcribe prevocalic *the* as monomoraic [ðu] not bimoraic [ði]. The prevocalic shape of *a* is [æn] or [ən], but this is not always bimoraic, since the final *n* often syllabifies with the following word: *an ice cream pie* and *a nice cream pie* may be homophonous, [ə.najs.krim.paj].[3]

The emphatic forms of *a* and *the*, [ej] and [ði], are of course bimoraic and do satisfy the minimal word requirement. The point here is that (near-)homophonous content words like *hay* and *bee* may not be reduced to [hə] and [bə], even when unemphatic: words-in-the-Lexicon must be bimoraic.

Non-lexical roots

Since *a* and *the* are both words and roots (just as *eat* and *frog* are both words and roots), they show that both FWs and the roots of FWs may consist of less than two moras. This suggests that English FWs and their roots are not subject to the same minimal requirements as English CWs and their roots. This is the case in language after language, as noted as early as Bloomfield (1933:243-4):

> The roots of a language are usually quite uniform in structure.... In a few languages, such as Chinese, the structure of the roots is absolutely uniform; in

[3] Resyllabification of final *n* in *an* (and *mine*) is long-standing in English. Historically this has given a newt < an ewt, a nickname < an ekename. Misanalysis occured in the other direction as well: an auger < a nauger, an adder < a nadder (Skeat 1910).

others, we find some roots that are shorter than the normal type. It is a remarkable fact that these shorter roots almost always belong to a grammatical or a semantic sphere which can be described, in terms of Enlgish grammar, as the sphere of pronoun, conjunction, and preposition. In German, which has much the same root structure as English [i.e., , the definite article contains a rood [d-], for in the forms *der, dem, den,* and so on, the rest of the word (*-er, -em, -en,* and so on) is in each case a normal inflectional ending, appearing also in the inflectional forms of an adjective like 'red': *rot-er, rot-em, rot-en*. The same applies to the interrogative pronoun 'who?' with forms like *wer, wem, wen*. In Malayan and in Semitic, many words in this semantic sphere have only one syllable [despite a two syllable minimal limit on content words--C.G.], as, in Tagalog, [at] 'and', or the syntactic particles [ang] 'sign of object-expression', [ajl 'sign of predication,' [na] 'sign of attribution.' This semantic sphere is roughly the same as that in which English uses atonic words.

In addition to *a* and *the*, English has a large number of FWs that reduce to monomoraic or non-moraic sequences in normal speech: bimoraic *and* reduces to monomoraic [n] (*Tom 'n' Jerry, Joe 'n' I);* bimoraic *will* reduces to nonmoraic [l] (*You'll go*) and so on. These forms are clear violations of the minimal word/minimal root requirement. Content words do not reduce to monomoraic or non-moraic strings (except, perhaps, in rapid speech). A comparison of homophonous FW/CW pairs makes this clear:

(7) Reducing and non-reducing homophones in English

Function Word	Stressed	Unstressed	Content Word	(Un-)Stressed
be	bi	bɪ	bee	bi
been	bɪn	bɛn	bin	bɪn
but	bʌt	bət	butt	bʌt
can (aux.)	kæn	kən, kn̩, kŋ	can (n.)	kæn
do	du	dʊ, də	dew	dju, dɯ, du
for	fɔr	fɾ	four	for, fɔr
have (aux.)	hæv	həv, əv, v, hæf	have (v.)	hæv[4]

[4] Kenyon & Knott do not give a separate listing for the main verb *have*, suggesting that it has the same full and reduced forms as the modal. In American dialects I am aware of (my own, for instance_ however, main verb *have* does not reduce: *I've a friend in town.

him	hɪm	ɪm, hɪm	hymn	hɪm
I	aɪ	aɪ, əɪ, ə	eye	aɪ
in	ɪn	ɪn, n̩	inn	ɪn
must	mʌst	məst, məs	must (n.)	mʌst
or	ɔr	ɾ	ore	or, ɔr
some	sʌm	səm, sm̩, sə	sum	sʌm
till (conj.)	tɪl	tl̩	till (v., n.)	tɪl
to	tu	tʊ, tə, ə	too, two	tu
we	wi	wɪ, wɨ	wee (adj.)	wi
will (aux.)	wɪl	wəl, əl, l̩, l	will (n.)	wɪl
would	wʊd	wəd, əd, d	wood	wʊd
you	ju	jʊ, jə	ewe	ju, jo
your	jʊr	jʊr, jɾ	yore	jor, jɔr

How the function words in (7) are reduced will be discussed in the following chapter (see also Golston 1990b). For now it is sufficient to observe that none of the content words in (7) have reduced forms, despite the fact that they are segmentally identical to the full forms of the corresponding function words. This is especially noteworthy in cases where the corresponding function word reduces to a monomoraic (fɾ, n̩, ɾ, sm̩, tl̩, etc.) or non-moraic (d, l, v) sequence. The fact that only the content words in (7) lack reduced forms adds support to the 2LH: the reducibility of function words but not of homophonous content words follows if only the latter are subject to minimal word requirements.

The Minimal Affix = μ. I have tried to show that content words and their roots in English are subject to a minimal prosodic requirement to which function words are not subject: words-in-the-Lexicon must be bimoraic, words-in-the-Phrasicon may be monomoraic and even non-moraic. A similar case can be made for derivational affixes (which are subject to a minimal requirement) as opposed to inflectional affixes (which are not). The minimal requirement for derivational affixes is that they consist of at least a light syllable:

(8) English: Min Aff = μ

This minimum requirement applies only to derivational affixes, i.e., to those affixes that are stored in the Lexicon. Affixes in the Phrasicon are not subject to the minimal affix requirement, just as words and roots in the Phrasicon are not subject to the bimoraic minimal word requirement.

I claim in (8) that English has no derivational affixes that consist of less than a mora. To see that this is so, let us begin with prefixes. Marchand (1969) gives what he claims to be an exhaustive list of 65 English prefixes (9):

(9) English prefixes (Marchand 1969)[5]

a-	a-	ante-	anti-	arch-	auto-	be-	bi-
circum-	cis-	co-	counter-	crypto-	de-	demi	di-
dis-	en-	epi-	ex-	extra-	fore-	hyper-	hypo- in-
inter-	intr-	mal-	meta-	micro-	mid-	mis-	mono-
multi-	neo-	non-	pan-	para-	per-	peri-	poly-
post-	pre-	preter-	pro-	pro-	pro-	proto-	pseudo-
re-	retro-	semi-	step-	sub-	super-	supra-	sur-
trans-	tri-	twi-	ultra-	un-	un-	uni-	vice-

Each of the prefixes in (9) is at least monomoraic, in accordance with (8). Similarly for suffixes; of the 80 or so that Marchand lists, all but two (underlined) are moraic:

(10) English suffixes (Marchand 1969)[6]

-able	-acy	-age	-al	-al	-an	-ance	--ancy	-ant	-
ard	-arian	-ary	-ate	-ate	-ate	-ate	-ation	-by	-
cy -dom	-ed	-ed	-ee	-een	-eer	-en	-en	-er -er	-
erel	-ery	-ese	-esque	-ess	-et	-ette	-fold	-ful	-
ful -hood	-iana	-ic	-ician	-ie	-fy	-ine	-ing	-ing	-
ish	-ism	-ist	-ister	-ite	-ity	-ive	-ize	-kin	-
le[7] -le	-less	-let	-ling	-ly		-ment	-mo	-most	-
ness	-ory	-ous	-ship	-some	-some	-ster	-th	-ton	-
ure	-ward	-y	-sy	-ety	-i				

Most of the suffixes in (10) are bimoraic, though a few contain only a single mora. The two apparent exceptions, *-th* and *-ed*, each admit of alternative analyses that render their status as exceptions dubious.

-ed. Marchand distinguishes two suffixes here, the type found in *feathered* and the type found in *palefaced*. Both are derived from the *inflectional* ending found on past participles (and thus exceptions that prove the rule, so to speak) and both have unpredicatble allomorphs in /id/ which satisfy the one-mora minimum. Comparative evidence that both types of *-ed* were originally inflectional comes from other IE languages in which the same alternation between denominal and deverbal adjectives occurs (L. *dentatus* 'toothed' and *amatus* 'was loved'). Internal evidence includes the fact that in Old and Middle English such denominals are also found with the inflectional prefix *ge-* (*gehlidod* 'lidded', *geswurdod* 'sworded'). Synchronic evidence that denominals in *-ed* are still felt to be inflection-like comes from the otherwise peculiar fact that these adjectives (and no others) may be modified with *well-* and *ill-*, two prefixes which otherwise modify only deverbal participles (*well-worn, ill-suited;* but **well-blue, *well-warm, *ill-big*, etc.). That is, the prefixes *well-* and *ill-* clearly subcategorize for verbs (*worn*) not adjectives (**blue*), showing that the *-ed* in *ill-suited* is a verbal not an adjectival suffix.

5 Prefixes that are listed more than once are ones that Marchand treats as separate affixes: *ablaze* vs. *asymmetric; proconsul* vs. *pro-amnion* vs. *pro-British; unfair* vs. *untie*.

6 I have not included alternate spellings here (*-ance, -ence; -ine, -in)*, nor all of the allomorphs of each morpheme (*-ery, -ry; -ify, -fy; -ety, -ity, -dy, -ty*). Marchand also lists a number of what he calls semi-suffixes, none of which violates the min aff requirement: *-like, -worthy, -monger, -way/-ways, -wise, -word/-wright*. I will not discuss these here; the interested reader is directed to Marchand (1969).

7 As Marchand (1969:323) points out, "*-le* is not a derivative suffix proper from existing roots" and is best treated as a recognizable but not segmentable symbolic element of a number of words. "*Twink* is not recorded before 1400, i.e., 500 years later than *twinkle*; *fizzle* is recorded 1532, *fizz* 1665, *quackle* 1564 is oder than *quack* 1617. Many verbs probably never had a simple root without the [l] element, as *drizzle, bustle, hustle, rustle, suffle, shuffle, trickle...*"

Additional evidence that the *palefaced*-type *-ed* suffix is felt to be inflectional rather than derivational comes from the fact that it is added not to true compounds (11a) but to syntactic phrases (usually N-bars consisting of an adjective-noun sequence) (b). This is made evident by the fact that bare nouns generally are not affixed with this type of *-ed* (c):

(11)	a.	*	[[báking-pòwder]$_N$ ed]	'having baking powder'
		*	[[wáll-pàper]$_N$ ed]	'having wall paper'
	b.		[[hèavy hánd]$_{N'}$ ed]	'having heavy hands'
			[[thrèe córner]$_{N'}$ ed]	'having three corners'
	c.	*	[[hánd]$_N$ ed]	'having a hand'
		*	[[córner]$_N$ ed]	'having a corner'

I will assume that the inflectional nature of both types of *-ed* is what allows them to exist as non-moraic affixes, insofar as they are non-moraic.

Many of the oldest and most common *feathered*-type words with *-ed* show an underlying suffix /ɪd/ rather than /d/:

(11)	/ɪd/:	crooked	<	crook	[krʊkɪd]	*[krʊkt]
		wretched	<	wretch	[rɛčɪd]	*[rɛčt]
		ragged	<	rag	[rægɪd]	*[rægd]
		jagged	<	jag	[ǰægɪd]	*[ǰægd]

An underlying suffix /ɪd/ conforms neatly to the one-mora minimum and accounts for the otherwise anomalous data in (11), given that other cases of *-ed*, such as the preterite (underlyingly /d/), never surface as syllabic, even after velars: *looked* [lʊkt], *[lʊkɪd]; *wagged* [wægd], *[wægɪd]. The same underlyingly moraic suffix is found with the palefaced-type words *one-legged, bow-legged*, etc.

Thus, *-ed* is the exception that proves the rule that derivational affixes in English must consist of at least one mora: it is neither clearly derivational (since it has a number of inflectional properties) nor clearly non-moraic (*wretched, one-legged*). Its status as an

exception to the minimal affix requirement is therefore seriously compromised.[8]

-th. This is not a productive affix in English. Marchand's (1969:349) discussion of its history bears this out:

> *-th* is a substantival suffix with a few coinages of doubtful currency. Productive in Old English and Middle English (*depth, health, length, strength, wealth*, etc.), it has in the Modern English period formed deverbal *growth* 1557 and *spilth* 1607. *Breadth* 1523 (no derivative for the present-day speech-feeling) is extended from now obsolete brede (OE brædu), influenced by *length*, and was followed by *width* 1627 (dialectal *widness* with a change of suffix rather than "a literary formation" (OED), as original [i:] had, by 1627, long become a diphthong. *Coolth* 1547 (now chiefly colloquial or jocular) was coined after *warmth*. *Illth* 1860 is "used by and after Ruskin as the reverse of *wealth* in the sense of 'well-being': Ill-being" (OED). Horace Walpole coined the word *greenth*, Wentworth quotes *lowth* as used by Bacon.

In light of this non-productive status, *-th* might well be dismissed as a serious counterexample to the minimal affix requirement: if it is not productive, there is little reason to list it as a separate morpheme in the Lexicon. And if it is not listed in the Lexicon, it will not be subject to a minimal affix requirement.

Note also that the most common words containing *-th* involve an unpredicable ablaut variant of the stem-vowel, indicating that words like *length* and *depth* are not synchronically

---

8 Tom Cornell (p.c.) points out the interesting case of adjectival passives (see Grodzinsky 1990:114-120 for a review and a discussion with regards to Agrammatism). In 'Mary was concerned', the adjectival status of concerned is shown by the following contrast:

a. Mary was concerned about John.
b. ?Mary was concerned by John.

If concerned were a true passive (2) should be grammatical (cf. ? John concerned Mary.). These adjectival passives are not (on most accounts) derived transformationally and should be formed in the Lexicon (V -> Adj) derivationally. If so, they are derivational suffixes in English that consist of less than a mora. As with the *-ed* in *one-legged*, however, this is clearly related to an inflectional affix; this, I would claim, is what allows it to consist of less than a mora.

derived from [*long* + *th*] and [*deep* + *th*] (cp. 12a, b).

(12) a. Ablaut vowels and *-th*

[ɔ]-[ɛ]		[i]-[ɛ]		[aʊ]-[ʊ]	
long	length	deep	depth	wide	width
strong	strength	heal	health	five	fifth
broad	breadth	weal	wealth		

If the alternations in (12) were to be derived synchronically, English would require phonological rules something like those in (13).

(13) Non-rules of English Phonology
a. /ɔ/ -> [ɛ]/ __CC$
b. /i/ ->[ɛ]/ __CC$
c. /aʊ/ ->[ʊ]/ __CC$

The rules in (13), however, are clearly not phonological rules in English, even if we restrict the domain of rule-application to words ending in the morpheme *-th*. (13a) is counter-exemplified by *fourth* (*[fɛrθ]) and *warmth* (*[wɛrmθ]); (b) is counterexemplified by *sixteenth* (*[sɪkstɛnθ]); (c) by *heighth* (*[hɪtθ]) and *ninth* (*[nɪnθ]). Note also that eighth (*[æɪtθ]) does not undergo an /e/ -> [ɛ]/ __CC$ rule. The absence of rules like those in (13) in English makes the synchronic derivation of *depth* from *deep* + *th*, *strength* from *strong* + *th*, *width* from *wide* + *th*, etc. highly unlikely.

Even if *-th* is treated as a synchronic affix of English, its status as a counterexample to the minimal affix requirement is not beyond dispute. Evidence discussed by Goldsmith (1990) suggests that *-th* may consist underlyingly of a CV sequence. Goldsmith notes that *-th* is exceptional not only by its apparently non-moraic status, but also by its ability to occur in the codas of syllables in positions normally restricted to [s, z, d, t]. Goldsmith points out that sequences of obstruents are not allowed word finally in English unless one of the obstruents is an alveolar:

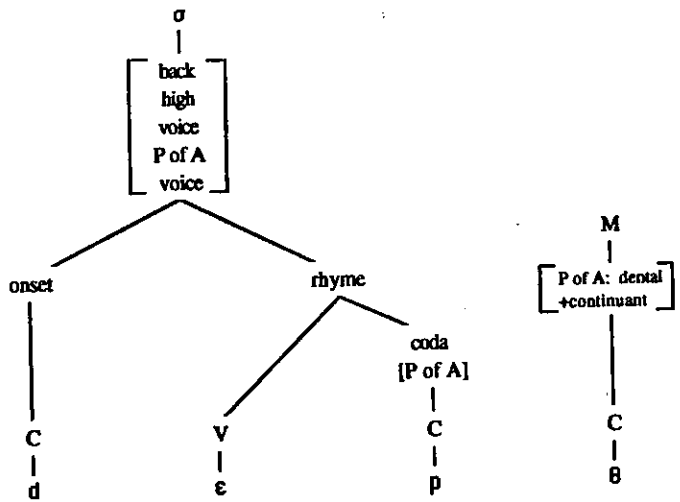
(14) [V C_1 C_2]$_\sigma$ -> (C_1 is alveolar) or (C_2 is alveolar)

That is, given a coda-cluster consisting of two obstruents, one of the obstruents must be [s], [z], [t] or [d]: *æfk, *æθk, *æʃk, *æfp, *æθp, *æʃp, and *ækf, *ækθ, *ækʃ, *æpf, *æpθ, *æpʃ are impossible English morphemes. (Note that 'alveolar' in (14) cannot be replaced by 'coronal' since coronal [ʃ] is not allowed in these positions.) Curiously, θ *does* occur after another non-alveolar obstruent but *only* when it occurs as a separate morpheme: *depth, length, strength,* etc. Goldsmith concludes that the morpheme *-th* somehow licenses [θ] in this position, since [θ] is not licensed in this position in monomorphemic words:

> This terminal θ (that is, one that appears after an obstruent) appears in *fifth, si[ks]th, (eithth), twelfth, hundredth, thousandth, length, strength, depth, width,* and *breadth*, and in no other words in the English language. Clearly there is a generalization here, of which the weakest possible is that the *-th* is a separate morpheme. We may state the connection between the otherwise exceptional distinctive point of articulation and the presence of a distinct morpheme by saying that *the dental point of articulation is licensed word-finally by the suffixal morpheme itself*, rather than by cononical syllable-internal structure or by the regular word-final extrasyllabicity.
>
> (1990:147--my emphasis)

Goldsmith formalizes the licensing of θ in the coda position when it is a separate morpheme as in (15) (= Goldsmith's (44)):

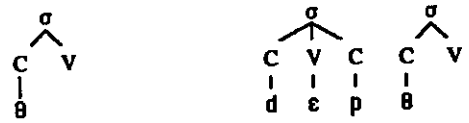
(15) licensing word-final θ in *depth*



In (15), the fact that -θ may occur where it may is a result of 'licensing features' associated with the morpheme; when [θ] is not part of this morpheme, its point of articulation is not licensed, explaining the lack of *monomorphemic* words like dɛpθ, wɪdθ, etc. For Goldsmith, then, it is not anything about the *phonological* shape of -θ that licenses its occurrence after p and k: it is a property of the *morpheme* that licenses these positions.

One problem with Goldsmith's view is that *any* point of articulation could be licensed, even one (say) that does not occur in English (linguo-labial, pharyngeal, etc.); his proposal is clearly too strong in that it allows far more than it needs to. All that is required is (a) that a segment that commonly occurs in onsets and simple rhymes be allowed to occur after [p] and [k] in the coda and, (b) that this be a property of the affix *-th*, since it does not occur with θ morpheme-internally. This may be accomplished with an underlyingly CV form of the suffix in which the V-slot is empty (16a):

(16) a. *-th* as underlyingly CV b. *depth*



The empty V-slot puts θ in the onset of a syllable, where it is licensed by the same licensing that permits words like *thing* and *thick*; the (merely) apparent coda cluster [pθ] in (16b) does not violate the restriction that coda cluster cannot have two non-alveolar points of articulation. Note that segmental material from *dep* (namely the vowel) cannot be associated with the V wihtout crossing the association line linking θ to its C-slot; consequently, the V-slot in *-th* goes unfilled and is lost, leaving the θ to be stray-adjoined into the root-syllable. This analysis is more restricted than Goldsmith's for two reasons. First, it utilizes a formal device (an empty timing-slot) independently established in the literature--Goldsmith himself discusses analyses involving empty C-slots in French, Seri (Marlett & Stemberger 1983) and Ondandaga (Michelson 1985). Second, it is less powerful insofar as it is not capable of introducing (say) a point of articulation not found in English.

How could an underlying form such as the one in (16) come about? Comparative evidence from Dutch and German suggests that the English form once contained an audible vowel, most likely a schwa: cf. Du *zesde* 'sixth', *zevende* 'seventh', achtste *'eighth'*; G zweite 'second', dritte 'third', vierte 'fourth'.

The same underlying form used to allow *-th* to occur after obstruents in a coda qualifies *-th* as underlyingly moraic: the final morpheme in *depth* is underlyingly CV (satisfying the min aff requirement) but surfaces only as θ because its V-slot is not linked to anytthing on the segmental tier.

Neither the *-ed*s in *palefaced* and *feathered* nor the *-th* in *width* and *length* provide solid examples of non-moraic derivational affixes in English. The latter fails to be clearly non-moraic; the former fail to be clearly non-moraic and fail to be clearly derivational. Hollander (1976:177), discussing derivational affixes from Old to Modern English, puts the general case

like this:

> There...seems to be a clear connection between productivity and syllabicity in English derivational morphology. Productive suffixes obviously tend to be syllabic rather than nonsyllabic.

Without convincing counterexamples, then, and in light of the sheer number of clear cases that support it, I conclude that the minimal affix requirement is true of English.[9]

English Inflectional Affixes < μ. English has eleven inflectional affixes, eight of which do not obey the minimal affix requirement required of derivational affixes.

(17) English inflectional affixes

-ed	/d/	bak-ed, claw-ed	past
-ed	/d/	(has) bake-d, (has) kill-ed	perfect participle
-ed	/d/	(was) bake-d, (was) kill-ed	passive participle
-en	/n/	(has) tak-en, (has) see-n	perfect participle
-en	/n/	(was) take-n, (was) see-n	passive participle
-er	/ər/	brown-er, slow-er	comparative
-est	/ɪst/	brown-est, slow-est	superlative
-ing	/ɪŋ/	(is) sink-ing, (is) flow-ing	present participle
-ly	/li/	deep-ly, quick-ly	adverb (adjectival inflection)
-s	/z/	duck-s, dog-s	plural
-s	/z/	eat-s, beg-s	3rd singular
-'s	/z/	duck-'s, dog-'s	possessive[10]

Note that none of the underlyingly non-moraic suffixes in (17) have moraic allomorphs like the [ɪd] allomorph found in *wretched* or *one-legged*. This is seen most clearly with the various *-s* and *-ed* suffixes but is equally true of the *-en* suffixes, which always surface as non-moraic [n] post-vocalically: see / seen, be / been, lay / lain, slay / slain.

English grammatical items are not subject to the same minimal prosodic requirements to which English lexical items are subject. Content words and their roots must be bimoraic, though function words may be monomoraic and may reduce to non-moraic status. Similarly, derivational affixes must be at least monomoraic, whereas inflectional affixes are frequently non-moraic. The 2 Lexicon Hypothesis allows this to be stated in a non-arbitrary fashion: morphemes and words in the Lexicon are subject to minimal prosodic constraints that do not hold for morphemes and words in the Phrasicon.

2.2 *Ancient Greek*

Like English, Ancient Greek (AG) allows no content words that consist of a single light syllable: well-formed lexical words are minimally bimoraic. Again like English, this holds not only of lexical words but of lexical roots. AG content words, like their counterparts in English, do not have 'reduced forms'.

A bimoraic foot can be shown independently to be the foot used in stress and pitch-accent

---

9 Sloat & Taylor (1975) list approximately 165 derivational affixes for English. Of these, five are non-moraic and deserve some discussion:

-l	'small'	gastrula
-sc	'becoming'	nascent
-m	NOUN	baptism, chasm, idiom, poem, sarcasm, spasm
-s	NOUN	chaos, crux, defense, diacoustics, diplomatice, economics, ethics, expense, larynx, linguistics, sex, tax
-t	'one who, that which'	anthropologist, architect, catapult, ventriloquist, dentist, enthusiast, florist, individualist, materialist, poet, suspect

-l and *-sc* are not synchronically recoverable in *gastrula* or *nascent*; other words in which these putative suffixes suggest moraic (*spittle*) or bimoraic (*pubescent*) suffixes. The last three suffixes are best analyzed as parts of other suffixes (following Marchand 1969 and others): *-m* is part of *-ism*, *-t* is part of *-ist*, *-s* is part of *-ics*. Words like *poet* and *suspect*, *sex* and *tax* are monomorphemic and do not have a synchronically recoverable *-t* or *-s*. Sloat & Taylor's list of roots and affixes is primarily pedagogical; more theoretically oriented works such as Aronoff (1976), Selkirk (1982) or Scalise (1984) do not treat *-l*, *-sc*, *-m*, *-s*, and *-t* as affixes.

10 I follow Emonds (1985) and others in considering *'s* an inflectional affix. It is, of course a phrasal rather than a lexical affix, since it is attached to NP rather than to N (see Anderson 1989, Golston 1989).

assignment (Allen 1973; Sauzet 1989; Golston 1990b). Thus, in McCarthy and Prince's (1986) terms, the minimum word equals the minimum foot in AG.

With the exception of a handful of imperatives and participles (on which see below), AG had no monomoraic nouns, verbs or adjectives. Examples of disyllabic and monosyllabic bimoraic nominals are given below:

(18) Ancient Greek minimal nouns (nominative singular)

CVV					
mnáa	'type of currency'	gée	'earth'	noó	'mind (dual)'
CVC(C)					
thrík(s)	'hair'	phlép(s)	'vein'	hál(s)	'salt'

(19) Ancient Greek minimal adjectives (masculine singular unless otherwise noted)

CVV					
sóo(s)	'safe'	eú(s)	'noble'	píoo(s)	'fat'
CVC(C) [11]					
mérop(s)	'mortal'	óinop(s)	wine-colored'	hélik(s)	'bent'

Notice that the minimal monosyllabic word attested for nominals is either CVV or CVCC; no cases of simple CVC nouns or adjectives occur in AG. I will argue that this is due to the extrametricality of word-final consonants in AG (Steriade 1988; Sauzet 1989). CVCV and CVV shaped words are unaffected by final consonant extrametricality; but CVC shaped words violate the bimoraic minimum because the segment that constitutes the second mora is extrametrical: CV(C). AG content words are thus minimally CVV or CVCC.

For morphological reasons, minimal word sized verbs are hard to find in AG: most verbs contain more than two moras because of moraic inflectional endings. Not all of the following,

[11] True CVC(C) adjectives are hard to find; I have included longer words here. The minimal requirement is satisfied trivially.

therefore, are truly *minimal;* this is especially true for verbs of the form CVCV.

(20) Ancient Greek minimal verbs

CVV verbs					
pléi	'it sails'	rhéi	'it flows'	psóo	'I rub'

No CVCC verbs are given because Greek verbal inflection never produces word-final consonant clusters. No lexical adverbs have been given here because they are all derived either from nominal or from verbal roots.

Two sets of counter-evidence to the bimoraic minimal word proposal must be countenanced. The first involves a small number of verbs that have CVC second person singular aorist imperatives:

(21) CVC imperatives (second person singular aorist)

dós	'give!'	thés	'put!'

If final Cs are extrametrical, such forms should be monomoraic and hence ill-formed. Hayes (19??) has proposed that degenerate feet (a monomoraic foot in this case) are allowed in some languages under heavy stress. I will assume that imperatives were contrastively stressed in AG (not an implausible assumption for imperatives) and that this extra degree of stress is what licensed otherwise monomoraic CV(C) syllables.

The second set of counter-evidence consists of a few neuter participles of the form CVC:

(22) CVC neuter participles (neuter nominative and accusative singular)

stán	'having set'	thén	'having placed'	dún	'having shown'

The contrastive stress hypothesis countered for imperatives is surely out of place here; there is no reason to think that participles would acquire any extra degree of stress that might license a monomoraic foot. Steriade (1988) provides compelling evidence, however, that such neuter participles were treated as CVCC with a final [t] as in (23):

(23) CVCC forms of neuter participles

/stánt/	'having set'	/thént/	'having placed'	/dúnt/	'having shown'

Steriade's evidence for the final *-t* in neuter participles comes from the accentuation of polysyllabic neuter participles:

> *-nt* finals surface as *-n*, but words ending in such sequences are accented exactly like those ending in *-ks, -ps* [i.e., in consonant clusters--C.G.]: their recessive accent never retracts beyond the penult. Examples of this type are neuter participles like *paidéu-on* 'educating' (from /paideu-ont/; compare the genitive *paideú-ont-os*) and neuter adjectives like *kharí-en* 'graceful' (from /khari-ent/; compare *kharí-ent-os*). (1988:275)

The CVCC forms in (23) are thus motivated by the accentuation of neuter participles like *paidéu-on*. The late deletion of the post-nasal [t] apparently does not affect the acceptability of these words.

Lexical Roots

AG lexical roots as well as words are subject to the bimoraic minimum. The only difference is that final consonants in roots are not extrametrical (only *word*-final consonants are extrametrical). CVC nominal and verbal roots are extremely common: the roots for some of the nouns, verbs and adjectives mentioned above are given below:

(24) Ancient Greek nominal roots

CVC nouns					
thrikh-	'hair'	phlep-	'vein'	hal-	'salt'
CVV nouns					
mnaa-	'type of currency'	gee-	'earth'	thee-	'serf'

(25) Ancient Greek adjectival roots

CVC					
kak-	'bad'	mak-	'blessed'	bath-	'deep'
CVV adjectives					
soo-	'safe'	eu-	'noble'	pioo-	'fat'

(26) Ancient Greek verbal roots

CVC					
ag-	'come'	tith-	'put'	phob-	'fear'
CVV verbs					
dee-	'bind'	luu-	'loosen'	thee-	'give'

(Again, I have not included roots of adverbs since these are themselves nominal or verbal.)

Bimoraic Lengthening. Additional evidence for a bimoraic minimal word in Greek comes from phonological processes that preserve bimoraicity in derived words. Consider the paradigm for *pod-* 'foot', for instance:

(27) *pod-* 'foot'

Singular		Plural	
nom	póu-s	nom	pód-es
acc	pód-a	acc	pód-as
gen	pod-ós	gen	pod-óon
dat	pod-í	dat	po-si

Note that only the nominative singular form has a long vowel; the lengthening is easily explained by the bimoraic minimal word requirement coupled with final consonant extrametricality. Alveolars delete before */s/* in AG. Consider the effect of such a rule on the nominative singular and the dative plural forms of *póus*. The dative plural resulting from T/D Deletion is a well-formed bimoraic CVCV; the resulting nominative singular, however, is an illicit monomoraic CV(C).

(28) Derivation of nominative singular and dative plural of *póus* 'foot'

	nom sg	dat pl
Underlying Form	pod-s	pod-si
T/D Deletion	ø	ø

	po -s	po -si
Lengthening	poo-s	-------
O-Raising	pou-s	-------

The nominative singular vowel is therefore lengthened (o -> oo) to satisfy bimoraicity. A later rule of O-raising (Sommerstein 1973) raises *oo* to *ou* [uu]. Similar cases can be seen in other monosyllabic forms:

(29)

		pur- 'fire'	*xho-* 'an amount'	*su-* 'pig'
Sing	Nom	púur	khóu-s	súu-s
	Acc	púur	khóu-n	súu-n
	Gen	pur-ós	kho-ós	su-ós
	Dat	pur-í	kho-í	su-í
Plural	Nom	pur-á	khó-es	sú-es
	Acc	pur-á	khó-as	sú-as
	Gen	pur-óon	kho-óon	su-óon
	Dat	pur-óis	?	su-sí

In the nominative and accusative, would-be CV(C) forms are lengthened to CVV(C) to achieve bimoraicity: pur -> puur; khós -> khóos -> khóus; sún -> súun, etc.

The same lengthening processes can be observed in 'clipped' forms of vocatives. Consider the following:

(30) Clipping

báa	<	ba.si.léu	'king-voc'
dií[12]	<	di.í	'Zeus-dat'
máa	<	má.teer	'mather-voc'

The clipped forms on the left are derived as follows: the first syllable of the full word is copied (*ba, di, ma*), lengthened to achieve the requisite bimoraicity (*baa, dii, maa*), and accented:

(31) Derivations of clipped forms

Full Word	ba.si.leu	di.i	ma.teer
Copy first syllable	ba	di	ma
Lengthening	baa	dii	maa
Accentuation	báa	díi	máa

[12] This accent here is not recessive (that would yield *dii*); I have no account of this but it does not materially affect the argument at hand.

A third souce of evidence for this lengthening rule comes from letter-names in AG, which provide interesting support for the claim that the minimal word was bimoraic. All AG letter names had at least two moras, even where the phoneme named was itself inherently *mono*moraic. The 24 letter names[13] are given below, with the Greek letter preceding:

(32) Greek letters and letter-names

α	ἄλφα	álpha	ν	νῦ	núu
β	βῆτα	béeta	ξ	ξεῖ	kséi
γ	γάμμα	gámma	ο	οὖ	óu
δ	δέλτα	délta	π	πεῖ	péi
ε	εἶ	éi	ρ	ῥῶ	hróo
ζ	ζῆτα	zdéeta	ς	σίγμα	sígma, sígma
η	ἦτα	éeta	τ	ταῦ	táu
θ	θῆτα	théeta	υ	ὖ	úu
ι	ἰῶτα	ióota	φ	φεῖ	phéi
κ	κάππα	káppa	χ	χεῖ	khéi
λ	λάμβδα	lámbda, lábda	ψ	ψεῖ	pséi
μ	μῦ	múu	ω	ὦ	óo

Of particular interest are the letters ε and ο (called 'epsilon' and 'omicron' in later Greek). These letters unambiguously represent *mono*moraic vowels[14] [ɛ] and [o] yet have *bi*moraic names [éi] and [óu]. The letter υ is a slightly different case: the letter may represent either a long or a short vowel, but the letter name is long [óu]. Equally suggestive are the letter names for μ and ν (why not [mú], [nú] ?), for α and ι (why not [á], [í] ?) and for π, φ, χ, ξ (why

[13] I omit here other letters such as ϝ digamma, koppa ϙ and sán ϻ (name not certain acc. to L&S) letters which were used only in very early Greek. kóppa is non-problematic, as is dígamma. For the name sán, recall that in early Greek nasals were moraic (Allen 1973:xx), so sán would probably have been bimoraic.

[14] In some early forms of the Greek alphabet, the letters ε and ο were used both for long and short vowels. When the letters η and ω were introduced for bimoraic vowels, ε and ο were used to represent only monomoraic vowels.

not [pf], [phf], etc.?). It seems that a well-formed letter-name in AG is minimally bimoraic.

The bimoraic minimal word requirement, then, is well-substantiated for content words in AG. It will be shown below, however, that an impressive number of both derived and underived function words do not obey this constraint.

Ancient Greek Function Words ≤ μμ. A large number of function words in AG are monomoraic; it is significant that they occur in almost every non-lexical grammatical category. In order not to be circular or arbitrary in defining 'word', I will use Sauzet's (1989) definition of a word: a string of segments associated with a single H tone. This will exclude marginal 'words' such as proclitics and enclitics; I should note, however, that all such clitics are function words and a number of them are monomoraic. Categories which have monomoraic members include:

(33) Monomoraic function words

Prepositions:	prós	'towards'
	sún	'with'
	pró	'in front of'
Articles:	tón	'the' (masc. acc)
	tó	'the' (neut. nom/acc)
	tá	'the' (neut. pl nom/acc)
Pronouns:	sú	'thou'
	sé	'thee'
	hé	'him' (acc)
Possessive pronouns:	sós	'your' (nom, masc)
	són	'your' (nom/acc, neut)
	hós	'his' (nom, masc)
	hón	'its' (nom/acc, neut)
Interrogative pronouns:	tís	'who ?'
	tí	'what ?'

Relative pronouns:	hós	'who' (masc sg)
	hón	'whom' (masc sg)
	hó	'which' (nom/acc sg)
	há	'which' (nom/acc pl)
Small Numbers:	hén[15]	'one' (neut nom/acc)
Conditional Particles:	án	'then'
Conjunctions	dé	'and'
Temporal conjunctions:	prín	'until'
Particles	gár	'for'
	má	'verily'
	mén	'surely'

The words above violate the bimoraic minimal word requirement in one of two ways. CV(C) words such as *mén* 'surely' are monomoraic because of their extrametrical final consonants. CV words such as *dé* 'and' and *tí* 'what' are monomoraic regardless of extrametricality.

Derived monomoraic function words. Itô (1989, 1990) has argued for Japanese that minimal word requirments are met only by derived words: monomorphemic words that do not undergo phonological rules may freely violate minimal word restrictions but derived words may not. Such an explanation is not viable for Ancient Greek because of the existence of a number of *derived* monomoraic function words.

AG pronouns have inflectional suffixes similar to those of nouns. In *tís* 'who, what', for example, the root *tin-* is clearly discernible and the case endings are essentially those of third declension nouns.

---

15 The nom masc sg form is [héis] < /hens/. Notice that compensatory lengthening of /e/ occurs when /n/ deletes. This does not occur with all function words, as will be seen below in discussing [tís] (< /tin-s/).

(34) *tin-* 'who' (interrogative pronoun)

	masculine/feminine	
	sg	pl
nom	tí-s	tín-es
acc	tín-a	tín-es
gen	tín-os	tín-oon
dat	tín-i	tí-si

AG /n/ deletes before an */s/* (nominative singular, dative plural). Thus CV(C) *tís* is a derived form, in violation of Itô's requirement that derived forms obey minimal word constraints. (A number of other derived function words are given below, (35) - (37).) For AG, then, the only words that are not subject to minimal word constrains are *function words*--i.e., words in the Phrasicon. The derived/underived distinction is not an issue in this language.

Non-lexical roots. Recall that lexical roots as well as lexical words obey the bimoraic minimum in AG: nonimal and verbal roots of the shape CV or C are unattested. Since function words do not obey bimoraicity, we might expect their roots not to obey it either. That is, just as both lexical words *and roots* obey bimoraicity, non-lexical words *and roots* might fail to obey it. And this is the case. Stripping off inflectional material (and an alternating o/e theme vowel that is not part of the root) for a number of function words reveals monomoraic and even non-moraic roots. The relative pronoun, for instance, is built on a root *h-*[16]:

(35) *h-* 'who' (relative pronoun)

		sg	pl
masc:	nom	h-ós	h-oí
	gen	h-óu	h-óon
	dat	h-óoi	h-óis
	acc	h-ón	h-oús

16 For clarity of expostion I have not fully parsed the inflectional elements here and in the following paradigms: tbey consist of a theme vowel (*-o-, -ee-, -a-* depending on gender and number) and a fused person-number-case marker. *hós* is thus properly h + ó + s (root + theme vowel + masc-nom-sg).

fem:	nom	h-eé	h-aí
	gen	h-ées	h-óon
	dat	h-éei	h-áis
	acc	h-eén	h-aás
neut:	nom	h-ó	h-á
	gen	h-óu	h-óon
	dat	h-óoi	h-óis
	acc	h-ó	h-á

Similarly, the inflected forms of the definite article are derived from a non-moraic root *t-*, followed by theme vowel and person-number-case:

(36) *t-* 'the' (definite article)

		sg	pl
masc:	nom	(h-o	h-oí)[17]
	gen	t-óu	t-óon
	dat	t-óoi	t-óis
	acc	t-ón	t-oús
fem:	nom	(h-ee	h-aí)
	gen	t-ées	t-óon
	dat	t-éei	t-áis
	acc	t-eén	t-aás
neut:	nom	t-ó	t-á
	gen	t-óu	t-óon
	dat	t-óoi	t-óis
	acc	t-ó	t-á

AG pronouns provide additional cases of non-moraic roots. The oblique singular paradigm consists of a 1st, 2nd or 3rd person root (*em-, s- and h-*, respectively) plus *-óu* (gen), *-oí* (dat), *-é* (acc).[18]

(37) *Pronominal Roots*

	singular:	*plural:*

17 The nominative forms of the masc and fem sg and plural are suppletive.

23 I have not attempted to sort out theme vowels here, as they are fairly fused with the person-number-case suffixes. The pronominal roots are invariant in both singular and plural and this is all that is of concern presently.

	1	*2*	*3*	*1*	*2*	*3*
	[egoó	sú	---][19]	heem-éis	hum-éis	sp^{h}-éis
gen	em-óu	s-óu	h-óu	heem-óon	hum-óon	sp^{h}-óon
dat	em-oí	s-oí	h-oí	heem-ín	hum-ín	sp^{h}-ín
acc	em-é	s-é	h-é	heem-áas	hum-áas	sp^{h}-áas

The plural paradigm is similar, consisting of root (*heem-*, *hum-*, or *sph-*) plus *-éis* (nom), *-óon* (gen), *-ín* (dat), *-áas* (acc).

The AG bimoraic minimal word requirement, although exceptionless for content words and their roots, has a large number of exceptions among the class of function words (*dé* 'and') and their roots (*t-* 'the'). The 2 Lexicon Hypothesis allows us to model this situation straightforwardly: minimal prosodic requirements are met only by morphemes in the Lexicon. Items in the Phrasicon may violate them freely, both underlyingly (*tí* 'what') and in derived forms (*tís* 'who' < /tín-s/). The following sections will show that this holds for AG affixes as well: as with English, derivational affixes are subject to a minimal affix requirement that inflectional affixes escape.

The Minimal Affix in Ancient Greek. AG provides a striking example of minimal-affix differences for derivation and inflection: no AG derivational affix consists of less than a mora whereas non-moraic affixes are common in all types of inflection. The minimal affix in AG, then, is (as in English) as follows:

(38) AG: Min Aff = μ

To reiterate the general discussion: while a minimal-affix requirement may be posited for derivational affixes in AG, it does not extend to inflectional affixes. This generalization provides support for the 2 Lexicon Hypothesis in the following way: the apportionment of roots, words and affixes between the Lexicon and the Phrasicon provides the exact division needed to state minimal prosodic requirements: elements of the Lexicon are subject to such

[19] These forms are suppletive.

restrictions, those of the Phrasicon are not. More concretely, affixes that change nouns to adjectives, verbs to nouns, nouns to adverbs, etc., must consist of at least a mora, while plural markers, tense/aspect markers, person/number markers need not.

(39) Minimal prosodic requirements in Ancient Greek

Lexicon	*Phrasicon*
Affix ≥ μ	---
Word, Root ≥ μ μ	---

Smyth (1920) provides a fairly exhaustive list of AG derivational affixes. AG had a great number of these and the fact that none of them consists of less than a syllable is evidence for the claim that the minimal derivational affix in AG was a mora. As with English, I begin with derivational prefixes:

(40) Ancient Greek prefixes

an-	'not'	an-áksios	'unworthy'
heemi-	'half'	heemi-thneés	'half-dead'
dus-	'ill, un'	dus-tukheés	'unfortunate'
a-	'very'	a-teneés	'very stretched'

Following Smyth, I have listed as prefixes only bound lexical formatives; 'prefixes' that also occur as separate words (*prós-* 'towards', *eu-* 'well', *pent-* 'five', etc.) will be treated (in Chapter 3) as members of compounds rather than as prefixes.

AG had a much larger number of suffixes than prefixes. I will present these according to the lexical feature of the affix: first, noun-forming suffixes, then adjective-forming, verb-forming and adverb-forming suffixes. Smyth tentatively divides AG suffixes into primary (cf. Level 1 in English) and secondary (cf. Level 2 in English) suffixes; since the distinction is irrelevant to the issue at hand, I have collapsed the two classes in the data that follow. Smyth categorizes noun-forming suffixes semantically into agentives, abstract substantives, patronymics, etc. I will follow him here for the sake of exposition, but the point of the present section is not semantic: all that is of interest here is the lack of derivational affixes consisting

of less than a mora. Often the example Smyth provides does not exactly match the citation form of the suffix: e.g., the agentive suffix *-taa-* occurs in the nominative singular form *kri-teé-s* 'judge' as *-teé-*. I will not discuss the many phonological and morphological rules that apply in these cases--it will be sufficient to show that in each case both forms consist of at least a mora. In a few cases, one or the other of these forms consists of a single consonant--these cases will be discussed in detail to show that they do not constitute actual counter-examples to the claim that the minimal derivational affix in AG is moraic.

Noun-forming suffixes. Category names are all those of Smyth; I have supplied only one example for each affix, where Smyth lists a large number. Final *-s* is an inflectional (nominative singular) suffix.

(41) Agentives

-eu	graph-eú-s 'writer'	*-taa*	kri-teé-s 'judge'
-teer	do-teér 'giver'	*-teiraa*	soó-teira 'savior'
-tid	hik-é-tid-os 'of a suppliant'	*-tor*	rheé-toor 'orator'
-triaa	poieé-tria 'poetess'	*-trid*	aulee-tríd-os 'of a flute-girl'
-tro	iiaa-tró-s 'physician'		

(42) Names of actions and abstract substantives

-ti	pís-ti-s 'faith'	*-si*	poíee-si-s 'poetry'
-siaa	dokima-sía 'examination'	*-tu*	ás-tu 'city'
-mo	dioog-mó-s 'pursuit'	*-maa*	gnoó-mee 'knowledge'
-maa[20]	tól-ma 'daring'	*-es*	dé-os 'fear'
-iaa	man-ía 'madness'	*-o*	arkh-ó-s 'leader'
-a	arkh-eé 'beginning'	*-ad*	tri-ád-os 'of a triad'
-iaa	aleéthe-ia 'truth'	*-iaa*	eu-daimoon-ía 'happiness'
-sunaa	dikaio-súnee 'justice'	*-teet*	philó-teet-os 'of friendship'

(43) Result of action

-es	gén-os 'race' (/genes-os/)
-mat	gram-ma 'letter' (/gram-mat/)

(44) Instrument or means of action

-tro	aro-tro-n 'plough'	*-th-ro*	kléi-thro-n[21]
-teer-i	po-teér-io-n 'cup'[22]	*-eio*	troph-éia 'pay for rearing'
-ro	pte-ró-n 'wing'	*-traa*	mák-traa 'kneading-trough'

(45) The person concerned

-eu	gram-mat-eú-s 'secretary'	*-taa*	naú-tee-s 'sailor'
-iaa	hiére-ia 'priestess'	*-id*	pharmak-íd-os 'of a witch'
-tid	oiké-tid-os 'of a house-maid'	*-ittaa*	thée-tta 'female serf'
-ainaa	lé-aina 'lioness'		

(46) Gentiles or place names.

-eu	Platai-eú-s 'Plataian man'	*-taa*	Sikeli-oó-tee-s 'Sciliote man'
-ios	Atheená-ios 'Athenian man'	*-iaa*	Atheena-ía 'Athen. woman'
-tid	Sikeli-oó-tid-os 'Sciliote woman's'	*-id*	Platai-íd-os 'Plataian woman'

(47) Patronymics

-adaa	Thesti-ádee-s 'son of Thestios'	*-daa*	Boreaádee-s 's.o. Boreaas'
-idaa	Tantal-ídee-s 's.o. Tantalos'	*-iadaa*	Persee-iádee-s 's.o. Perseus'
-iion[23]	Kron-iion-os 'of the s.o. Kronos'	*-ideo*	adelph-idou-s 'nephew'
-ad	Thesti-ad-os 'd.o. Thestios'	*-id*	Boreá-d-os 'd.o. Boreaas'
-id	Tantal-íd-os 'of the d.o. Tantalos'	*-iad*	Persee-ídos 'd.o. Perseus'
-ioonaa	Akris-ioónee 'd.o. Akrisios'	*-iinaa*	Adreest-iínee 'd.o. Adrastos'
-ideaa	adelph-idée 'd.o. sibling, niece'		

(48) Place

-io	Dionuús-io-n 'temple of Dionysus'	*-oon*	andr-oón 'apartment for men'
-iitid	androon-íitis 'apartment for men'	*-ooniaa*	rhod-ooniaá 'rose-bed'
-traa	orkheé-s-traa 'dancing-place'		

20 The two suffixes -ma- listed here differ in whether their nominative singular has a long (gnoó-mee) or a short (tól-ma) vowel.

21 As Smyth points out (1929:§832) *-th-* here is probably part of the root. The suffix is therefore properly *-ro*.

22 Properly two suffixes, agentive *-teer* and the normally adjectival *-io* 'pertaining to'.

23 A poetic form that also appears with a short i as -ion-, depending on the requirements of the meter (Smyth 1920:§845.5).

(49) Diminutives

-io	paid-ío-n 'little child'	-id-io	ksiph-íd-io 'small sword'
-ar-io	paid-ár-io-n 'little child'	-ud-rio	mel-údrio-n 'little song'
-ullio	ep-úllio-n 'little epic'	-isko	anthroop-ísko-s 'manikin'
-id-eu	luk-id-eú-s 'wolf's whelp'	-ikho	ortál-ikho-s "young bird'
-iskaa	paid-ískee 'little girl'	-aknaa	pitháknee 'wine-jar'
-id	hamaks-íd-os 'of a little wagon'	-ikhnaa	kulíkhnee 'little cup'

80 of the 81 noun-forming suffixes in (41) - (49) consist of at least one mora. Moreover, the one sub-syllabic affix that Smyth gives, the patronymic *-d-* in *Boreádos* 'daughter of Boreaas', is better analyzed as an underlying */ad/*, making the generalization exceptionless. I turn now to the evidence for such an analysis.

Consider the full list of Patronymics containing a [d] that Smyth offers:

(50) Patronymics formed with [d] (Smyth 1920:§845.1-2)

	son-of		daughter-of	parent
daa	Boreá-dee-s	*d*	Boreá-d-os	Boreaa-
adaa	Thesti-ádee-s	*ad*	Thesti-ad-os	Thestio-
idaa	Tantal-ídee-s	*id*	Tantal-íd-os	Tantalo-
	Kekrop-ídee-s		Kekrop-íd-os	Kekrop-
	Oine-ídee-s		Oinee-íd-os	Oineú-
	Leeto-ídee-s		Leetoo-íd-os	Leetoo-
iadaa	Persee-iádee-s	*iad*	Persee-íd-os	Perseu-
	Phereet-iádee-s		Phereet-iád-os	Phereet-
	Telamoon-iádee-s			Telamoon-
ideo	adelph-idou-s	*ideaa*	adelph-idée	adelpho-
iion[24]	Kron-iion-os			Kron-
		ioonaa	Akris-ioónee	Akris-ios
	iinaa		Adreest-iínee	Adrast-os

Smyth himself notes that "stems in *o* drop *o*; stems in *eu* (*eeu*) drop *u*; stems in *oi* (*ooi*) drop *i* " in forms like Thesti(o)ádees, Oine(u)ídees. Thus there is evidence for a rule that deletes a stem-final vowel before a vowel-initial suffix as in (51):

[24] A poetic form that also appears with a short i as -ion-, depending on the requirements of the meter (Smyth 1920:§845.5).

(51) Vowel-Deletion

V -> Ø / __ + V

Vowel-Deletion applies to stems like Thestio-, Oineu- and adelpho- as in (52). I have separated morphological and phonological processes that take place in the lexicon from those that take place later.

(52) Sample derivations (all nominative singular)

Derivation	*Der. Affixation*	[Thestio + adee]	[Oineu + idee]	[adelpho + idou]
	Vowel Deletion	[Thesti + adee]	[Oine + idee]	[adelph + idou]
Inflection	*Inf. Affixation*	[Thestiadee + s]	[Oineidee + s]	[adelphidou + s]
	Accentuation	[Thestiádees]	[Oineídees]	[adelphídous]

Vowel-deletion also deletes the stem-final (long) vowel of *Boreaa-* before the suffix *-adaa*, as shown in (37):

(53) Derivation

Derivation	*Der. Affixation*	[Boreaa + adee]
	Vowel Deletion	[Bore + adee]
Inflection	*Infl. Affixation*	[Boreadee + s]
	Accentuation	[Boreádees]

The *a* that appears in *Boreádees*, then, may be interpreted as part of the suffix rather than as part of the stem, bringing it in line with the 80 noun-forming suffixes that consist of at least one mora. Note that both the suffix *-adaa-* and the Vowel Deletion rule are independently needed as the derivation of *Thesti(o)-adee-s* makes clear.

This brief discussion does not explain all of the data in (50), of course. *Leeto-ídee-s* 'son of Leetoo' and *Leetoo-íd-os* 'of the daughter of Leetoo' should undergo Vowel-Deletion but do not--and the former inexplicably shortens its vowel (oo -> o). *Oine-ídee-s* 'son of Oineús' and *Oinee-íd-os* 'of the daughter of Oineús' also do not undergo Vowel-Deletion and the latter inexplicably lengthens its vowel (e -> ee). Also, *Akris-ioónee*, from *Akrisio-*, seems to require

deletion of both the final vowels in the stem. A treatment of these cases follows.

Leetoídees, Leetooídos, Oineídees and *Oineeídos* may be explained on independent grounds: as Smyth points out (1929:§848), 'most genuine patronymics are poetical' and thus subject to well-formedness constraints imposed by meter. AG epic meter allows only dactyls (– ∪ ∪) and spondees (– –). Since no combination of these feet produces a single light syllable (∪) between two heavies (– –), words with such stranded short syllables (– ∪ –) cannot be used. Now consider the forms one would predict if the Vowel-Deletion rule posited above were to apply everywhere:

(54) Derivations

Derivation	*Affixation*	[Leetoo + idee]	[Leetoo + id]
	Vowel-Deletion	[Leet + idee]	[Leet + id]
Inflection	Affixation	[Leetidee + s]	[Leetid + os]
	Accentuation	[Leetídees]	[Leetidos]

**Leetídees* (hypothetical) has a structure that is metrically ill-formed in a dactylic hexameter (– ∪ –), whereas *Leetídos* (attested) forms a dactyl (– ∪ ∪). If [Leetoo + idee + s] were not to undergo any rules, however, another metrically ill-formed structure would be produced: **Leetooidees* (– – ∪ –). The only way to make [Leetoo + idee + s] metrically acceptable is to shorten, rather than delete the stem-final vowel: [Leeto + idee + s] (– ∪ ∪ –). A similar case can be made for *Oineídees*:

(55) Derivations

Derivation	*Affixation*	[Oineu + idee]	[Oineu + id]
	Vowel-Deletion	[Oine + idee]	[Oine + id]
Inflection	Affixation	[Oineidee + s]	[Oineid + os]
	Accentuation	[Oineídees]	[Oineídos]

Oi.ne.í.dees is metrically well-formed (– ∪ ∪ –) since it contains a dactyl and a heavy syllable

that can be used as the beginning of the next foot, whereas **Oi.ne.í.dos* (– ∪ ∪ ∪) is ill-formed because of the three adjacent lights. Hence **Oineídos* is ruled out on metrical grounds; the form is made metrically acceptable by lengthening the remaining stem-final vowel *e* to *ee*, yielding a metrically acceptable *Oi.nee.í.dos* (– – ∪ ∪). Similar lengthening can be observed in *Per.see.i.á.dees* (– – ∪ ∪ –) and *Per.see.í.dos* (– – ∪ ∪), from (short-e) *Perseu-*, avoiding **Perseiádees* (– ∪ ∪ ∪ –)and **Perseídos* (– ∪ ∪ ∪) with their sequences of three lights.

(The suffix in *Akrisioónee* (from *Akrisio-*), which Smyth treats as *ionaa*, is better analyzed as *oonaa*, the *i* belonging to the stem. Vowel Deletion deletes the final *o* of the stem, leaving [Akrisi + oonaa], which surfaces correctly as *Akrisioónee*.)

To conclude, all of the patronymioc suffixes are at least monomoraic, making the generalization that noun-forming suffixes in AG consist of at least a mora exceptionless.

Adjective-forming suffixes. Smyth lists twenty-seven adjective-forming suffixes, all of them minimally monomoraic; these are given below for completeness but require ao discussion.

(56) Adjective-forming suffixes (Smyth 1920:§858)

-o, -aa[25]	leuk-ó-s 'bright'	-io, -iaa	ág-io-s 'sacred'
-es	pseud-és 'false'	-lo	dei-ló-s 'cowardly'
-a-leo	thars-aléo-s 'bold'	-mo	ther-mó-s 'warm'
-i-mo	dók-imo-s 'approved'	-mon	mneé-mon 'mindful'
-no	dei-nó-s 'fearful'	-ro	ekhth-ró-s 'hostile'
-u	heed-ú-s 'sweet'	-oodes	prep-oódes 'proper'
-io, -iaa	tím-io-s 'worthy-masc'	-ent[26]	kharí-ent-os 'graceful-gen'
-eo	khrusús-eo-s 'golden'	-ko	manti-kó-s 'prophetic'

25 When two suffixes are given, the first is masculine, the second feminine.

26 The adjective *páas* 'all' seems to have a suffix *-nt-* (cf. pa-nt-ós 'all-gen'). The -nt- in this case is probably best treatedas part of the root (*pa* is not itself a root). Other adjectives all seem to have *-ent-*: *pter-ó-ent-os* 'winged-gen', *phoonee-ént-os* 'voiced-gen', *dakru-o-ént-os* 'tearful' (Smyth 1920:§299).

-ako	Korinth-akó-s	-iko	basil-ikós 'royal'-s
-lo	apat-ee-ló-s 'deceitful'	-mo	hébd-o-mo-s
-si-mo	khreé-si-mo-s 'useful'	-no	skotei-nó-s
-ino	líth-ino-s 'of stone'	-ro	phobe-ró-s
-teer-io	hormee-teér-io-n 'starting-place'		

Insofar as this is a complete list of adjective-forming suffixes in AG, it lends support to the claim that the minimal derivational affixes in AG is monomoraic. To complete the picture, verb- and adverb-forming suffixes must be included.

Verb-forming suffixes. AG verb-stems are formed in one of three ways: (1) directly from verbal roots, (2) directly from nominal stems, and (3) by affixing verb-forming suffixes to nominal stems.

Verb-stems formed directly from verbal root involve no (derivational) affixation and thus do not bear on the minimal size of affixes in AG. Examples are given below--note that all affixes here are inflectional and thus not subject to the minimal affix requirement ex hypothesi.

(57) Verb-stems consisting only of verbal roots

blep-s-oo	'I will see'	see-Fut-1sg
páu-e	'stop!'	stop-2sgImperative
é-*luu*-s-as	'you loosened'	Past-loosen-Aorist-2sg

Verb-stems formed directly from nominal stems also involve no derivational suffixes. Examples are given below:

(58) Denominal verbs

Verb	Nominal stem	Noun
oiké-oo 'dwell'	oiko/oike-	óiko-s 'house'
douló-oo 'I enslave'	doulo-	dóulo-s 'slave'
basileú-oo 'I am king'	basileu-	basileú-s 'king'
dakruú-oo 'I weep'	dakru-	dákru 'tear'
tiimá-oo 'I honor'	tiimaa-	tiimeé 'honor'

AG does have some denominal verb-forming suffixes and they are all at least monomoraic. Most if not all seem to have been formed by analogy with pre-existing forms

like those directly above. From stems ending in *-e* (eg, *oike-* 'house') comes a verb-forming suffix *-e-*; from stems ending in *-o-* (eg, *doulo-* 'slave') comes a verb-forming suffix *-o-*; and so on. These verb-forming suffixes are best seen on consonant-final roots and vowel-final roots that have a different final vowel in their nominal stem than they have in their verbal stem:

(59) Verbal suffixes and consonant-final roots

Suffix	Example	Root
-e-	martur-é-oo 'I bear witness'	martur-
-o-	mastíg-ó-oo 'I whip'	mastig-
-eu-	paid-eú-oon 'I educate'	paid-

(60) Verbal suffixes and vowel-final roots

Suffix	Example	Root
-o-	zdeemi-ó-oo 'I punish'	zdeemiaa-
-eu-	boul-éu-oo 'I counsel'	boulee-

Similarly, verb-stems ending in *-azd-* (eg, *harpázd-oo* 'I seize') and *-izd-* (*elpízd-oo* 'I hope')[27] gave rise to the suffixes *-azd-* and *-izd-*[28]. Denominal verbs formed with *-azd-* and *-izd-* are given below:

(61) Denominal *-izd-* and *-azd-*

agor-ázd-oo	'I buy'	< agorá 'market'
anank-ázd-oo	'I compel'	< anánkee 'necessity'
atim-ázd-oo	'I dishonor'	< átimo-s 'dishonor'
hybr-ízd-oo	'I am angry'	< hybri-s 'anger'
nom-ízd-oo	'I consider'	< nomo-s 'custom, law'

[27] Such stems were themselves derived originally from stems ending in -d or -g: elpízdoo < elpid-ioo, arpázdoo < arpag + ioo (Smyth 1920:§866.6).

[28] Buck (1933:§360.2) points out that '*-ázdoo* is more common from *aa*-stems and neuter n-stems, *ízdoo* from other stems'. This suggests that the suffix might properly be *-zd-* and that the *a* or *i* is a property of the stem. Several observations argue against this: first, stems ending in *-aa* show up as *-azdoo* not **-aazdoo*. Second, consonant-final stems take *-izdoo* not *-zdoo*. Third, stems ending in final vowels other than *a* generally take *-izdoo* (nomo- gives *nomízdoo* not **nomózdoo*); as Smyth points out, 'verbs in *-ezdoo*, *-ozdoo*, and *-uzdoo* are rare' (1920:866.6b).

teikh-ízd-oo	'I fortify'	< téikho-s 'wall'
helleen-ízd-oo	'I speak Greek'	< hélleen 'Greek'

Note that the final vowel (long or short) of the stem is deleted when *-azd-* and *-izd-* are added; this is the same Vowel Deletion rule that applies in patronymics (q.v.).

Note on 'Root Determinatives'. 'Root determinatives' constitute a coherent class of prima facie counterexamples to the claim that AG derivational affixes must consist of at least a syllable. Smyth (1920:§832) defines root-determinative as follows: 'A consonant standing between root and suffix (or ending), and not modifying the meaning of the root, is called a *root-determinative*.' Examples include:

(62) Root-Determinatives in AG

-t^h-	bá-t^h-ro-n	'pedestal'	< ba- 'go'
	és-t^h-oo	'I eat'	< ed- 'eat'
	pleé-t^h-oo	'I am full'	< plee- 'full'
	pleé-t^h-os	'crowd'	
	plee-t^h-óree	'satiety'	
	stá-t^h-mo-s	'day's journey'	<sta- 'stay, stand'
	stá-t^h-mee	'a rule'	
-k^h-	smeé-k^h-oo	'I wipe'	< sma- 'wipe'

If root-determinatives are derivational suffixes, the generalization that derivational suffixes are minimally monomoraic cannot hold.

A number of considerations, however, suggest that root-determinatives are best treated as part of the root (as the name suggests). First, root-determinatives always occur immediately next to the root--no other affixes occur between root and root-determinative. Second, the appearance of root-determinatives is idiosyncratic to roots rather than to grammatical category. Third, root-determinatives add no meaning to a root. Fourth, they never change the lexical category (N, A, V, Adv) of a root. Fifth, they are not reconstructable as affixes for other Indo-

European languages (Smyth 1920:§832).[29] All of these properties are predicted if root-determinatives are parts of the roots.

A related case involves the apparent insertion of *-s-* and *-t-*. Examples are given below:

(63)

-s-	sk^bi-s-mó-s	'cleaving'	< sk^bid- cleave'
	spá-s-ma	'spasm'	< spa- 'rend'
	kéleu-s-ma	'command'	< keleú- 'command'
	mía-s-ma	'stain'	< miaín- 'stain'
	spa-s-mós	'spasm'	< spa- 'rend'
	keleu-s-mós	'command'	< keleu- 'command'
	dú-s-mee	'setting'	< du- 'set'
	keleu-s-teés	'signal-man'	< keleu- 'command'
	ork^hee-s-teés	'dancer'	< ork^hee- 'dance'
	duná-s-tees	'lord'	< dun- 'power'
	dra-s-teérios	'efficacious'	< dra- 'do'
	$ork^heé$-s-traa	'dancing-place'	< ork^hee- 'dance'
	plee-s-mónee	'fulness'	< plee- 'full'
-t-	ep^h-e-t-meé	'command'	< epi-hiee- 'send'
	lái-t-ma	'depth of the sea'	< (?) laim- 'throat'
	au-t-meé	'breath'	< aee- 'blow'

Whereas root-determinatives are best treated as part of the root, *-s-* and *-t-* are best treated as part of the suffix. Similar considerations apply to *-s-* and *-t-* as apply to root determinatives: they co-occur with certain suffixes rather than with lexical categories suggesting that they are parts of those suffixes rather than suffixes themselves that select for N, A or V. They add no meaning to either the root or the suffix. They do not change the lexical category (N, A, V) of the root or affect the lexical category of the suffixed word. They are not reconstructable as affixes for other Indo-European languages.

---

[29] Cf. also the putative suffix -gg- in *phálanks* 'phalanx'(< *phalagg-s*), *sálpinks* 'trumpet' (< *salpigg-s*), *lárunk-s* 'larynx' (< *larygg-s*) (Smyth 1920:§864.11). Smyth claims *-gg-* denotes 'something hollow', but this seems to be a case of sound-symbolism at most (*phalank-s* is from *phallang-io-n* 'spider'): in any case, stripping off *-gg-* as a suffix leaves no attested roots in these cases: **phala-*, **salpi-*, **lary-*.

> 'This parasitic letter [-s-] spread from the perfect middle, where it is properly in place only in stems in *t, d, t*^{h}, or *s*....This *s* appears before many suffixes, and usually where the perfect middle has acquired it.... In a few words *t* is inserted before the suffixes *mo, ma, mee, meen*...In *eret-mó-n* 'oar' the *t* may be part of the verb stem...and have spread thence to the other words' (Smyth 1920:§836-7).

There is thus no reason to assume that the *s*s and *t*s that occur above are separate morphemes; thus they constitute no counterexamples to the claim that all affixes in the Lexicon are minimally bimoraic.

A similar case can be made for a few suffixes with *-d-*: in *outi-d-anó-s* 'a nobody', *rhiig-e-d-anó-s* 'chilling', *allo-d-apó-s* 'foreign', etc. *-d-* is best analyzed as part of the following suffix (Smyth 1920:§863b.1). *-d-* occurs only in conjunction with other suffixes (*-dano-, -dapo-, -daa-, -dio-, -doon-*, and *-doonaa-*)[30]; it adds no meaning to words in which it occurs and does not change the lexical category of roots or stems it is attached to; and is not reconstructable for other IE languages.[31]

This completes the discussion of derivational affixes in Ancient Greek. The 130 or so derivational affixes include none that consist of less than a mora. This provides strong support for the minimal affix requirement in AG.

Ancient Greek Inflectional Affixes < μ. In sharp contrast to derivational affixes, a good number of inflectional affixes in AG consist of a single consonant. As in English single-

30 For the putative occurance of -d- in patronymics (Smyth 1920:§845.1) see discussion above under that heading.

31 Buck (1933) mentions *-t-* suffixes (§475), *-d-* suffixes (§491-2), *-k-* suffixes (§501) for AG. In each case these occur either as part of a larger (syllable sized) suffix (*-t-, -d-*) or are analyzed as part of the root (*-k-*). Suffixes with *-nt-* (§477) are derived from the inflectional -nt- of the active participle--cf. English adjectival passives, discussed above.

consonant affixes in AG include the dentals *s, d,* and *n*. The inflectional affixes in AG may be divided into adverbial, nominal, adjectival, verbal and comparative/superlative affixes.

Adverb-forming suffixes. AG has a number of suffixes that mark denominal adverbs. Smyth lists the following:

(64) Adverb-forming suffixes (Smyth 1920:§341 ff)

		Place	
-i	ofko-i 'at home'	-then	ofko-then 'from home'
-de	ofka-de 'homeward'	-thi	állo-thi 'elsewhere'
-se	állo-se 'elsewhither'	-si	Atheénee-si 'at Athens'
-zde	Atheéna-zde 'to Athens'	-óu	hom-óu 'at the same place'
		Manner	
-oos	kak-óos 'badly'	-a	tákh-a 'quickly'
-akis	poll-ákis 'very often'	-deen	sulleéb-deen 'in short'
-don	skhe-dón 'almost'	-ei	pandeem-ei 'in full levy'
-te	hó-te 'when'	-ti	ethelon-ti 'willingly'
-sti	helleeni-sti 'in Greek fashion'		

In what sense are these adverb-forming suffixes *inflectional*? Although they do not form parts of larger paradigms, they do close-off the stem to derivational morphology: none of the words above undergoes derivational affixation . More importantly,

> adverbs, like prepositions and conjunctions, were originally case forms, made from the stems of nouns and pronouns... It is sometimes uncertain whether we should speak of *adverbs* or of *nouns with local endings*.(Smyth 1920:§341)

The local endings are all inflectional. The adverbs in (64) are thus frozen forms of *inflected* words.

Nominal inflection. Case-endings for vowel-final noun stems are given in (65), those for consonant-final noun stems in (66). Representative examples for masculine, feminine and neuter nouns are given for each.

(65) Case-endings of vowel-final noun stems

	Masc and Fem Neuter	Masc	Fem	Neut
		'road'	*'mina'*	*'gift'*

Nom	-s, ø	-n	hodós	mnáa	dóoron
Gen	-s, -io	-s, -io	hodóu	mnáas	dóorou
Dat	-i	-i	hodóoi	mnáai	doóroοi
Acc	-n	-n	hodón	mnáan	dóoron
Voc	ø	-n	hodé[32]	mnáa	dóoron
Nom, Acc, Voc	ø	ø	hodoó	mnáa	doóroo
Gen, Dat	-in	-in	hodóin	mnáin	doóroin
Nom	-i	-a	hodoí	mnái	dóora
Gen	-oon	-oon	hodóon	mnóon	doóroon
Dat	-is(i)	-is(i)	hodóis	mnáis	doórois
Acc	-ns	-a	hodoús[33]	mnáas	dóora
Voc	-i	-a	hodoí	mnái	dóora

(66) Case-endings of consonant-final noun stems

	Masc and Fem	Neuter	Masc	Fem	Neut
			'guard'	*'beast'*	*'body'*
Nom sg	-s, ø	ø	phuúlaks	theér	sóoma
Gen	-os	-os	phuúlakos	theerós	soómatos
Dat	-i	-i	phuúlaki	theerí	soómati
Acc	-n, -a	ø	phuúlakn	théera	sóoma
Voc	ø[34]	ø	phuúlaks	theér	sóoma
N, A, V dual	-e	-e	phuúlake	théere	soómate
Gen, Dat	-oin	-oin	phuulákoin	theeróin	soomátoin
Nom pl	-es	-a	phuúlakes	théeres	soómata
Gen	-oon	-oon	phuulákoon	theeróon	soomátoon
Dat	-si, -ssi, -essi	-si, -ssi, -essi	phuúlaksi	theersí	soómasi
Acc	-ns, -as	-a	phuúlakas	théeras	soómata
Voc	-es	-a	phuúlakes	théeres	soómata

Non-moraic affixes are used to mark nominative (*-s, -n*), genitive (*-s*), accusative (*-n*) and vocative (*-n*) singulars as well as accusative plurals (*-ns*). There is nothing to suggest that any

[32] 'The stem vowel o varies with e, which appears in the vocative sing' (Smyth 1920:§229b).
[33] for /hodo-ns/. n deletes before s with compensatory lengthening of the preceding vowel.
[34] or like the nominative form.

of these is underlyingly moraic (e.g., has a vowel underlyingly).

Adjectival inflection. Case-endings for adjectives are essentially those of nouns. Most vowel-final adjectives take case endings like those in (66): *agathós* 'good-masc' is declined like *hodós* 'road'; *agatheé* 'good-fem' like *mnáa* 'mina', *agathón* 'good-neut' like *dóoron* 'gift'. Thus a number of adjectival inflectional affixes consist of less than a mora but these are identical to those given above.

Verbal inflection. AG verbal inflection is extensive and I will not review it all here. For the present, I merely want to establish that it makes use of a number of non-moraic afffixes. As with nominal and adjectival affixes, AG non-moraic affixes are always dentals (s, n, t^h).

AG verbal inflection consists of tense-aspect prefixes and a large number of suffixes. The prefixes are both moraic and I will not discuss them here. Suffixes that consist of a single consonant include two tense markers (future *-s-* and aorist *-s-*), an aspect marker (perfect *-k-*) and a voice marker (passive *-t^h-*). Each of these is non-moraic, as the following discussion will show.

The future is formed by inserting *-s-* after the root, as a comparison of the present active indicative and future active indicative shows (67):

(67) Future *-s-*

		Present	Future
Singular	1	luú-oo 'I loosen'	luú-*s*-oo 'I shall loosen'
	2	luú-eis	luú-*s*-eis
	3	luú-ei	luú-*s*-ei
Dual	2	luú-eton	luú-*s*-eton
	3	luú-eton	luú-*s*-eton
Plural	1	luú-omen	luú-*s*-omen
	2	luú-ete	luú-*s*-ete
	3	luú-ousi	luú-*s*-ousi

Comparison of the aorist active indicative and perfect active indicative reveals the non-moraic affixes *-s-* and *-k-*:

(68) Aorist *-s-* and Perfect *-k-*

		Aorist	Perfect
Singular	1	é-luu-*s*-a 'I loosened'	lé-lu-*k*-a 'I have loosened'
	2	é-luu-*s*-as	lé-lu-*k*-as
	3	é-luu-*s*-e	lé-lu-*k*-e
Dual	2	e-luú-*s*-aton	le-lú-*k*-aton
	3	e-luu-*s*-áteen	le-lú-*k*-aton
Plural	1	e-luú-*s*-amen	le-lú-*k*-amen
	2	e-luú-*s*-ate	le-lú-*k*-ate
	3	é-luu-*s*-an	le-lú-*k*-aasi

The marker *-t^h-* is added to the root to form passives. This is evident in comparing, e.g., the present active subjunctive with the aorist passive subjunctive:

(69) Passive *-t^h-*

		Active	Passive
Singular	1	luú-oo	luu -t^h-óo
	2	luú-ees	luu -t^h-ées
	3	luú-ee	luu -t^h-ée
Dual	2	luú-eeton	luu -t^h-éeton
	3	luu-eeton	luu -t^h-éeton
Plural	1	luú-oomen	luu -t^h-óomen
	2	luú-eete	luu -t^h-éete
	3	luú-oosi	luu -t^h-óosi

The other source of non-moraic verbal inflection is the large number of fused person-number suffixes. Consider the following paradigms for active and meddle/passive voices.

(70) Subject agreement: Active

		Present, Perfect, Future Future Perfect, Subjunctive	Imperfect, Pluperfect, Aorist, Optative	Imperative
Singular	1	ø, -mi	-n	
	2	-s, -t^ha	-s	ø, -t^hi, -s
	3	-si	ø	-too
Dual	2	-ton	-ton	-ton
	3	-ton	-ton	-toon
Plural	1	-men	-men	
	2	-te	-te	-te
	3	-nsi	-n, -san	-ntoon

Non-moraic suffixes include *-n* in the 1st singular and 3rd plural, and *-s* across the board in the 2nd singular. Middle and passive subject agreement is given in (71).

(71) Subject agreement: Middle/Passive

		Present, Perfect, Future Future Perfect, Subjunctive	Imperfect, Pluperfect, Aorist, Optative	Imperative
Sing.	1	-mai	-meen	
	2	-sai	-so	-so
	3	-tai	-to	-st^{h}oo
Dual	2	-st^{h}on	-st^{h}on	-st^{h}on
	3	-st^{h}on	-st^{h}een	-st^{h}oon
Plural	1	-met^{h}a	-met^{h}a	
	2	-st^{h}e	-st^{h}e	-st^{h}e
	3	--ntai	-nto	-st^{h}oon

There are no non-moraic affixes here, though remnants of non-moraic person marking can be seen in the repetition of *m* (1st person), *s* (2nd person) and *t* (3rd person).

AG inflectional affixes are not subject to the minimal affix requirement of the language. Whereas non-moraic derivational affixes in these languages are non-existent, inflectional affixes that consist of a single dental consonant are commonplace. The fact that inflectional affixes, function words and the roots of function words in AG are not subject to the minimal prosodic constraints placed on derivational affixes, content words and the roots of content words is easily modeled with the 2 Lexicon Hypothesis: minimal prosodic requirements are met by elements of the Lexicon but need not be met by those of the Phrasicon.

2.3 Latin

Latin provides another case in which content words and derivational affixes are subject to

minimal prosodic requirements to which function words and inflectional affixes are not subject. As with English and AG, Latin content words must consist of at least a heavy syllable, though function words do not conform to this requirement; derivational affixes consist minimally of a light syllable while inflectional affixes do not conform to this requirement.

The Minimal Word in Latin. Like English and AG, Latin allows no content words that consist of a single light syllable. Again, this holds not only of lexical *words* but of lexical *roots*; since words are formed from roots, the minimal word requirement in Latin clearly follows from the miminal root requirement.

As with English and AG, a bimoraic foot can be shown independently to be the foot used in stress (Allen 1973; see also Hayes 1981). Thus, the minimum word = the minimum foot in Latin. It appears that word-final obstruents were extrametrical for the purpose of satisfying thius bimoraic minimum; consequently, content words consist minimally of either CVV or of CVR where R is a sonorant.

Latin had no monomoraic content words. Examples of bimoraic nominals are given below:

(72) Latin minimal nouns (nominative singular)

CVV					
vii	'force (abl)'	ree	'thing (abl)'	dii	'gods'
CVC(C)					
nik(s)	'snow'	trab(s)	'beam'	ar(s)	'art'
CVR					
vir	'man'	cor	'heart'	rem	'thing(acc sg)'

Obstruent-final CVC nouns are extremely rare and will be discussed below. In general,

"nouns and adjectives of one syllable are long" (Allen & Greenough 1903§604):

(73) Monosyllabic content words are heavy

oos	'mouth'	boos	'cow'	vaas	'vessel'
viis	'strength'	sool	'sun'	veer	'spring'

The lack of CVC nouns and adjectives in Latin can be attributed to final consonant extrametricality. In AG, however, any final consonant is extrametrical whereas in Latin, only final obstruents are extrametrical. There is little additional evidence for the extrametricality of final obstruents in Latin: final *syllables*, regardless of weight, are extrametrical for Latin stress rules (Hayes 1981) and thus the issue of final *consonant* extrametricality does not arise here. Still, final obstruent extrametricality allows us to explain the absence of Latin obstruent-final CVC nouns and adjectives rather than stipulate it.

Latin adjectives are also minimally bimoraic. Indeed, most adjectives exceed this minimum due to stem-final theme vowels, derivational affixes or moraic inflection. The adjectives in (74) are among the only monosyllabic adjectives in the language.

(74) Latin minimal adjectives (neuter singular unless otherwise noted)

CVV(C) adjectival roots					
pluu(s)	'more'	dii(s)	'rich'	paar	'equal'

Note that paar is only near-minimal since the final consonant is moraic. Minimal CVC and CVV adjectival *roots* occur quite frequently in Latin, as will be seen below.

Minimal word sized verbs are given in (75):

(75) Latin minimal verbs

CVV					
stoo	'I stand'	staa	'stand!'	noo	'I swim'
CVR					
for	'I say'				

(*for* may well be better classified as a function word; truly minimal verbs are difficult to find in Latin since most inflected forms have additional moras.) Again, the lack of CV and CVC

(content) provides evidence for a bimoraic minimal word requirement if final consonants are taken to be extrametrical.

There is one set of obstruent-final CVC nouns that appear problematic for the bimoraic minimal word proposal. If final obstruents are indeed extrametrical, as I have argued, such words seem to violate the minimal word requirement of Latin.

Obstruent-final CVC nouns in Latin. Allen & Greenough (1903§604) list a small number of words as exceptions to their observation that monosyllabic nouns and adjectives are heavy. These are given in (76) and (77)--I have added *as* 'unit' as well..

(76) Apparent CVC nouns in Latin

a.			b.		
	cor	'heart'		as	'unit'
	fel	'gall bladder'		lac	'milk'
	mel	'honey'		os	'bone'
	vir	'man'		vas	'bail'
	ter	'thrice'			
	rem	'thing (acc)'			
	vim	'strength (acc)'			

(77) tot 'as many'
quot 'as'

The words in (77) are closed-class correlative adverbs and will be treated here as function words. The four words in (76a) all end in sonorants. These words are bimoraic and do not violate the minimal word requirement. The words in column (76b) all end in obstruents. Comparison of their genitive forms reveals, however, that each of these words ends in a consonant cluster underlyingly:

(78) CVC ~ CVCC stem alternations

as	'unit'	lac	'milk'	os	'bone'
ass-is	'unit' (gen)	lact-is	'milk' (gen)	oss-is	'bone' (gen)

The underlying nominative forms are thus /lact/, /oss/ and /vad-s/; the rules of final cluster simplification that create the surface CVC forms are well attested elsewhere in Latin. If we

assume that these rules apply late enough in the derivation, i.e., after minimal word reuirements need to be met, none of these words poses a serious problem for a bimoraic minimal word in Latin. Like the AG neuter participles *stán, thén*, etc. discussed earlier, these words would be bimoraic for enough of the derivation to satisfy minimal word requirements. (Note that the nominative forms for *cor, fel* and *mel* also end in clusters underlyingly: cf. genitives *cord-is, fell-is, mell-is.*) Derivations for *lac, os* and *as* are given below:

(79) Final cluster simplification and minimal word requirements

UR	as	lact	oss
Min Word Requirement	√	√	√
Degemination	as		os
Cluster Simplification		lac	

This analysis hinges crucially on the *late* application of Degemination and Cluster Simplification, since they are claimed here to apply *after* the Minimal Word requirement must be satisfied. If the Minimal Word requirement is a constraint on the output of the Lexicon, the final consonants in these words should survive at least until the first stage of lexical insertion (LexIns). In fact, there is some evidence that such final consonant clusters survive even later.

Evidence that word-final geminates were simplified late in the derivation of a sentence comes from the behavior of word-final geminates in Roman comedy. Buck notes that in Plautus, "*es* [you are] regularly, and sometimes *ter* [thrice] and the last syllable of *miiles* [soldier], have the value of long syllables, which means the survival of *ess, miiless, terr,* before vowels" (1933§212.6a). That is, at least some word-final geminates survive (in regular speech) up to the sentence level. even though they are simplified in spelling. Thus *oss* 'bone' and *ass* 'unit', at least, may have been bimoraic quite late in the phonology of Latin.

Similar evidence exists for a bimoraic form of *lac.* Latin 'milk' seems to have had three nominative forms: *lac, lact,* and *lacte*. *lac* is 'standard' Latin; but *lact* and *lacte* occur as well. Plautus, for instance, has nominative *lacte* as early as the beginning of the second century BCE. As Ernout and Meillet point out,

> La variation entre *lac* et *lacte* a dû dépendre à l'origine de l'initiale du mot suivant...: *lac* devant consonne, *lacte* devant voyelle; *lact* semble être une construction de grammairien [and thus is to be discounted for our purposes--C.G.]. A l'epoque classique, la première forme parait plus littéraire; c'est la seconde qui est représentée dans les langues romanes. (1959:335)

The literary form *lac* fails to meet the bimoraic minimum but this is of little concern; the actual form in the spoken language seems to have been lacte and thus to have met the bimoraic minimum.

This leaves anomalous *vas*. At present, *vas* must be treated as exceptional, since I am aware of no evidence that such underlying clusters were retained. Indeed, similar forms such as /ped-s/ 'foot' lengthen their vowel after T/D deletion to yield forms like *pees* (see below for a fuller discussion). What could account for the exceptionality of *vas* 'bail'? I believe the answer lies in the existence of another word *vaas* 'vessel'. Lengthening of *vas* 'bail' to *vaas* would render the words for 'bail' and 'vessel' homophonous. Avoidance of this homophony seems to be what licenses monomoraic *vas* in the Lexicon. (A similar problem with *os* 'bone' and *oos* 'mouth' would be avoided by the retention of the underlying geminate in 'bone': *oss* 'bone' contrasts fully with *oos* 'mouth'.)

None of the words in (76) or (77), then, seriously violate the bimoraic minimal word requirement in Latin. This is equally ture of lexical *roots*.

Lexical Roots. The bimoraic minimum in Latin extends to roots as well as to derived words. As with AG, final consonants in roots are not extrametrical (only *word*-final consonants are extrametrical), so a bimoraic minimum translates into CVC or CVV for monosyllables. CVC nominal and verbal roots are especially common. Adjectival roots are comparatively rare since most adjectives are formed from nominal roots--I have therefore included near-minimal CVVC roots in (81).

(80) Latin nominal roots

CVC					
vad-	'bail'	fac-	'torch'	urb-	'city'
CVV					
ree-	'thing'	suu-	'sow'	vii-	'force'

(81) Latin adjectival roots

CVC					
bon-	'good'	mal-	'bad'	nig-	'black'
CVV					
paar-	'equal'	pluus-	'more	pee-	'bad'

(82) Latin verbal roots

CVC					
voc-	'call'	cap-	'take'	lav-	'wash'
CVV					
flaa-	'blow'	naa-	'swim'	cree-	'believe'

(Again, I have not included roots of derived adverbs since these are derived from nominal or verbal roots.)

I am aware of only two lexical roots that are monomoraic: *lu-* 'loosen' and *flu-* 'flow'. The first, according to Ernout and Meillet, "bien qu'ancien et classique, est d'un emploi rare et a été remplacé par son composé *solv-oo*, qui indique le procès parvenu às son terme" (1959:370). Note that *solv-* is straightforwardly bimoraic. Similary, "cool-oo s'est substitutué à flu-oo dans les langues romanes, ou il est partout attesté" (p. 134). In both cases, the common form of the word is bimoraic; only the more literary form is monomoaic.

Bimoraic Lengthening. Like AG, Latin has a lengthening rule that preserves bimoraicity in derived words. Recalling the discussion above of AG póus 'foot', consider the Latin singular paradigms below:

(83)

	pod- 'foot'	*mas-* *'male'*
nom	pee-s	maas
acc	ped-em	mar-em
gen	ped-is	mar-is
dat	ped-ii	mar-i
abl	ped-e	mar-e

The oblique cases of both words reveal roots with short vowels: pod- and mas-. Consider *ped-* first: Latin dentals delete before *s*, so that the underlying nominative singular form *ped-s* yields an intermediate form *pe-(s)*; this form is monomoraic, however, and is therefore lengthened to *pees*. Similarly with mas-: the nominative singular is unmarked here, causing root-final [s] to be extrametrical. Consequently, the word is monomoraic and is therefore lengthened to maas. (The oblique cases with root-final *r* come from a rule that rhotacizes intervocalic *s*.) Derivations are given below.

(84) Derivation of nominative singular *pees* 'foot', maas 'male'

Underlying Form	ped-s	mas
T/D Deletion	∅	-----
	pe -s	-----
Lengthening	pee-s	maas

Similar cases can be seen in other monosyllabic forms: *laar* 'household god', *paar* 'equal', *saal* 'salt' all derive from roots with short vowels (genitive *laris, paris, salis*). In general, of course, word-final sonorants are not extrametrical; consequently, the extrametricality of these sonorants must be taken as an idiosyncratic property of these roots: *la(r), pa(r), sa(l)*. Once this step is taken, the the otherwise anomalous lengthening of the vowel follows straightforwardly.

Letter names in Latin are less well-established than for AG. What seems beyond dispute is that the letter names for *a, e, i, o,* and *u* (which represent either short or long vowels in the orthography) were all bimoraic: [aa], [ee], [ii], [oo], [uu]. Vowel-final leter names for letters like *b* and *d* seem also to have been bimoraic, [bee] [dee] (but see 19xx). There is evidence

that some letters had no names but were represented merely by a sound: eg, f was [ffff...]. These were clearly not names as such, but mimetic utterances; as such, they do not bear on the minimal word requirement.

Latin Function Words ≤ μμ. Latin has a large number of monomoraic function words; as with AG, they occur in almost every non-lexical grammatical category. The words below are monomoraic in one of two ways. CV words such as *qua* 'what' have a single moraic vowel; obstruent-final CVC words like *it* 'goes' violate bimoraicity because of their extrametrical final consonants.

(85) Monomoraic function words

Prepositions:	ab	'from'
	ad	'to'
	cis	'this side'
	ob	'on acount of'
	sub	'under'
Demonstratives::	is	'this' (masc nom sg)
	id	'this' (neut nom, acc sg)
Interrogative pronouns:	quis	'who' (masc, fem nom sg)
	quid	'what' (neut nom, acc sg)
Indefinite pronouns:	qua	'what' (neut nom, acc pl)
Conjunctions	ac	'and'
	at	'but'
	et	'and'
	nec	'and not'
	-que	'and'
	quod	'because'
	sed	'but'
	ut	'so that'
	-ve	'or'

Correlatives	tot...quot	'so many...as'
Negatives	-ne	'not'
Closed-Class Verbs	es	'you are'
	sit	'may s/he be'
	es	'be!'
	fit	'becomes'
	it	'goes'
	dat	'gives'
	det	's/he should give'
	quit	's/he cannot'

(Note that, except for *qua* 'what', all CV function words in Latin are enclitic: e.g., -ne 'not', -*ve* 'or'.)

One word that does not appear in (85) requires some discussion: *fac* 'make' (2nd sg imper) is also monomoraic, though it is not clear whether it should be treated as a function word. Evidence that is is a function word in Latin includes (1) its simple semantics, (2) its suppletive paradigm (passive is supplied by active forms of *fio* 'be made, become' and (3) the fact that it appears to violate the bimoraic minimum (though Plautus has both *fac* and *face*, suggesting that it may have been bimoraic in spoken Latin). I will not pursue this further here.

Derived monomoraic function words. As with AG, a number of Latin function words take inflectional suffixes. This occurs both with nominal (qui-s 'who') and verbal (i-t 's/he goes') function words.

Non-lexical roots. The roots of Latin function words also often fail to obey bimoraicity. As with AG, stripping off inflectional material from a number of function words reveals monomoraic roots. The copula, for instance, is built on three roots, one that obeys the bimoraic minimum (*es-*) and two that don't (*su-* and *fu-*).

(86) Monomoraic function word roots

su- 'be'	su-m	'I am'
	su-mus	'we are'
	su-nt	'they are'
fu- 'was'	fu-ii	'I was'
	fu-erit	'he should have been'
da- 'give'	da-re	'to give'
	da-nt	'they give'
i- 'this'	i-s	'this man'
	i-d	'this thing'
qui- 'which'	qui-s	'who'
	qui-d	'what'

(see Buck 1933§30ff. for discussion of pronominal roots).

The bimoraic minimal word requirement in Latin holds for content words and their roots but not for function words and their roots. The 2 Lexicon Hypothesis models this situation straightforwardly by having minimal prosodic requirements met only in the Lexicon, not in the Phrasicon. The following sections will show that Latin affixes in the Lexicon also are subject to a minimal requirement of one mora that affixes in the Phrasicon are not subject to.

The Minimal Affix in Latin. The minimal affix in Latin (as in English and AG) is monomoraic:

(87) Latin: Min Aff = μ

Allen & Greenough (1903) provide a fairly exhaustive list of Latin derivational affixes. None of these (ultimately) consists of less than a syllable. Consider the derivational prefixes:

(88) Latin prefixes

amb-	'around'	amb-iire	'go around'
dis-	'apart'	dis-ceedere	'depart'
dii-	'apart'	dii-videre	'divide'
por-	'foward'	por-tendere	'predict'
red-	'again'	red-iire	'return'

re-	'again'	re-ficere	'redo'
seed-	'apart'	seed-itioo	'secession'
see-	'apart'	see-cernoo	'I separate'

(I have listed as *prefixes* only bound lexical formatives.)

Latin had a large number of derivational suffixes that formed nouns, verbs, adjectives and adverbs. None of these is less than monomoraic. Category names are all those of Allen & Greenough. I have included only what they term "significant endings", i.e., those which "were used in Latin with more or less consciousness of their meaning; relics of PIE suffixes that are not synchronically recoverable in Latin have not been included below (though they too are monomoraic--see Allen & Greenoough §234)

Noun-forming suffixes. Latin nominalizing suffixes are given below. All of these consist of at least one mora.

(89) Agentives

-tor	can-tor 'singer'	*-trik*	can-trik-s 'singer' (fem)
-et	teg-et-is 'cover' (gen)	*-oon*	ger-oon-is 'carrier' (gen)

(90) Actions and abstract nouns (deverbal)

-or	tim-or 'fear'	*-ee*	seed-ee-s 'seat'
-us	gen-us 'birth'	*-ioon*	ref-ioon-is 'direction' (gen)
-tioon	vocat-ioon-is 'calling' (gen)	*-tuur*	scriip-tuur-a 'writing'
-su	seen-su-s 'feeling'		

(91) Acts, means and results

-men	ag-men 'line of march'	*-ment*	regi-ment-um 'rule'
-mooni	testi-mooni-um 'testimony'	*-mooni*	queri-mooni-a 'complaint'

(92) Means or instrument

-bul	paa-bul-um 'fodder'	*-cul*	vehi-cul-um 'vehicle'
-br	candeelaa-br-um 'candlestick'	*-cr*	sepul-cr-um 'tomb'
-tr(o)	araa-tr-um 'plough'	*-ber*	tuu-ber 'tube'

(93) Abstracts (de-adjectival)

-i	audaac-i-a 'boldness'	*-ieet*	pauper-ieet-is 'poverty' (gen)
-ti	triisti-ti-a 'sadness'	*-tieet*	seegni-tieet-is 'laziness' (gen)
-taat	boni-taat-is 'goodness' (gen)	*-tuut*	senec-tuut-is 'age' (gen)
-tuudoon	maagni-tuudoo 'greatness'	*-li*	auxi-li-um 'help'
-goon	lumbaa-goon-is 'lumbago' (gen)	*-ni*	contici-ni-um 'hush of night'
-cini	laatroo-cini-um 'robbery'		

The forms in (92) require some discussion. The *-br* and *-cr* affixes in *candeelaabrum* and *sepulcrum* are dissimilated forms of *-bul* and *-cul*, respectively (Buck 1933:330). Derivations appear in (94), Dissimilation is given in (95), U-syncope in (96).

(94) Derivations

UR	candeelaa-bul-um	sepul-cul-um
Dissimilation	candeelaa-bur-um	sepul-cur-um
U-syncope	candeelaa-br-um	sepul-cr-um
SR	candeelaabrum	sepulcrum

(95) Dissimilation

l -> r / l ... __

(96) Syncope

V -> ∅ / C__rV (where both vowels are stressless)

Syncope and Dissimilation are needed independently. (97) shows the application of Dissimilation on a derived adjective *populaaris* 'popular'.

(97) Dissimilation

UR	popul-aal-is
Dissimilation	popul-aar-is
SR	populaaris

The application of Syncope on an inflected adjective (*rub-* 'red') and noun (*maater-* 'mother') is illustrated below.

(98) U-syncope

	masc.	*fem.*	*nominative*	*genitive*
UR	rub-er	rub-er-a	maater	maater-is
Syncope	------	∅	------	∅

SR	ruber	rubra	maater	maatris

This leaves- tr(o) in *araatrum* 'plough' as an isolated exception.

Adjective-forming suffixes. Below is a list of Latin adjectival suffixes.

(99) Diminutives

-ul	riiv-ul-us 'streamlet'	*-cul*	auri-cul-a 'little ear'
-ol	gladi-ol-us 'small sword'	*-ell*	lib-ell-us 'little book'
-ill	coodic-ill-ii 'writing tablets'	*-cioon*	homun-cioon-is 'dwarf' (gen)

(100) Other adjectival endings

-oos	form-oos-us 'beautiful'	*-lent*	viino-lent-us 'given to drink'
-e	aur-e-us 'golden'	*-i*	patr-i-us 'paternal'
-tic	domes-tic-us 'dometic'		
-aace	ros-aace-us 'of roses'	*-iici*	later-iici-us 'of brick'
-aane	subterr-aane-us 'subterranean'	*-ne*	saliig-ne-us 'of willow'
--aan	mont-aan-us 'of the mountains'	*-een*	terr-een-us 'earthly'
--iin	div-iin-us 'divine'	*-eensi*	for-eensi-s 'of the Forum'
-aali	naatuur-aali-s 'natural'	*-eeli*	patru-eeli-s 'cousin'
-iili	host-iili-s 'hostile'	*-uuli*	cur-uuli-s 'curule'
-aat	iinfim-aat-is 'lowest ranking' (gen)		
-ic	fulloon-ic-us 'of a fuller'	-iic	am-iic-us 'friend'
-oon	naas-oon-is 'with a large nose' (gen)		
-aac	mer-aac-us 'pure'	*-e*	feemin-e-us 'feminine'
-eei	pleeb-eei-us 'plebeian'	*-ici*	patr-ici-us 'patrician'
-ter	paluus-ter 'of the marshes'	*-ester*	silv-ester 'woody'
-tri	seemees-tri-s 'lasting six months'	*-estri*	silv-estri-s 'woody'
-tim	fiini-tim-us 'neighboring'		
-aari	ordin-aari-us 'ordinary'	*-oori*	merit-oori-us 'profitable'

-aac	puugn-aac-s 'pugnacious'	*-id*	cup-id-us 'eager'
-ul	bib-ul-us 'thirsty'		
-u	proter-w-us 'violent'	*-iiw*	recid-iiw-us 'restored'
-ili	frag-ili-s 'fragile'	*-bili*	noo-bili-s 'well-known'
-i	exim-i-us 'choice'		

A non-moraic version of *-ic*, *-c*, occurs in *buubul-c-us* 'ox tender'. I assume it is a derived form of *-ic*, reduced to *-c* by syncope. Note also that the the *-u* in *proter-w-us* 'violent' is consonantal [w] not syllabic [u]; [u] regularly becomes a glide pre-vocalically.

Enlargements. As can be seen above, Latin has a number of suffixes that are built on smaller suffixes: e.g., *-aali*, *-eeli*, *-iili* and *-uuli* are built on a suffix *-li*. The larger suffixes are sometimes referred to as "enlargements". Generally, the "enlarged" affixes are the productive ones:

> One of the most characteristic features of Latin suffixes is their growth by misdivision: for instance, the elementary suffix *-n(us)* gives rise to a group of secondary suffixes *aan(us)*, *-iin(us)*, *-ern(us)*, *-tin(us)* [my parentheses--C.G.]. It is in many casses impossible to determine the historical facts, since in the classical period *-aan(us)* was clearly felt as a living suffix, whereas *-n(us)* was no longer employed in new formatives.
>
> (Oxford Latin Dictionary, p. xviii)

At least some of the 'elementary' suffixes underlying these enlargements are non-moraic. They are not synchronically recoverable as affixes in Latin, but deserve some attention nonetheless. What is noteworthy about them is that they are derived from *inflectional* affixes in the parent language. That is, a number of non-moaraic inflectional affixes are reanalyzed as derivational affixes in the history of the language: when they are, they are 'enlarged', presumably to meet the minimal affix requirement. Examples follow:

(101) Adjectival suffixes derived from inflectional affixes

'Elementary' affixes	*'Enlargements'*

-nd	secu-nd-us 'second'	*-bund*	viitaa-bund-us 'avoiding'
	rotu-nd-us 'round'		treme-bund-us 'trembling'
			mori-bund-us 'dying'
	-cund		faa-cund-us 'eloquent'
			fee-cund-us 'fruitful'
			iiraa-cund-us 'irascible'
-n	veer-n-us 'vernal'	*-ern*	hodi-ern-us 'of today'
		-urn	di-urn-us 'daily'
		-tern	hes-tern-us 'of yesterday'
	-turn		diuu-turn-us 'lasting'
-mn	alu-mn-us 'nursling'	*-min*	fee-min-a 'nourisher'

Thus, the non-moraic *-nd-* in *secundus* and *rotundus* is from an old inflectioanl affix in the parent language. That affix is still used in Latin as an inflectional affix (future passive participle) in obligation-constructions (*audie-nd-us est* 'he must be heard', *deelige-nd-us erat* 'he should have been chosen'); but the derivational enlargments of *-nd-* (i.e.,, *-bund* and *-cund*) are moraic. Similarly, the *-n-* in veernus 'vernal' "form[s] perfect participles in other languages and in Latin mak[es] adjectives of like participial meaning, which often become nouns" (Allen & Greenough §234.II.4); derivational enlargements of the -n- suffix (*-ern, -urn, -tern, -turn*) again obey the minimal affix requirement. The *-m(e)n* in *alumnus* and *feemina* is from a present participle suffix in the parent language; neither is recoverable synchronically in Latin, however, and they are given here merely for completeness.

One final set of suffixes require some discussion, the Latin froms corresponding to English *-ed* adjectives (*palefaced, bearded*):

(102) -t

hones-t-us 'honorable'
barbaa-t-us 'bearded'
turrii-t-us 'turreted'
cornuu-t-us 'horned'

This non-moraic suffix is in fact *inflectional*, however, used in forming the supine stem of

Latin verbs:

(103) -t (supine stems)

amaa-t-us	'loved' (masc sg)
deelee-t-um	'deleted' (neut sg)
audii-t-um	'heard' (neut sg)

Thus the adjectives in (102) essentially "imply reference to an imaginary verb-stem" (Allen & Greenough §246n). *barbaatus* 'bearded' is (so to speak) the past participle of barba 'beard' .

Verb-forming suffixes. Latin verb-stems are formed in one of three ways: (1) directly from nominal stems (zero derivation), (2) from nominal stems by suffixation and (3) from other verbal stems by affixation. Examples of zero derived verbs are given below. Zero derivation will not be of major concern here, since it involves no affixes and thus no evidence for or against a minimal affix requirement.

(103) Zero-derived denominal verbs

fug-aa-re 'to put to flight'	*from*	$[[fug]_{root}\ aa]_{stem}$	'flight'
stell-aa-re 'to be thick with stars'	*from*	$[[stell]_{root}\ aa]_{stem}$	'star'

Zero derivation involves a nominal stem (*fugaa-, stellaa-*), itself composed of a root (*fug-, stell-*) and a thematic-vowel (*-aa*). The thematic vowel often takes on a life of its own, however, in forming verb-stems:

(104) *-aa* stem verbs from *-o* stem nominals

stimul-aa-re 'to incite'	*from*	stimulo- 'a goad'
nov-aa-re 'to renew'	*from*	novo- 'new'

-aa stem verbs may also be formed from nominal stems that end in consonants, in *-i* or in *-u*:

(105) other *-aa* stem verbs

vigil-aa-re 'to watch'	*from*	vigil- 'awake'
auspic-aa-rii 'to take auspices'	*from*	auspic- 'augur'
lev-aa-re 'to lighten'	*from*	levi- 'light'
aestu-aa-re 'to surge'	*from*	aestu- 'tide'

The data in (104) - (105) shows that *-aa* has been made into a a verbalizing suffix in Latin, freed from its origins in *-aa* stem nouns; it now derives verbs stems (*fugaa-, stimulaa-, vigilaa-, auspicaa-*) from noun stems of many types (*fugaa-, stimulo-, vigil-, auspic-*). Other such verbalizing suffixes include *-ee, -e* and *-ii*:

(106) *-ee* stem verbs from nominals
claar-ee-re 'to shine' *from* claar- 'bright'
claud-ee-re 'to be lame' *from* claud- 'lame'

(107) *-e* stem verbs from nominals
metu-e-re 'to fear' *from* metu- 'fear'
acu-e-re 'to sharpen' *from* acu- 'needle'

(108) *-ii* stem verbs from nominals
iinsaan-ii-re 'to rave' *from* iinsaano- 'mad'
cuustood-ii-re 'to guard' *from* cuustood- 'guardian'

These verbalizing affixes lend more support to the claim that derivational affixes in Latin are minimally monomoraic.

This completes the discussion of derivational affixes in Latin. The 100 or so derivational affixes include none that consist (underlyingly) of less than a mora. This provides strong support for a minimal affix requirement Latin.

LatinInflectional Affixes < μ. But a number of central Latin inflectional affixes are non-syllabic.

Adverb-forming suffixes. Most adverbs in Latin are inflected nominals: the neuter accusative of adjectives and pronouns may serve as an adverb (*facil-e* 'easily'), as does the ablative singular neuter and feminine of adjectives, pronouns and nouns (*fals-oo* 'falsely', *reect-aa* 'straightway'). Adjectives may also be inflected with *-ter* and both adjectives and nouns take adverbial *-ee* (itself a generalized ablative suffix).

(109) Adverb-forming suffixes

		Manner	
-ee	caar-ee 'dearly'	*-ter*	aacri-ter 'eagerly'
	amiic-ee 'like a friend'		forti-ter 'bravely'

Nominal inflection. Case-endings for 1st and 2nd declension noun stems are given below.

(110) 1st and 2nd Declension Nominal Inflection

	1st Declension	*'star'*		2nd Declension *'slave'*	
Nom sg	*-ø*	stella	*-s*	servo-s	(servos)
Gen	*-i*	stella-e	*-i*	servo-i	(servii)
Dat	*-i*	stella-e	*-V*	servo-o	(servoo)
Acc	*-m*	stella-m	*-m*	servo-m	(servom)
Abl	*-V*	stella-a	*-V*	servo-o	(servo)
Nom pl	*-i*	stella-e	*-i*	servo-i	(servii)
Gen	*-Vrum*	stella-arum	*-Vrum*	servo-orum	(servoorum)
Dat	*-iis*	stell-iis	*-iis*	servo-iis	(serviis)
Acc	*-Vs*	stella-as	*-Vs*	servo-os	(servoos)
Abl	*-iis*	stell-iis	*-iis*	servo-iis	(serviis)

Note that non-moraic sufixes are used to mark nominative (*-s*) and accusative (*-m*) singular. I have diverged somewhat from traditional analyses of these declensions, which tend to fuse stem vowel and person marking; the divergence is most evident in the first declension where I assume a short *-a* stem vowel rather than long *-aa*. Nominative singular *stella* is thereby the bare stem (rather than a shortened stem, as on the traditional analysis). The various forms *stellae* (genitive and dative singular, nominative plural) are derived from *stella-i* by Diphthong Lowering:

(111) Dipthong Lowering: ai -> ae

The stem vowel in the dative and ablative plural is lost by a rule of Vowel Deletion, which deletes the first of three tautosyllabic vowels:

(112) Vowel Deletion: V -> ø / __VV]σ

[-hi]

Derivations for *stellae* and *stellis* are given below:

(113) Derivations

UR	stella-i	stella-iis
Dipthong Lowering	stellae	--------
Vowel Deletion	------	stell-iis

The ablative singular and genitive and accusative plural suffixes of the first declension contain empty Vowel positions that serve to lengthen the preceding vowel. This is accomplished by a rule that spreads vowel quality rightwards onto an empty V position.

(114) Spreading:

(115) Derivations

UR	stella-Vrum	stella-Vs
Spreading	stellaarum	stellaas

Such empty V-positions are also found in some of the 2nd Declension suffixes (note that the empty V-position allows us to state the first and second declension differ only in their affixes for the nominative and dative singulars). Note that the diphthong *-oi* surfaces as a long *-ii* word-finally (cf. Buck 1933§90).

(116) Unrounding: oi -> ii]ω

Derivations for *servii* and *servoo* folow:

(117) Derivations

UR	servo-i	servo-V
Spreading	--------	servoo
Unrounding	servii	

Some 3rd Declension inflectional suffixes are also non-moraic:

(118) Case-endings, 3rd Declension

	C-stems		i-stems	
Nom sg	-s	dux (duc-s)	-s	turri-s
Gen	*-is*	duc-is	-s	turri-s
Dat	*-ii*	duc-ii	*-V*	turri-i
Acc	-m	duc-em	-m	turri-m
Abl	*-V*	duc-e	*-V*	turri-i
Nom pl	*-ees*	duc-ees	*-ees*	turr-ees
Gen	*-um*	duc-um	*-um*	turri-um
Dat	*-ibus*	duc-ibus	*-bus*	turri-bus
Acc	*-ees*	duc-ees	*-Vs*	turri-is
Abl	*-ibus*	duc-ibus	*-bus*	turri-bus

Non-moraic affixes mark nominative (*-s*), genitive (*-s*), and accusative (*-m*) singulars. I take the [e] in *ducem* to be epenthetic (/ducm/ -> [ducem]), [e] being the unmarked vowel quality in Latin.

The Latin 4th and 5th Declensions also make use of non-moraic inflectional suffixes in the nominative (*-s*) and accusative (*-m*) singular and plural (*-s*)

(119) Case-endings, 4th and 5th Declensions

	4th Declension (u-stems)		5th Declension (ee-stems)	
Nom sg	-s	lacu-s	-s	ree-s
Gen	*-Vs*	lacu-us	*-ii*	ree-ii (reii)
Dat	*-ii*	lacu-ii	*-ii*	*ree*-ii (reii)
Acc	-m	lacu-m	-m	ree-m (rem)
Abl	*-V*	lacu-u	*-ϕ*	ree
Nom pl	*-Vs*	lacu-us	-s	ree-s
Gen	*-um*	lacu-um	*-rum*	ree-rum
Dat	*-bus*	lacu-bus	*-bus*	ree-bus
Acc	*-Vs*	lacu-us	-s	ree-s
Abl	*-bus*	lacu-bus	*-ibus*	ree-bus

Note that long vowels shorten before other vowels (*reeii* -> *reii*) and before *-m* (*reem* -> *rem*).

Adjectival inflection. Case-endings for adjectives are essentially those of nouns and will not be re-discussed here. A number of them, of course, are non-moraic.

Verbal inflection. Latin verbal inflection is extensive and I will not review it all here. For the present, I merely want to establish that it makes use of a number of non-moraic afffixes.

Affixes that consist of a single consonant include *-s* PERFECT and -n- PRESENT. The latter is found only in isolated words and is probably not a synchronic affix in Latin. Examples of each appear below. (Note that *-n-* is, or was, an infix).

(120) Non-moraic tense-aspect markers

Perfect *-s*		Present *-n-*	
carp-s-ii	'I have seized'	find-e-re	'to find' (fid-)
teek-s-ii	'I have touched'	tang-e-re	'to touch' (tag-)

Verbal subject agreement in Latin also makes extensive use of non-moraic suffixes. Consider the following:

(121) Subject agreement: Present Active

		Indicative (am-aa-)		Subjunctive (am-ee-)	
Singular	1	*-oo*	am-oo	*-m*	ame-m
	2	*-s*	amaa-s	*-s*	amee-s
	3	*-t*	ama-t	*-t*	amee-t
Plural	1	*-mus*	amaa-mus	*-mus*	amee-mus
	2	*-tis*	amaa-tis	*-tis*	amee-tis
	3	*-nt*	ama-nt	*-nt*	amee-nt

Non-moraic suffixes include *-m* in the 1st singular, *-s* in the 2nd singular, *-t* in the 3rd; and *-nt* in the 3rd plural. Note that *-m* is not a dental and thus that not all non-moraic affixes are dentals; this was not the case for English or AG. The passive utilizes a first-person non-moraic affix *-r*:

(122) Subject agreement: Imperfect Passive

		Indicative (am-aa-baa-)		Subjunctive (am-aa-ree-)	
Singular	1	*-r*	am-aa-ba-r	*-r*	am-aa-re-r
	2 *-ris*	am-aa-baa-ris		*-ris*	am-aa-ree-ris
	3	*-tur*	am-aa-baa-tur	*-tur*	am-aa-ree-tur
Plural	1	*-mur*	am-aa-baa-mur	*-mur*	am-aa-ree-mur
	2	*-minii*	am-aa-baa-minii	*-minii*	am-aa-ree-minii
	3	*-ntur*	am-aa-baa-ntur	*-ntur*	am-aa-re-ntur

The final use to which non-moraic suffixes are put in the verbal system of Latin is in the creation of participles from verb stems. Three of the four participles in Latin are formed from non-moraic suffixes:

(123) Latin Participles

Present (-nt)	*Future (-tuur)*
ama-nt-is 'is loving'(gen)	fu-tuur-us 'what is to be'
mone-nt-is 'is reminding' (gen)	fac-tuur-us 'to do'
Perfect (-t)	*Gerundive (-nd)*
amaa-t-us 'was loved'	ama-nd-us 'must be loved'
moni-t-us 'was reminded'	mone-nd-us 'must be reminded'

Before closing this section on Latin, I should point out remnants of some earlier inflectional affixes that remain in a number of Latin verbs. Among these are the affixes below:

(124) Deverbal verbal suffixes

Inceptive -sc

calee-sc-oo 'I grow warm'	*cf.*	cale-oo 'I am warm'
labaa-sc-oo 'I begin to totter'	*cf.*	lab-oo 'I totter'
scii-sc-oo 'I determine'	*cf.*	sci-oo 'I know'

Intensives -t

iac-t-oo 'I hurl'	*cf.*	iaci-oo 'I throw'
dormii-t-oo 'I am sleepy'	*cf.*	dormi-oo 'I sleep'
veendi-t-oo 'I try to sell'	*cf.*	veend-oo 'I sell'

It is clear that these affixes violate the monomoraic minimal affix requirement I have argued for in Latin. But are they derivational or inflectional? The answer is clearly inflectional.

Inceptive *-sc-* occurs only in the present tense and is thus in complementary distribution with past tense markers; it must therefore be inflectional, as suspected. Intensive *-t* is related to the perfect participle formant *-t* and is in complementary distribution with other participial markers; it too is clearly inflectional.

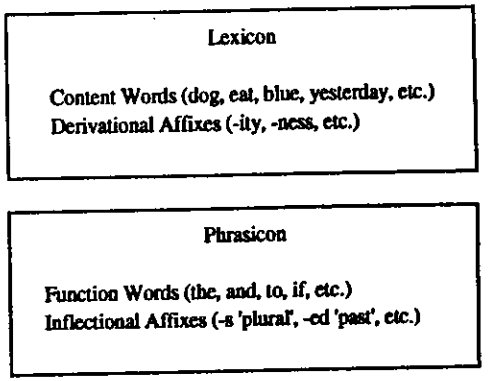
2.4 Conclusion. English, Ancient Greek and Latin have remarkably similar minimal prosodic requirements. The only difference in this respect among the languages is whether word-final consonants are extrametrical or not. Once this has been factored out, minimal word and affix requirements for all three languages are identical:

(125) Minimal prosodic requirements in English, Ancient Greek, Latin
Affix = μ
Word, Root = μ μ

The counterexamples to (125) are legion, but they fall into two categories.

Affixes that fail to meet the μ requirement are all inflectional affixes; words and roots that fail to meet the μμ requirement are all function words. If inflectional affixes and function words are not stored in that part of the grammar in which derivational affixes and content words are stored, this is to be expected. The 2 Lexicon Hypothesis claims not just that content words and function words are different, nor that derivation and inflection are different. The claims it makes is that content words and derivational affixes form a *natural class* and that function words and inflectional affixes form a different natural class:

(126) The 2 Lexicon Hypothesis



The 2LH also asserts that these natural classes are central to the organization of the grammar: they form separate *components* of the grammar. Minimal parosodic requirements, I have argued, are met only in one of these components and thus affect only on of these natural classes.

Taken together, (125) and (126) can be unpacked into a list of separate claims, as follows.

(127) Of English, Ancient Greek and Latin...
a. None has lexical words that consist of less than two moras.
b. None has lexical roots that consist of less than two moras.
c. None has derivational affixes that consist of less than one mora.
d. All have non-lexical words that consist of less than two moras.
e. All have non-lexical roots that consist of less than two moras.
f. All have non-lexical affixes that consist of less than one mora.

It is highly unlikely that the statements in (127) should be true *coincidentally* of these three distantly related languages. The 2 Lexicon Hypothesis provides a straightforward model to account for all of them at once.

3. *Level-Ordered Lexical Insertion*

In Chapter 2, I presented evidence for the existence of two lexicons in the grammars of English, Ancient Greek and Latin. In the present chapter I will provide evidence that the elements of these lexicons are inserted into syntactic structure at different stages. This I will call Level-Ordered Lexical Insertion (LOLI). As the review in Chapter 1 made clear, the idea itself is not new: similar proposals have been made in the processing and aphasic literature (Fromkin 1971, Garrett 1975, 1980; Lapointe 1985; Lapointe & Dell 1989); in the syntactic literature (Stockwell, Schachter and Partee 1972 [function words only]; Emonds 1985); and in the morphology literature (Anderson 1977, 1982 [inflection only]). The idea has yet to be exploited, however, in *phonology*.

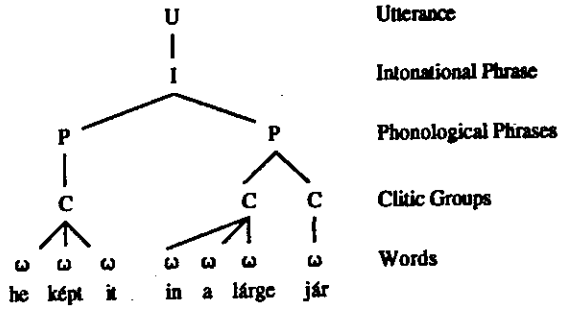
The late insertion of function words is crucial for the construction of prosodic constituents above the level of the phonological word. Recall that on the 2-Lexicon model proposed here, prosodic constituency is built on a partial phonological representation--i.e., on lexical stems organized into an S-structure representation. Now it has been claimed that these prosodic constituents--phonological phrases, intonational phrases and so on--are constructed *as if* function words were invisible (Selkirk 1984). LOLI makes it possible to change the *as if* to a more explanatory *because:* function words and inflectional affixes take no part in defining prosodic constituents above the word because they have not yet been inserted at the stage at which those constituents are formed. And LOLI is not motivated *solely* on phonological grounds: morphological, syntactic, processing and aphasic evidence points to the same thing.

The framework I will assume here is that of *Prosodic Phonology* (Selkirk 1978, 1980; Nespor & Vogel 1982, 1986; Hayes 1989; see also Inkelas & Zec (eds.) 1990 for a collection of recent work). I will begin by discussing the general framework and incorporating LOLI into it (3.1). From there I will move to a discussion of phrasal stress that illustrates the construction of prosodic constituents with LOLI (3.2). I end with a discussion of reduced forms of function words, a topic which also involves LOLI (3.3).

3.1 Prosodic Phonology and LOLI

The theory of Prosodic Phonology organizes phonological representations into constituents that comprise a *prosodic hierarchy*. The prosodic constituency of a sentence like *He kept it in a large jar* may be represented as in (1) (Hayes 1989, Nespor & Vogel 1986).

(1) The Prosodic Hierarchy



Each morpho-syntactic word (*he, kept, it,* etc.) constitutes a phonological word (ω). ωs are grouped into Clitic Groups (Cs), Cs into Phonological Phrases (PPs), PPs into Intonational Phrases (IPs), IPs into Utterances (Us). (See Nespor & Vogel 1986 for a more complete discussion of these constituents; I have left out the syllable and the foot from discussion here for clarity.)

Among other things, prosodic constituents mark natural breaks for pauses. Thus, the sentence aboves may be spoken without a pause, with one pause, etc.:

(2) Pausing and prosodic constituency

He kept it in a large jar	(I)
He kept it...in a large jar	(P...P)
He kept it...in a large...jar	(C...C...C)
He...kept...it...in...a...large...jar	(ω...ω...ω...ω...ω...ω...ω)

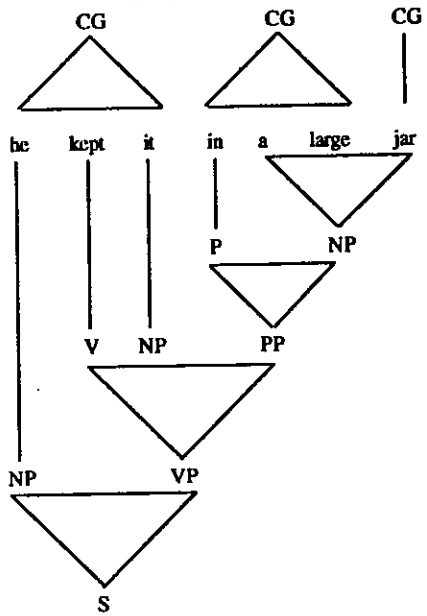
Although any of the phrasings above is natural, ones that violate prosodic constituents are generally ill-formed:

(3) Pausing vs. prosodic constituency
?He kept...it in...a large jar
?He kept it in a...large jar
?He kept...it in a...large...jar
etc.

Prosodic constituents serve two other main functions. First, each prosodic constituent defines a domain in which certain phonological rules are operative. German final obstruent devoicing, for instance, applies at the end of the prosodic domain defined by the constituent ω. Second, each prosodic constituent constitutes a domain in which exactly one unit is prominent. In English, for example, the constituent ω defines a stress domain in which only one syllable is prominent i.e., receives primary stress; the constituent C defines a domain in which a single ω is prominent: (he képt it) (in a lárge) (jár); and so on.

A crucial observation about the prosodic hierarchy is that the constituents that make it up are *not* isomorphic to syntactic constituents. Notice, for instance, that [he képt it] and [in a lárge] form prosodic constituents but not syntactic constituents:

(4) Prosodic constituency ≠ syntactic constituency



A more principled difference between prosodic and syntactic constituents is that only the latter seem to allow embedding: e.g., in *the dog with a hat with a flower*, the NP *dog* dominates the NP *hat*, which dominates the NP *flower*. Prosodic constituents, on the other hand, are hypothesized to obey *strict layering* (Selkirk 1984), such that constituents of one type are all immediately dominated by constituents of the next type on the hierarchy: the ωs are all gathered into non-overlapping Cs, which are all gathered into non-overlapping Ps, etc.

Although prosodic constituents are not the same as syntactic constituents, they are defined in terms of them. That is, at least some information about syntactic constituency is required in constructing prosodic constituency. The algorithms by which syntactic constituency is rewritten as prosodic constituency mark the interface between syntax and phonology. How is this rewriting done?

Building Prosodic Constituents without LOLI

For the sake of exposition, I will assume the version of the Prosodic Hierarchy given above in (1). The existence of some of these constituents, especially the Clitic Group, may well be debated, but the discussion which follows will not be crucially affected by this. What is said below of the constituent C holds equally well for the constituent ω, for those formulations of the prosodic hierarchy that omit it. Central to the discussion at hand is the fact that on any formulation of the prosodic hierarchy, function words are systematically *subordinated* to content words. Whether the constituent that immediately dominates e.g., *the boy* is C or ω is not relevant to the point at hand.

Let us begin, then, with the general form of the algorithms which build prosodic structure, which Nespor & Vogel give as follows (1986:7):

> *Prosodic Constituent Construction*
> Join into an n-ary branching X^P all X^{P-1} included in a string delimited by the definition of the domain of X^P.

X^P refers to some prosodic constituent; X^{P-1} is thus the prosodic constituent immediately under X^P, as defined by the prosodic hierarchy. In all work to date in Prosodic Phonology, prosodic constituents are built off of S-structure, with all words and affixes (grammatical and lexical) in place. Let us see how this is done first, so that later the advantages of incorporating LOLI into these derivations will be clear.

In the simplest case, ω, the domain of the constituent is simply the terminal elements of the syntactic tree.

(5) ω-domain construction: each terminal elements of the syntactic tree is a ω

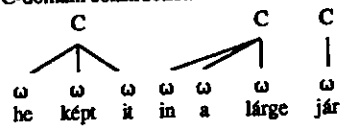
ω	ω	ω	ω	ω	ω	ω
he	kept	it	in	a	large	jar

The next constituent to be formed is the Clitic Group. Hayes (1989) defines a clitic group "roughly as a single content word together with all contiguous grammatical words in the same syntactic constituent". His formal definition of a C is as follows:

(6) Clitic Group Formation
- a. Every content word (lexical category) belongs to a separate Clitic Group.
- b. Defn.: The host of a Clitic Group is the content word it contains.
  Defn.: X and Y share category membership in C if C dominates both X and Y.
  Rule: Clitic words are incorporated leftward or rightward into an adjacent Clitic Group. The group selected is the one in which the clitic shares more category membershiups with the host.

According to this formulation, *kept, large* and *jar* belong to separate Cs and are the *hosts* of those Cs. The function words *he, it, in* and *a* are incorporated into the Cs with which they share the most dominating syntactic nodes: *in* and *a* are dominated by the Prepositional Phrase node that dominates *large*, [PP in a large], and therefore form a C with *large*; *it* is dominated by the VP node that dominates *kept*, [VP kept it...], and so forms a C with *kept*; *he* is dominated by the S node that dominates *kept*, [S he kept it...], and so forms a C with *kept* (and *it*).

(7) C-domain construction



Note that on a strict reading of "Lexical", the preposition *in* would form its own clitic group. Nespor & Vogel (1986) suggest that Lexical be interpreted as a head "with at least one positive specification according to the categorial feature system". Thus only N, A, V are to count as lexical, P falling out of the class. Presumably this account could be extended in some way to distinguish between proper names (which form their own Cs) and pronouns (which do not), both of which are NPs.

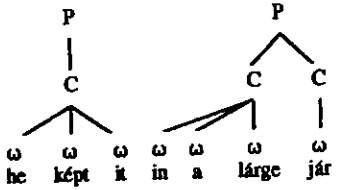
What is crucial here is that the rule that forms clitic groups must check the syntactic category of *each* ω in the string. Furthermore, the rule serves only to demote function words so that they will be invisible to future rules: the lexical word in a C is the head--any function words in the C are non-heads.

Construction of the Phonological Phrase also involves explicit mention of the content word vs. function word distinction and it too serves to demote function words. Following Nespor & Vogel (1986) the domain of the Phonological Phrase can be expressed as:

(8) Phonological Phrase domain
The domain of PP consists of a C which contains a lexical head (X) and all Cs on its nonrecursive side up to the C that contains another head outsice of the maximal projection of X.

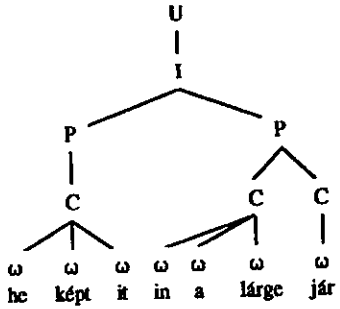
This converts ω-level structure into the following PP-constituents:

(9) PP-domain construction



PPs are then grouped into Intonational Phrases,which are then grouped into one or more Utterances by rules which need not concern us here (see Nespor & Vogel 1986 for formulation and discussion).

(10) Intonational Phrase and Utterance-domain construction



It is important to note that function words are *irrelevant* for determining the number or distribution of prosodic constituent (other than ω) in a sentence. Cs consist of exactly one content word each and it is these Cs that the prosodic constituency of the rest of the sentence depends on. Since PPs consist exclusively of Cs, function words are irrelevant for determining the distribution of PPs in a sentence; likewise for IPs and for Us.

(11) The Prosodic Hierarchy without function words

U	consists of one or more IP
IP	consists of one or more PP
PP	consists of one or more C
C	consists of exactly 1 Content Word

Essentially, function words play no role in creating prosodic constituency.

Selkirk's "Principle of Categorial Invisibility of Function Words"

An especially elegant formulation of prosodic constituency is Selkirk's (1986) end-based theory, in which "the relation between syntactic structure and prosodic structure above the foot and below the intonational phrase is defined in terms of the *ends* of syntactic constituents of designated types (1986:385). In the typical case, clitic groups are defined by left or right syntactic brackets labeled $[_{X^0}$ or $]_{X^0}$, where X^0 is the syntactic notion *head*; phonological phrases, on the other hand, are marked by brackets labeled $[_{X_{max}}$ or $]_{X_{max}}$, where Xmax is the syntactic notion *maximal projection*.

The analysis only works if all function words are left out of the mapping from syntax to prosodic constituency. Selkirk notes that "function words are not identified as 'real' words and so do not count" in the mapping from syntactic to prosodic structure (1986:387). Since the effect of this not counting as a real word is so widespread, she raises it to the status of a principle of grammar (Selkirk 1984):

> What we suggest is that the syntactic category labels for function words are simply "invisible" to principles of the syntax-phonology mapping. This means tha if a function word has the labeled bracketing (a) or (b), it will be treated as though it had the labeled bracketing (c). (FW stands for the syntactic category feature complex(es) of function words.)
> a. $_{fw}$[$_{root}$.....]$_{fw}$$_{root}$ b. $_{fw}$[$_{word}$.....]$_{fw}$$_{word}$ c. [.....]
> This is the *Principle of the Categorial Invisibility of Function Words* (PCI). Given this general principle, any rule that crucially mentions the category name or type associated with the labeled bracketing in its structural description will simply not apply to function words. (1984:343)

The usefulness of the PCI can be seen in how it simplifies the writing of prosodic costituency formation rules. If function words are invisible to these rule the C and PP-formation rules may be simplified as follows (Selkirk 1986).

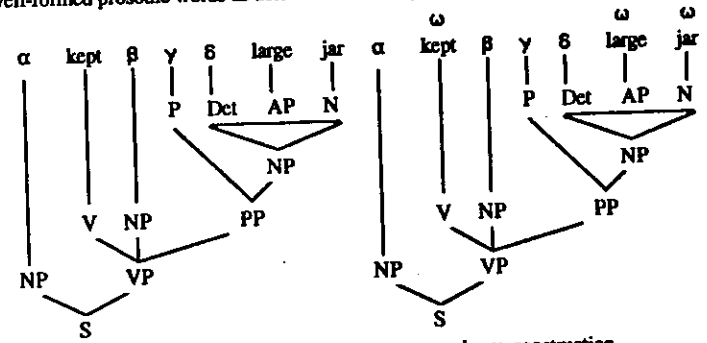
(12) Clitic Group: $]_{X^0} =]_C$

(13) Phonological Phrase: $]_{X^{max}} =]_P$

That is, the rightmost edge of a real word defines the rightmost edge of a clitic group (12) and the rightmost edge of the maximal projection of a real word defines the rightmost edge of a phonological phrase (13).

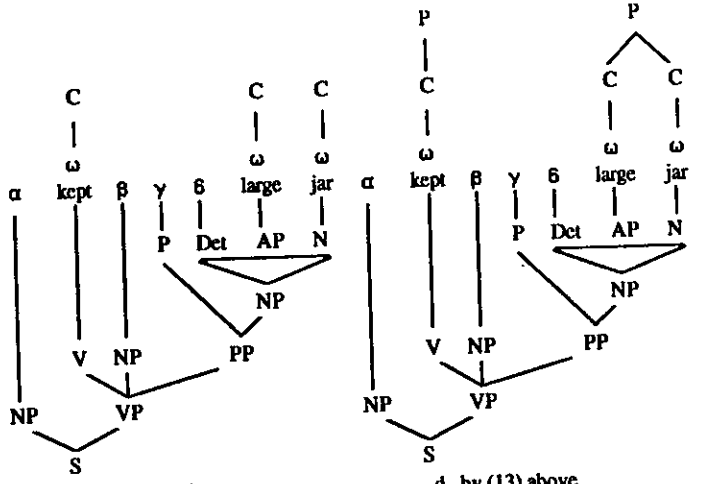
The PCI is clearly very useful in simplifying rules that construct prosodic constituents. What is disturbing about it is its completely stipulative nature within Selkirk's (or anyone else's) framework. Why should function words should be made *invisible* by the syntactic category features they bear?

LOLI allows us to *derive* Selkirk's PCI rather than stipulate it: the phonological content of function words (and inflectional affixes) is not present at the stage at which prosodic constituents are built. According to the PCI, it is the *presence* of syntactic features that makes function words invisible to phonological rules. LOLI allows us to say that it is merely the *absence* of any phonological material at a certain level in the derivation that makes grammatical items "invisible" to phonological rules. Consider a LOLI derivation of the prosodic constituents in *he kept it in a large jar*. S-structure is a partial phonological repreacntation with only lexical stems represented phonologically (a). Function words are marked here by Greek letter place-holders that stand for morpho-syntactic bundles of features. Notice that the first stage in the construction of prosodic constituents (b) is *guaranteed* by the minimal word requirements discussed in Chapter 2. That is, the minimal word and affix requirements operative in the Lexicon insure that the input to prosodic constituency construction consists of well-formed prosodic words as defined for the language in question.



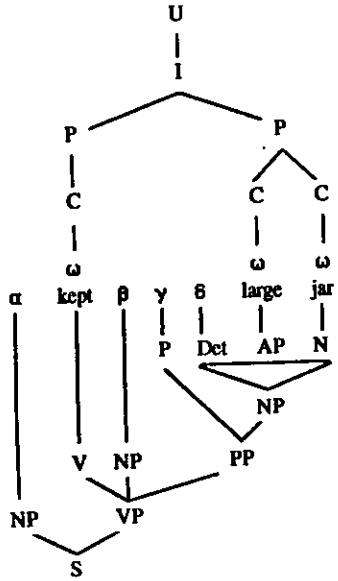
a. Partial Phonological Representation b. ω-construction

Each phonological word becomes the head of a clitic group (c). Clitic Groups are joined into phonological phrases (d), and ultimately into intonational phrases and utterances (e).



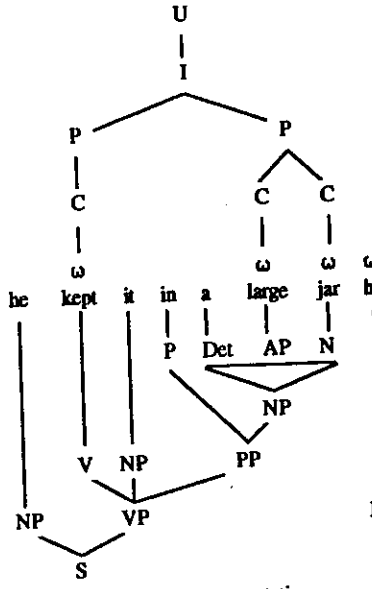
c. by (12) above

d. by (13) above

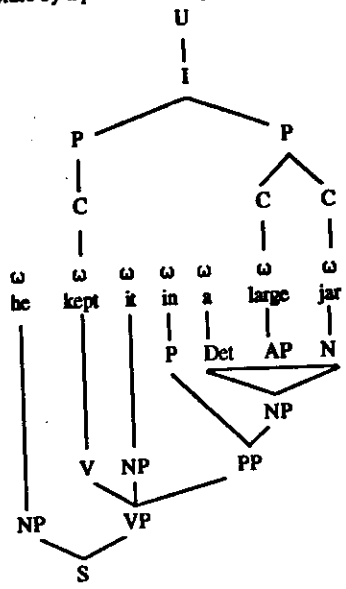


e. PPs -> IPs -> UP

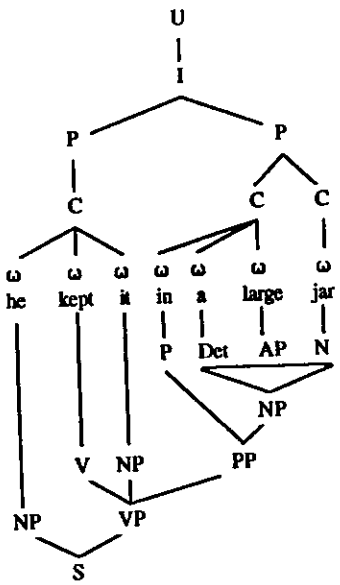
This completes the construction of prosodic constituents. When function words and inflectional affixes are added to yield a full phonological representation (f), they are assigned ω-status and accomodated by the prosodic structure by a process of stray adjunction (h).



f. Full Phonological Representation



g. ω-construction



g. Stray Adjunction

Two aspects of the derivation above require special comment here. The first is the iteration of ω-construction in (f). I will argue below that this is not the usual case and only occurs in fairly careful speech; in normal and rapid speech function words are stray adjoined at levels lower than ω (see below 3.3)

The second is the complex nature of stray adjunction in (g). Note that stray adjunction must take into account both syntactic constituency and prosodic constituency: the prepositon *in*, for example must i) be adjoined to a prosodic constituent C and ii) be adjoined within the syntactic constituent PP (*he kept it in...a large jar). Although the phonology will be satisfied with adjunction to any C, the syntax requires that it be the C that follows rather than the C that precedes. Speech error evidence for this analysis will be given in Chapter 5.

3.2 Phrasal stress

In most language, function words form prosodic constituents with adjacent content words; in many languages, including Enlgish, these constituents are the domain of sentence-stress assignment. Normal intonation places noticably more stress on content words (kept, large, jar) than on function words (he, it, in, a).

(14) Stressless function words: He kèpt it in a làrge jár.

How can LOLI be used to derive stressed content words and stressless function words?

The wrong approach

In earlier work (Golston 1990, 1991) I proposed LOLI could directly derive the streaslessness of function words in English and in Ancient Greek. The idea was to order phrasal stress rules after content word insertion but before function word insertion:

(15) LOLI and phrasal stress

a. Insert content words		kept			large	jar
b. Assign phrasal stress		képt			lárge	jár
c. Insert function words	he	képt	it	in	a	lárge	jár

This analysis, though simple, cannot be completely right for two reasons. First, the second stage of lexical insertion inserts not only function words but inflectional affixes as well. In a highly inflected language like Ancient Greek, inflectional affixes affect the location of primary word stress. Since phrasal stress is realized on the primary stressed syllable of a word, the latter must be determined before the former. Thus inflectional affixes must be added before phrasal stress is assigned. Consider the following one word sentences:

(16) a. é-lip-on
PAST-leave-1sg
'I left'

x
. (x .)
e li pon
| |
H L*

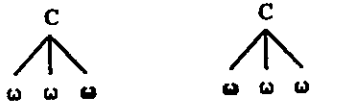
b. e-líp-omen
PAST-leavc-1pl
'We left'

```
          x
(x .)   (x  .)
e  li   po  men
   |    |
   H    L*
```

(Columns of x's mark prominent syllables; (x .) designates a left-prominent foot or trochee; L* marks the prominent syllable in each word, H marks the syllable immediately *before* that syllable (Sauzet 1989).) The penultimate syllables are stressed in both words, *élipon* and *ellpomen*. But notice that in the singular form the penult is part of the lexical stem *lip*, while in the plural it is part of the inflectional suffix *-omen*. The latter case makes it clear that inflectional affixes must be present when phrasal stress is assigned.

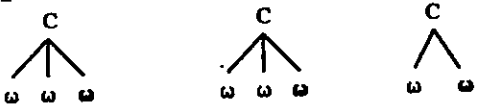
Additional evidence against the analysis in (15) comes from stressed clitic pronouns (examples and discussion based on Nespor & Vogel 1986:155ff). Consider the following sequences of content word plus clitic(s) in Italian and French:

(17) Italian



a. "(he) gives it to me"' b. "give it to me"'

(18) French



a. ".... gives it to me"' b. "go away"' c. "take it"'

(Stressed ωs in boldface.)The Italian cases are easy to derive with the analysis in (15): stress is assigned to the content word when no function words are present; function words are added later but phrasal stress does not reapply. But this will not work for French, which stresses the

rightmost element in the Clitic Group ragardless of whether it is a content word (a) or a function word (b, c).

A better approach

The problem with the analysis in (15) is that it invokes LOLI as a means of assigning phrasal stress rather than as a means of building phrasal constituents. Once the correct constituents are built, phrasal stress is relatively straightforward. The Italian cases involve stressing the host (or head) of the clitic group (Hayes 1989), while the French cases involve stressing the right-most member of the clitic group: the phrasal domain is the same for both languages, but the rule assigning relative prominence within that domain is different.

Now consider English phrasal stress. Again, phrasal constituents are built before function words (and inflectional affixes) are inserted. But phrasal prominence is not assigned until after funtion words are in place. As in Italian, and in most languages according to Nespor & Vogel (1986:155), phrasal prominence within the clitic group is assigned to the host: i.e., to kept, large and jar.

(19) LOLI and phrasal stress

a. Insert content words		kept				large	jar
b. Form prosodic constituents		[kept]				[large]	[jar]
c. Insert function words	he	[kept]	it	in	a	[large]	[jar]
d. Adjoin function words	[he	kept	it]	[in	a	large]	[jar]
e. Assign phrasal stress	[he	képt	it]	[in	a	lárge]	[jár]

Now let us return to Ancient Greek inflection. Prosodic constituents are built on partial phonological representations consisting only of lexical stems (*lip* 'leave'); then function words and inflectional affixes are added and adjoined.

(20) Derivation

a. Insert content words	lip	lip
b. Form prosodic constituents	[lip]	[lip]
c. Insert function words, inflection	e [lip] on	e [lip] omen
d. Stray Adjunction	[e lip on]	[e lip omen]

e. Assign lexical prominence	x (x .) [e lip on]	x (x .) (x .) [e lip omen]
f. Tone Assignment	x (x .) [e lip on] | | H L*	x (x .) (x .) [e lip omen] | | H L*
g. Assign phrasal prominence	x x (x .) [e lip on] | | H L*	x x (x .) (x .) [e lip omen] | | H L*

(See Sauzet 1989 for details of tone-assignment and lexical prominence rules.) At this point, lexical prominence (penultimate stress in this case) is assigned, followed by phrasal prominence, which is built "on top of" lexical prominence.

In this way, LOLI is used to derive stressless function words, albeit indirectly. Phrasal constituents are built on lexical stems, function words and inflectional affixes are inserted and adjoined, and prominence relations are determined on the basis of phrasal constituency.

3.3 Reduced forms of function words

In addition to providing a treatment of unstressed function words, LOLI can also be used to treat reduced forms of function words. In many language function words but not content words have 'reduced' forms (cf. Chapter 2 above). In English, for instance, the modal *will* reduces to syllabic [l] in normal speech and the auxiliary *is* occurs in the reduced from [z] or [s]:

(21) Reducible function words:
Max will be hoping Tom is asleep. (careful speech)
Max'll be hóping Tóm's asléep. (normal speech)

Content words do not undergo such reduction: e.g., the noun *will* (*last will and testament*) does not reduce to a syllabic [l] in normal speech.

Most analyses of reduced function words (e.g., Selkirk 1984) have treated them as exactly that: full forms that are *reduced* by rule in certain environments. There is something unnatural about such analyses, however: at least in English, all but the most formal registers *prefer* reduced function words. This is especially relevant in production: does it make sense to say that a speaker prepares to say *Max will be hoping Tom is asleep* and then goes to the trouble of reducing it to *Max'll be hoping Tom's asleep*? Why treat the full forms as basic?

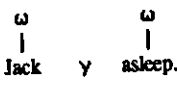
An obvious answer is that only the full forms preserve unpredictable information: if *is* were to be derived from [z], how would we know what vowel to insert? The reduced form must come from the unreduced form because the opposite cannot be true. Intuitively, however, [z] is a simpler (sloppier, etc.) form of *is*. It is not so much that we go through an extra step just to make *is* shorter: we just don't put in the effort to say the whole thing and end up with only the [z].

How may this conflict be resolved? The role of LOLI in the construction of prosodic constituents above the word provides a useful clue.

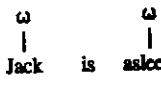
Building prosodic constituents above the word requires an even foundation--each level of structure must be built directly on an earlier layer, like bricks in a wall (Selkirk's Strict Layer Hypothesis). For this reason, formulations of ω-level structure have assigned the category ω to every lexical terminal; this allows clitic groups to exhaustively dominate phonological words; and phonological phrases to exhaustively dominate clitic groups; etc. But LOLI obviates the need for assigning the category ω to function words. If prosodic structure is built only on lexical stems, function words may enter the derivation later as whatever they like. If they enter as consonants, they are adjoined into syllable; if they enter as syllables, they are adjoined into ωs; if they enter as ωs, they are adjoined into clitic groups (cf. Berendsen 1986).

Consider, for example, *Jack is asleep*:

(22) Jack ɣ asleep. Partional Phonological Representation

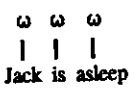


Prosodic Constituents built

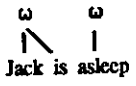


Full Phonological Structure

At the final stage in (22) at least two possibilities arise. The speaker could (i) save the whole word *is* or, (ii) save only part of the word. That is, the speaker can either add *is* as another prosodic word (23i) or just incorporate part of it into one of the prosodic words already present (23ii).

(23) (i) 

Jack is asleep

(ii) 

Jack's asleep

Clearly option (ii) requires less effort and planning. The speaker adds as little to the previously built structure as possible. (23i) is characteristic of careful or formal speech; (23ii) models normal to rapid speech.

Reduced Function Words in English

It should be noted, first, that not all function words in English are reducible. Reduction is lexically idiosyncratic. Consider the following prosodically identical pairs of words.

(24)	Reducible	Non-Reducible	Prosodic Composition
	would ('d)	what	CVC
	has ('s)	his	CVC
	am ('m)	on	VC
	is ('s)	as	VC

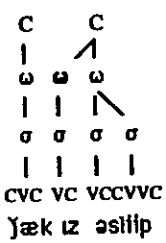
Those function words that do reduce, however, tend to reduce in the same way. First, more sonorant segments are deleted while less sonorant segments are retained; second, it is the right-most segments that are retained.

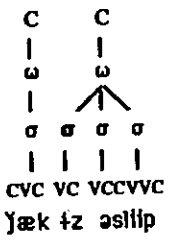
(25) Reduction and Sonority

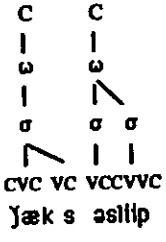
Obstruents never[1]:	should, this, to, does	
Nasals never[2]:	in, on, can, than	
Liquids never:	will, all, or, were	
Glides, h sometimes:	would ('d), will ('ll), him ('im)	(reducing)
	we, why, you, who	(non-reducing)
Vowels often:	am ('m), is ('s), or (r), and (n)	(reducing)
	at, off, as	(non-reducing)

That is, the segments that are easiest to add to existing prosodic structure, namely single consonants, are retained, while those that require adding another syllable to existing prosodic structure, namely vowels, are often deleted.

How much of a word is retained is a function of the rate of speech. For the sake of exposition, let us assume three rates of speech here: careful, normal and rapid. In careful speech, a function word is granted full ω status and adjoined into an existing clitic group (26i); in more normal speech, function words are assigned syllable status and adjoined into existing ωs (26ii); in rapid speech they are assigned only status as strings of consonants (Cs) and vowels (Vs), in which case they must be adjoined into existing syllables.

(26) (i) 

(ii) 

(ii) 

In careful speech, is is given full word status and retains its full vowel. In normal speech, it is accorded only syllable status and is realized with a high central vowel. In rapid speech, it is

1 The one exception is *of*, which generally deletes before C-initial words and is retained before V-initial words: *friend o' Bob's, friend of Ed's.* (see Selkirk 1972).

2 The one exception is *an*, clearly a sandhi form (see Rotenberg 1978).

given only segmental status and must be incorporated into an existing syllable; a syllable-internal rule of voicing assimilation changes the underlying */z/* to [s].

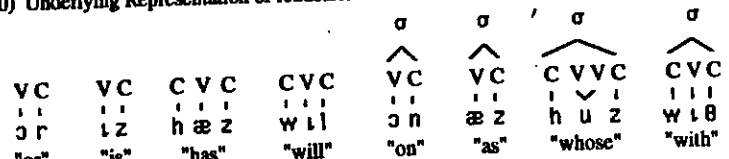
Note that a stray consonant is *enclitic*, i.e., is adjoined leftward into a preceding syllable, while stray syllable and words are *proclitic*, i.e., are adjoined rightwards into following words and clitic groups. This split occurs generally when the function word is syntactically bracketed with the following word:

(29)

	[Pat]$_{NP}$	[has won]$_{I'}$	[Joe]$_{NP}$	[would go]$_{I'}$	[I]$_{NP}$	[will leave]$_{I'}$
Careful:	[pæt]	[hæz wən]	[ʃow]	[wʊd go]	[aj]	[wɪl liv]
Normal:	[pæt]	[ɨz wən]	[ʃow]	[wɨd go]	[aj]	[wɨl liv]
Rapid:	[pæts]	[wən]	[ʃowd]	[go]	[əjl]	[liiv]

Recall that not all function words in English have reduced counterparts. These irreducible function words may be represented as syllables *underlyingly*. The underlying differences between reducible and irreducible function words may then be represented as follows:

(30) Underlying Representation of reducible and non-reducible function words

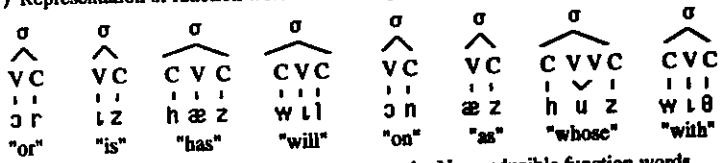


a. Reducible function words b. Non-reducible function words

That is, in rapid speech the words in (30a) are generally reduced to sub-syllabic strings. But at the same rate of speech, the words in (30b) are not so reduced.

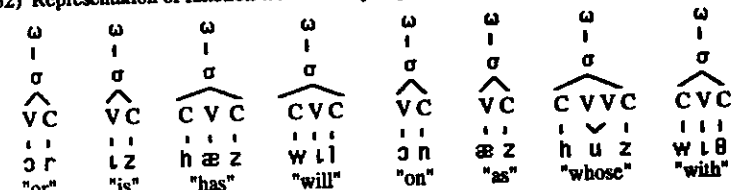
In what I have called normal speech, the difference is neutralized and all of the forms in (30) surface as full syllables (31). Similarly for careful speech, where function words surface as full words (32).

(31) Representation of function words: *normal speech*



a. Reducible function words b. Non-reducible function words

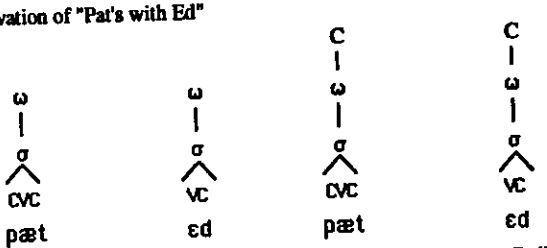
(32) Representation of function words: *careful speech*



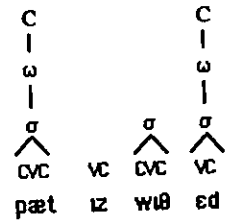
a. Reducible function words b. Non-reducible function words

We are now in a position to give a full derivation of a sentence with both reduced and stressless function words in what I have called rapid speech. Take the sentence ***Pat's with Ed*** (prosodic structure above the clitic group omitted for clarity):

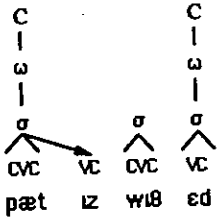
(33) Derivation of "Pat's with Ed"



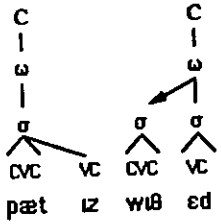
a. Partial Phonological Representation b. Prosodic Constituents Built



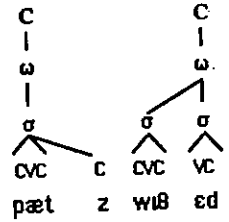
c. Function Words Inserted



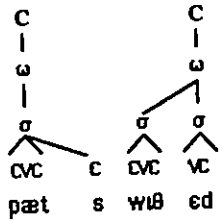
d. Stray Adjunction to σ



e. Stray Adjunction to ω



f. Deletion of Unlicensed Segments



g. Voicing Assimilation

The derivation proceeds as follows: beginning at Partial Phonological Representation (a), prosodic constituents are built (b); function words and inflectional affixes are inserted (c); stray segments are adjoined to syllables (d); stray syllables are adjoined to words (e); segments that are not prosodically licensed are deleted (f); and regular phonological rules such as voicing assimilation apply (g).

The 'reduction' of function words in fast speech on this view is not accomplished by a reduction rule but by a rule that saves prosodic strays. The process is characterized not as one that *destroys* structure and deletes information ('reduction') but one that *preserves* structure and

retains information. The number of syllables and prosodic words is undiminished from the time function words are added to the surface representation; reduction is thus a method of retaining the maximal amount of information with the minimal amount of work.

Reduced function words in Ancient Greek

There is a process in AG traditionally referred to as Crasis ('mingling') in which the final vowel of a function word coalesces with the initial vowel of the following word (Smyth 1920:§173; Vendryes 1945:§320; Bally 1945:§237; Allen 1973:228; Sommerstein 1973:59, etc.). Crasis is similar to the reduction of English function words and can be modeled using the same underlying structure used for English reducible function words. An important point to notice is that the output of crasis has one less syllable than the input:

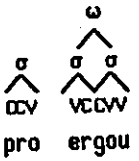
(34) Crasis

a.	to ónoma the name	->	toúnoma	(o + o = ou)
b.	tée heeméra the day (dative)	->	theeméra	(ee + ee = ee)
c.	prò érgou for use 'useful'	->	proúrgou	(o + e = ou)

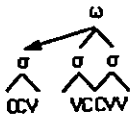
Crasis is merely the prosodic incorporation of a function word into a following word: it follows straightforwardly, via stray-adjunction, from a representation in which the function word consists of only a syllable (cf. English *non*-reducing function words). Consider the following, somewhat truncated, derivation:

(35) Derivation of proúrgou < /prò érgou/ 'for use'

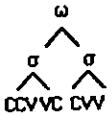
a. Partial Phonological Representation



b. Function words inserted



c. Stray-adjunction to ωs

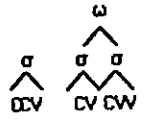


d. 'contraction'

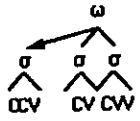
As the derivation makes clear, crasis has two parts: the incorporation of the stray syllable into the following word and the contraction of the (now-)adjacent vowels. For comparison, consider the derivation of a phrase with a consonant-inital content word:

(36) Derivation of *prò phóbou* 'for fear'

a. insert content words



b. insert function words



c. stray-adjoin σs

d. 'contraction': NA

Contraction cannot apply because its structural description (ω-internal adjacent vowels) is not met.

Crasis occurs within Clitic Groups as defined by Hayes (1989): a content word and adjacent function words (the same process occurs word-internally; this will be discussed below). Examples of the different parts of speech to which words undergoing crasis belong bear this out. Crasis occurs between a definite article and the noun that follows, between a relative pronoun and the word which follows, etc.:

(37) Domain of Crasis ("." represents syllable boundary)

a.	definite article			
	[tà ál.la]	->	taál.la	(a + a = aa)
	the others			
b.	relative pronoun			
	[há e.gó]	->	haa.gó	(a + e = aa)
	which I			
c.	pronoun			
	[e.gò óida]	->	e.góoi.da	(o + oi = ooi)
	I think			
d.	enclitic pronoun			
	[soi es.tí]	->	sous.tí	(oi + e = ou)
	you:dat is			
d.	vocative particle			
	[óo á.ner]	->	óo.ner	(oo + a = oo)
	O man			
e.	prepostion			
	[prò é.khoon]	->	proó.khoon	(o + e = ou)
	for having			
f.	conjunction			
	[kaì hi.ke.teú.e.te]	->	khii.ke.teú.e.te	(ai + i = ii)
	and beseech:2sg			

Crasis applies across the boundary between a function word and a content word, but not across the boundary between two content words:

(38) Domain of crasis

a.	FW CW	tée hee.mé.ra	=>	thee.mé.ra
		the day (dative)		
b.	*CW##CW	tée de.ká.tee hee.mé.ra	≠>	*tée de.ká.thee.mé.ra
		the tenth day (dative)		

The domain of crasis is thus internal to the the clitic group.

The domain of contraction, that sub-part of crasis in which adjacent vowels coalesce, is the phonological word ω. This is made evident by the fact that contraction applies not only in the environments in (39), but also across word-internal morpheme boundaries:

(39) Contraction

a.	géra + a	->	géraa	(a + a = aa) cf. taálla

b. tiimá + ete	->	timáate	(a + e = aa)	cf. haagó
c. deelóo + oo	->	deelóo	(o + oo = oo)	cf. egóoida
d. heéroo + a	->	heéroo	(oo + a = oo)	cf. óoner
e. edeélo + e	->	edeélou	(o + e = ou)	cf. proúkhoon

There are a few minor differences between crasis and such morpheme-boundary contraction, but they need not concern us here; a full discussion is found in Sommerstein (1973). His conclusions, however, are worth noting:

> For the most part the rules determining the output vowel [of crasis--C] are the same as for normal contraction, and for this reason the best way of accounting for the general phenomenon of crasis is to have a rule, earlier than Contraction, reducing the word boundaries within such phrases...to formative boundary so that Contraction can operate. (1973:59ff)

If AG function words are entered into the derivation of a sentence not as words but as syllables, as I argued above for English, Sommerstein's special rule reducing the word boundary between a function word and an adjacent content word to a mere formative boundary is not needed. The effect of the rule follows from the underlying status of function words as mere syllables rather than prosodic words.

Reduced function words in Latin

Latin also has a number of function words with reduced forms. Due to the conservative nature of Latin orthography, however, these reduced forms are preserved only in comedies, which were meant to capture the spoken language of the day. The reduced forms which follow all come from the plays of Plautus, but similar forms can be found in the works of other authors who wrote naturalistically.

Vowel-initial forms of the copula are regularly enclitic on the preceding word. When the preceding word is vowel final, the initial vowel of the copula is deleted.

(40)	a.	acceptumst	<	acceptum est	'is accepted'
	b.	lubidost	<	lubido est	'desire is...'
	c.	sist	<	si est	'if (x) is'
	d.	mihist	<	mihi est	'is mine'

e.	stultas	<	stulta es	'you are stupid'
f.	tutes	<	tu te es	'thout thee art'

(n.b.: final *-um* represents a nasalized vowel.) Like Greek crasis, this process involves a syllable-sized function word (*est, es*) being adjoined to the ω that dominates a content word--the second of two adjacent vowels then deletes. In a few cases (c, d, f), two function word syllables are joined together and adjoined to a word. Example (c) is particularly interesting in this regard as it apparently involves the adjunction of four function words to a single content word. The full context of (c) is given below and derived in (42).

(41) sist quod mihi cor
if-is a to me heart
'If I have a heart'

(42) Derivation

a. Function words inserted b. Adjunction to ω c. vowel deletion

In addition to these enclitic forms of the copula, Latin has a proclitic, *sii* 'if' that undergoes the same process:

(43)	a.	sodes	<	sii audes	'if you please'
	b.	sultis	<	sii vultis	'if you (pl) wish'
	c.	siis	<	sii viis	'if you (sg) wish'

Note that intervocalic *v* [w] deletes in (b) and (c). In (43) it is the *first* vowel (the vowel in *sii*) which deletes. What *es, est* and *sii* have in common is thus that their vowel deletes in deference, as it were, to the vowel of the clitic group *host* (the content word).

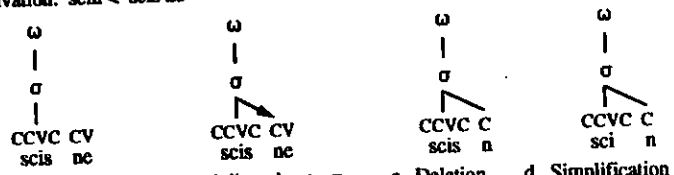
Latin also has an enclitic (*ne* 'not') that reduces to a non-syllabic form. It can be analyzed like English *is* and *has*, i.e., as consisting underlyingly of unsyllabified strings.

(44)	a.	dixin	<	dixi ne	'haven't I told (you)?'
	a.	ain	<	ais ne	'don't you say?'

b.	scin	<	scis ne	'don't you know?'
c.	potin	<	potis ne	'aren't you able?'
e.	audin	<	audis ne	'don't you hear?'

A derivation for *scin* is given below.

(45) Derivation: scin < scis ne



a. Function word insertion b. Adjucntion to σ c. Deletion d. Simplification

As before, the stray segments are adjoined to the syllable node dominating *scis* (b) and the unlicensed final vowel in *ne* is deleted (c); finaly, the illicit coda *-sn* is simplified to *-n* (d).

3.4 Conclusion

Speakers of a language can memorize all the (uninflected) words in their language and the stress patterns that go with them. In a language such as English, they probably do. But speaker *cannot* memorize all the sentence in their language and thus *cannot* memorize the intonational patterns that go with sentences. Thus, although the prosodic constituency of words can be memorized, the prosodic constituency of sentences cannot be. It must be generated on line each time a speaker utters a sentence.

For this reason, the rules which translate syntactic structure to prosodic structure must be maximally simple. Selkirk's end-based theory of prosodic domains looks like a promising characterization of the process a speaker uses to map prosodic constituency onto syntactic constituency. And yet it contains an *ad hoc* stipulation that function words are invisible to these rules; moreover, it is just this stipulation, the PCI, which allows the simple formulations used in the end-based theory. Without the categorial invisibility of function words, the end-based theory of prosodic constituents is as clunky and unwieldy as its predecessors.

LOLI allows us to derive the invisibility of function words rather than stipulate it. The phonological strings that make up function words are 'invisible' to prosodic constituency formation rules because *they are not there* when those rules apply.

4. *Word Formation*

The 2 Lexicon Hypothesis makes strong predictions about what types of word-formation processes should occur in natural languages. In particular, it predicts that content words and derivational affixes form a natural class in word formation and that function words and inflectional affixes form another. This chapter will attempt to substantiate that claim and thus provide support for the 2LH. I will argue that Affixation is different in the Lexicon than it is in the Phrasicon (4.1); that Compounding is different in the Lexicon than in the Phrasicon (4.2); and that other processes of word-formation (acronyms, blends, etc.) found in the Lexicon are not found in the Phrasicon and vice versa. This amounts to saying that *word-formation in general* is different in the Lexicon than it is in the Phrasicon, supporting the distinction between the two lexicons. An interesting result of the discussion in these sections is that some words are hypothesized to be in both the Lexicon and the Phrasicon--this will be called the Double-Listing Hypothesis. Finally, I will show that other types of word-formation including blending, acronyms, portmanteaux, etc. occur either in the Lexicon or the Phrasicon, but not in both (4.3), another claim that supports the 2LH.

4.1 Affixation

There is a vast literature on the differences between inflectional and derivational affixation (see Scalise 1984, Anderson 1988 for overviews). Anderson (1982:) has defined inflectional morphology as "what is relevant to the syntax". This notion was refined by Emonds (1985), though I have disputed some of his conclusions. As discussed above in Chapter 1, I will assume that inflection is "what is relevant to the discourse" in the appropriate sense.

The rest of this section adds additional evidence for a distinction between derivation

(affixation-in-the-Lexicon) and inflection (affixation-in-the-Phrasicon). My general claim will be that derivational affixation is restricted to content words; only function words that are doubly listed in both the Lexicon and the Phrasicon undergo the type of affixation found in the Lexicon. The rareness of derivational affixes that attach to function words is taken as strong support that derivational affixes and function words never 'meet'--i.e., that they are in different lexicons. Inflection, on the other hand, sees both content words and function words, though a language with improverished morphology, such as English does not show this. Ancient Greek and Latin, however, provide abundant evidence of inflected function words.

4.1.1 Affixation in the Lexicon

Derivational affixes create stems from roots or from other stems. Unlike inflectional affixes, they may change the category of the words or roots to which they attach: *-ness* attaches to adjectives to form nouns, *-able* attaches to verbs to form adjectives, *-al* attaches to nouns to form adjectives. Most languages place a rather severe restriction, however, on both the input to and the output of derivational affixation: derivational affixes generally take only content word stems (CWs) and yield only CWs (1a).

(1) Affixation: input, output

	input			*output*	
a.	CW + der affix	=>	CW		COMMON
b.	CW + der affix	=>	FW		RARE
c.	FW + der affix	=>	CW		RARE
d.	FW + der affix	=>	FW		RARE

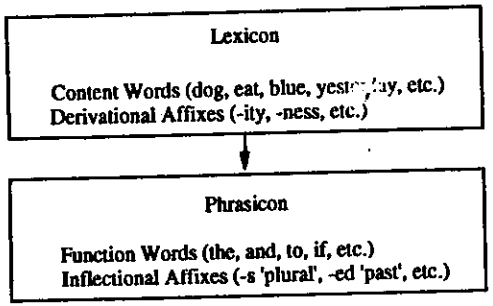
What is *not* generally found are affixes that take CWs as input and yield function words (FWs) (1b), affixes that take function words as input and yield content word stems (1c) or affixes that take function words as input and yield other function words (1d):

Evidence for (1a-d) will be drawn from English, Ancient Greek and Latin. Note that (1b)

and (1d) are related to the claim that function words in most languages are 'closed-class': speakers rarely make up new function words from old function words or from old content words. On a model of grammar in which all words and affixes of a language are stored and processed in the same component, the situation described in (1) is curious. Why should derivational affixes attach only to content words and their roots? And why shouldn't derivational affixation result in the creation of function words?

The 2 Lexicon Hypothesis allows the generalizations in (1) to fall out as a natural consequence of the form of the grammar. On the 2LH content words and derivational affixes are stored and processed in the Lexicon while function words and inflectional affixes are stored and process in the Phrasicon (2).

(2) The 2 Lexicon Hypothesis



The form of the grammar in (2) derives the observations in (1) as follows. (a) CW stems may appear as input to derivational affixation. Derivational affixes have access to CW stems because they are in the same component of the grammar. (b) The output of derivational affixation is limited to CW stems. The creation of FWs in the Lexicon would violate the 2LH, since FWs only appear in the Phrasicon. (c, d) FWs may not appear as input to derivational affixation because FWs and derivational affixes are in different components of the grammar. According to the 2LH, function words and

derivational affixes never 'meet': derivational affixation occurs only in the Lexicon and thus excludes function words in principle.

Generative treatments of English morphology including Aronoff (1976), Selkirk (1982), and Scalise (1984) stipulate that derivational affixes attach *only* to N, A, and V and that the output of such affixation is always N, A or V. These are the claims behind (1a) and (1b). (1a) allows the types of affixation shown below:

(3)	CW	+	der	>	CW	examples
	N	+	der	>	N	king-dom
	N	+	der	>	A	education-al
	N	+	der	>	V	motor-ize
	A	+	der	>	N	fastidious-ness
	A	+	der	>	A	green-ish
	A	+	der	>	V	short-en
	V	+	der	>	N	invent-ion
	V	+	der	>	A	like-able
	V	+	der	>	V	re-arrange

Other languages show the same patterns. Ancient Greek and Latin also productively form nouns, verbs and adjectives stems from other noun, verb and adjective stems (see Chapter 2 for lists); the process is very common in most languages.

(1b) *disallows* the types of affixation in (4). Such formations are generally unattested.

(4) Unattested function words created by derivational affixation

CW	+	der	>	FW
*N	+	der	>	Auxiliary
*N	+	der	>	Modal
*N	+	der	>	Pronoun
*N	+	der	>	Article
*N	+	der	>	Conjunction
*N	+	der	>	Comp
*N	+	der	>	Neg

CW	+	der	>	FW
*V	+	der	>	Auxiliary
*V	+	der	>	Modal
*V	+	der	>	Pronoun
*V	+	der	>	Article
*V	+	der	>	Conjunction
*V	+	der	>	Comp
*V	+	der	>	Neg

CW	+	der	>	FW
*A	+	der	>	Auxiliary
*A	+	der	>	Modal
*A	+	der	>	Pronoun
*A	+	der	>	Article
*A	+	der	>	Conjunction
*A	+	der	>	Comp
*A	+	der	>	Neg

(Prepositions will be treated separately from other function words: see 4.3.)

The 2LH allows us to *derive* the generalizations in (3) and (4) rather than stipulate them. Only CW stems undergo derivational affixation because only CW stems are stored with derivational affixes. FWs are excluded from derivational affixation by the form of the grammar.

Problem cases

The 2LH incorrectly allows two sets of words to undergo derivational affixation: lexical adverbs and irregularly inflected words. Lexical adverbs like *yesterday* and *twice* are stored in the Lexicon according to the 2LH (they are not defined purely by discourse features, do not have reduced forms, are not skipped over in the formation of prosodic constituents, etc.) and thus should be eligible for derivational affixation. But they are not. Irregularly inflected content words (*slept, fell, women, children*) must also be listed in the Lexicon but they too fail to undergo derivational affixation. The problem is shared by other models such as Lexical Phonology: if these words are stored with *dog* and *eat*, they should be able to undergo the same morphological processes.

Irregularly inflected words are the easier case. They are memorized chunks that include both lexical and grammatical material that is no longer availabe for productive processes in the language. In this, they are similar to idioms. *It takes two to tango* is a memorized lexical chunk: modification of any of its parts is anomalous (**It took two to tango, *It takes two to waltz, *It's been taking two to tango for a long time now,* etc.). Just as we do not expect *It takes two to tango* to take derivational affixes like *-ity* and *-ness*, we should not expect memorized chunks like *slept* and *men* to take affixes like those in *sleepy* and *manly*. Part of knowing the words *slept* and *men* involves knowing that they are fully inflected syntactic heads; as such they are as incligible for derivational affixation as fully inflected syntactic phrases like *It takes two to tango.*

The same argument can be extended to lexical adverbs. If regularly formed adverbs like *quickly* are inflected adjectives (see Chapter 1), then irregularly formed adverbs like *once* and *twice* must bear the same relation to *one* and *two* that *slept* and *men* bear to *sleep* and *man*: i.e., they are fully inflected syntactic heads that are memorized as chunks. As such, they are ineligible for word formation processes. Just as a speaker knows that a sentence (however memorized) is not a stem, she knows that an inflected word (however memorized) is not a stem.

Most of (1a) and (1b), then, is straifghtforwardly modeled by the 2LH. Function words are stored and processed in a lexicon distinct from the one that contains content words and derivational affixes. Content words thus form a natural class that is susceptible to derivational affixation, a process which turns CWs into other CWs. More importantly, perhaps, the 2LH ties the exclusion of FWs from derivation to other aspects of the grammar discussed already: first, that function words and inflectional affixes are not subject to minimal prosodic rquirements (Chapter 2) and second, that function words and inflectional affixes are 'invisible' to the construction of prosodic domains (Chapter 3).

It remains to be shown, of course, that function words as a class do not undergo derivational affixation; if they did, the 2LH would be severely undermined because function words (stored in the Phrasicon) and derivational affixes (stored in the Lexicon) should never 'meet' according to the 2LH.

Affixes that derive words from function words are extremely rare. (1c) and (1d) (repeated here) jointly state that function words are not the input to affixation that derives new words from old words.

(5) Affixation: input, output

c.	FW	+	der affix	=>	CW	RARE
d.	FW	+	der affix	=>	FW	RARE

Modeling (5c) and (5d) with a single lexicon requires some stipulation: if derivational affixes attach to some words in the Lexicon, why not to all? The 2LH, on the other hand, captures (5c) and (5d) straightforwardly: derivational affixes do not attach to function words because derivational affixes and function words are stored and processed in different components of the grammar.

With only a handful of exceptions, content words in languages such as English, Ancient Greek and Latin are never composed of a function word and a derivational affix (=5c). The 2LH correctly predicts the non-occurrence of the forms in (6):

(6) Content word types created by derivational affixation (* = unattested)

*Auxiliary	+	der	>	N		*Auxiliary	+	der	>	V
*Modal	+	der	>	N		*Modal	+	der	>	V
*Pronoun	+	der	>	N		*Pronoun	+	der	>	V
*Article	+	der	>	N		*Article	+	der	>	V
*Conjunction	+	der	>	N		*Conjunction	+	der	>	V
*Comp	+	der	>	N		*Comp	+	der	>	V
*Neg	+	der	>	N		*Neg	+	der	>	V

*Auxiliary	+	der	>	A
*Modal	+	der	>	A
*Pronoun	+	der	>	A
*Article	+	der	>	A
*Conjunction	+	der	>	A
*Comp	+	der	>	A
*Neg	+	der	.>	A

Alternative Explanations

I have tried to argue that the non-occurence of the structures in (1b-d) above is evidence for two lexicons: the central fact here is that derivational affixes rarely attach to function words--this is predicted by the form of the grammar given the 2LH, but must be stipulated with a single lexicon.

Undoubtedly, this sort of explanation will strike some as baroque. Isn't an alternative explanation availabe that does not require two lexicons. Perhaps something in the nature of function words is responsible for their inability to host derivational morphology.

Alternative 1: the Unitary Base Hypothesis.

Aronoff has argued that "the syntacticosemantic specification of the base, though it may be more or less complex, is always unique" (1976:48ff). That is, an affix that subcategorizes for, e.g., transitive verbs may not attach to nouns or adjectives and vice versa. If the UBH is correct, it follows that any affix that subcategorizes for nouns, verbs or adjectives cannot attach to pronouns, modals, determiners, conjunctions, etc. The restriction against [FW + der aff] would seem to follow from the UBH; if so, it adds no independent evidence for the 2LH.

But this sort of syntactic feature argument does not go through. The claim in (1c-d) is not that *-ity, -able* and *-ness* do not derive new words from function words but that *no* affixes derive new words from function words. The UBH does not preclude affixes whose sole function is to derive, e.g., nouns from modals; it merely precludes affixes that attach to words of one class from attaching to words of another.

Alternative 2: the Anaphor Hypothesis.

Postal (1969) argued that the inability of (at least certain) function words to host derivational affixes comes from their anaphoric nature. This is a semantic argument. On

the basis of sentences like those in (7) and (8), Postal claimed that lexical items in general are *anaphoric islands*, "where such an enity is a sentence part which cannot contain an anaphoric element whose antecedent lies outside of the part in question and which cannot contain the antecedent for anaphoric elements lying outside".

(7) a. Max's parents weren't married when he was conceived and yours weren't married then either.
b. *Max is a bastard and yours got married afterward too.

(8) a. Max's parents are dead but my parents are alive.
b. *Max is an orphan but mine are alive.

Thus in (7b) *bastard* is taken to be an anaphoric island which contains the antecedent (*unmarried parents*) for another anaphoric element *yours*; in (8b) *orphan* is an anaphoric island that contains the antecedent (*parents*) for *mine*.

Now consider the sentences below:

(9) a. McCarthy$_i$ was glad that him$_i$ites were in the majority in the room.
b. Iroquoian$_i$ is such an interesting language family that the number of it$_i$ists has been growing rapidly.
c. When Murphy$_i$ entered the room all of the him$_i$ists began to applaud.
d. When he poked her in the leg$_i$, the long-them$_i$ed girl started to scream.

Anaphoric function words may not be embedded in morphological structure since any embedding would make them anaphoric islands; if function words in general are anaphoric, the prohibition against [FW + der aff] follows from the fact that derived words are anaphoric islands. Again, there is no need for the 2LH.

While Postal's analysis holds for function words insofar as they are anaphors, it does not rule out non-anaphoric aspects of function words. Function words are composed both of anaphoric and non-anaphoric features. *he*, for instance, has an anaphoric feature that refers to a set of previously mentioned non-participants in the discourse as well as gender and number features that limit the referent to a single male member of that set. These gender and number features are not anaphoric and can be used in forms like *he-man*,

where the relevant feature of *he* is, say, [-female]) and *she-wolf*, where the relevant feature of *she* is, say, [+female]. Thus while himite may not mean 'someone who adheres to his way of thinking', nothing prevents it from meaning, e.g., 'someone who adheres to male ways of thinking', 'someone devoted to men', etc.

Taking such non-anaphoric features into account, it is easy to construct any number of a priori possible content words derived from function words. We can imagine an affix *-ark*, for instance, that derives abstract nouns from modals:

(10)	can-ark	'potential'
	will-ark	'necessity'
	won't-ark	'impossibility'
	should-ark	'probability'

or affixes that derive nouns from auxiliaries (11), determiners (12) or pronominals (13):

(11)	be-ert	'the present'
	will-ert	'the future'
	was-ert	'the past'
(12)	the-ter	'definiteness'
	a-ter	'indefiniteness'
	some-ter	'plurality'
(13)	who-niss	'character'
	what-niss	'type'
	when-niss	'time'
	where-niss	'place'
	why-niss	'reason'
	how-niss	'manner'

or affixes that derive adjectives from deictic adverbs (14) or pronouns (15):

(14)	here-an	'close'
	there-an	'distant'
	now-an	'contemporary'
	then-an	'archaic'

(15) he-lic 'male'
she-lic 'female'
it-lic 'inanimate'

or affixes that derive verbs from conjunctions (16):

(16) and-ler 'add'
but-ler 'subtract'

In short, nothing in the syntax or semantics of function words makes them unfit for derivational affixation.

Alternative 3: the ω Hypothesis

A final possibility is that the phonology of function words, rather than their syntax or semantics, is responsible for their not hosting derivational morphology.

Inkelas (1989) proposes that affixes impose both morphological and phonological requirements on their hosts. We might then propose that derivational affixes impose a minimal prosodic requirement on their hosts such that the host consist of at least a phonological word ω (or Foot, or whatever). Since function words generally do not meet this minimal requirement (see above, Chapter 2), they will generally not be compatible with the prosodic subcategorization requirements of derivational affixes. Again, no appeal to a second lexicon is necessary, since *[FW + der aff] is ruled out on independently needed grounds.

But this cannot be right. At best it would work for monomoraic *a* and *the*. But the underlying status of a function word like *can* (*eat*) is the same as that of a content word like *can* (*of soup*)--namely, [kæn]. Any affix that fails to attach to one should fail to attach to the other.

Nor can the distinction be one of *derived* prosodic constituency. That is, one might propose that *can* (*of soup*) undergoes a rule of prosodic constituency formation that

makes it a freestanding phonological word; *can* (*eat*), on the other hand, fails to undergo this same rule (Chapter 3) and thus fails to become a freestanding phonological word; since derivational affixes only attach to freestanding phonololgical words, they cannot attach to function words.

The obvious problem here is that derivational affixes do not only attach to freestanding words. As Inkelas points out, many derivational affixes subcategorize for bound roots, i.e., for roots that may not stand on their own. Examples might include the *-fer* of *infer, transfer, confer, defer,* the *-mit* of *transmit, emit, permit,* the *-ology* of *psychology, biology, phrenology,* etc. And in a highly inflected word such as Ancient Greek or Latin, derivational affixes *never* attach to free-standing content words, since the only free-standing content words are those that are inflected. In such a language, derivation attaches to CW stems; the question remains, then, why derivation does not attach to FW stems.

Moreover, the existence of inflected function words (e.g., pronouns and articles) in languages such as Ancient Greek (ho-s 'who-nom') raises another problem for a prosodic analysis such as this. If inflectional affixes can attach to prosodically light stems, why can't derivational affixes do the same? The problem is clearly not one of adding affixes to light stems, but of adding *derivational* affixes to light stems. The best that a prosodic analysis could do is stipulate that derivational affixes subcategorize for heavy stems.

I conclude that syntactic, semantic and phonological considerations are unable to independently rule out [FW + der aff] structures. Since the 2LH rules them out straightforwardly, I will take the fact that they do not occur (or occur only very rarely) as support for the existence of two lexicons in the grammar.

Prepositions and the Double Listing Hypothesis.

Following Emonds (1985), I have assumed that there are both lexical and grammatical

nouns (*joke, he*), verbs (*buy, be*), adjectives (*small, so*) and prepositions (*under, of*). It follows that only some of these (*joke, buy, small, under*) should undergo derivational affixation while others (*he, be, so, of*) should not. This is a distinctly different claim than the one which states that only N, A, V undergo affixation. The 2LH makes a distinction between content words and function words, *not* between {N, A, V} and {P, det, Neg, Aux, Conj...}.

Particularly interesting is the split between lexical and grammatical prepositions. Pretheoretically, prepositons seem to fall between the cracks with respect to the distinction between *lexical* and *grammatical*. On the lexical end are prepositions like *concerning* and *regarding* (which contain verbal stems); on the grammatical end are prepostions such as *to* and *of* (which are hard to picture); straddling the fence, as it were, are prepositions like *over* and *under* (which are easy to picture, but hard to paraphrase). Word-formation processes provide evidence that prepositions split into these three categories: those that are lexical, those that are grammatical and those that are both.

Let us assume, then, that some words may appear both in the Lexicon (as content words) and in the Phrasicon (as function words). I will call this the Double Listing Hypothesis (DLH) and define it as follows:

(17) Double-Listing Hypothesis:
Some words are listed both in the Lexicon and Phrasicon

The Double-Listing Hypothesis (DLH) clearly weakens the 2 Lexicon Hypothesis in terms of the predictions it makes. The affixation of some prepositions in English, however, requires it. Fortunately, a number of independent considerations including the split behavior of prepositions in compounding (below), in sentence-level stress (START), and in the speech of Agrammatics (Chapter 6) point to the same conclusion: some prepositons behave both like function words and like content words. The DLH is meant to model this.

The 2LH and the DLH, then, are meant to account for a certain skewing in the affixation of words in English. Whereas content words receive derivational affixes very commonly, only a small handful of function words receive them. Function words are postulated on the 2LH to be in a component of the grammar (the Phrasicon) that derivational affixes have no access to. The few function words that do get such affixes are postulated to be doubly-listed both in the Phrasicon (with other function words) and in the Lexicon (with content words and derivational affixes). The 2LH models the general case, that function words do not undergo derivational affixation. The Double Listing Hypothesis models the exceptional case, that a certain number of function words do undergo derivational affixation.

I turn now to the affixal evidence for Double Listing. This falls into two class: affixed prepositions (*inner, outing*) and prepositional affixes (*under-achiever, outdo*).

Affixed prepositions

If prepositions are stored and processed only in the Phrasicon, they should not have access to derivational affixes such as *-ity, -able* and *-ness* in the Lexicon. But this is not the case. Derivational affixes attach to a number of prepositions to yield both nouns and adjectives:

(18) Affixed prepositions: English

after	+	most	>	aftermost	'located nearest the end'	A
down	+	er	>	downer	'barbituate'	N
	+	ward	>	downward	'descending'	A
in	+	er	>	inner	'located further inside'	A
	+	ing	>	inning	'period of a baseball game'	N
	+	most	>	inmost	'located furthest inside'	A
	+	ward	>	inward	'located inside'	A
	+	y	>	inny	'type of navel'	N
off	+	ing	>	offing	'distant part of the sea'	N
	+	ish	>	offish	'aloof'	A

on	+	ward	>	onward	'moving foward'	A
out	+	age	>	outage	'temporary loss of electricity'	N
	+	er	>	outer	'located further outside'	A
		ing	>	outing	'excursion'	N
		most	>	outmost	'located furthest outside'	A
		ward	>	outward	'located or moving outside'	A
		y	>	outy	'type of navel'	N
over	+	age	>	overage	'surplus'	N
under	+	ling	>	underling	'inferior'	N
up	+	er	>	upper	'amphetamine'	N
	+	ish	>	uppish	'arrogant'	A
	+	ity	>	uppity	'arrogant'	A
	+	most	>	upmost	'located nearest the top'	A
	+	ward	>	upward	'directed to a higher place'	A

The data in (18), then, provide a modicum of evidence that at least *after, down, in, off, on, out, over, under* and *up* are listed both in the Lexicon (where they take derivational affixes) and in the Phrasicon (where they may undergo late lexical insertion).

Affixed prepositions also occur in Ancient Greek though perhaps somewhat less frequently than in English:

(19) Affixed prepositions: Ancient Greek

anti-ázdoo	against:V:1 s	'meet face to face'	V
anti-áoo	against:V:1 s	'go to meet'	V
antí-os	against:N:m sg	'distant'	A
en-tea	in:N:n pl	'entrails'	N
epi-ssai	upon:N:f pl	'early-born'	N
kat-ooteros	down:COMP:m sg	'lower'	A
kat-ootatos	down:SUPER:m sg	'lowest'	A
meta-ssai	after:N:f pl	'late-born'	N
pros-thios	towards:ADV:m sg	'foremost'	A

(*katoóteros* 'lower' and *katoótatos* 'lowest' are formed from comparative and superlative stems (respectively) *as if* from an adjective *kata-* meaning 'low'; *kata* does not occur as an adjective on its own. *prosthíos* is built from an adverbial stem *pros-thio-* 'towards'.) Latin has affixed prepositions as well, mostly abstract nominals and adjectives.

(20) Affixed prepositions: Latin

ant-ae	against:f p	'pillars'	N
citer-ia	this side:N:f sg	'caricature'	N
contraa-rietas	against:N:f sg	'opposition'	N
contraa-rius	against:ADJ:m sg	'contrary'	A
ex-ter	out-DEIC	'outside, strange'	A
extraa-neus	outside:ADJ:m sg	'extraneous'	A
extraa-rius	outside:ADJ:m sg	'external'	A
inter-ior	inside:COMP:	'interior'	A
inter-iores	inside:COMP:m pl	'those who live further inland'	N
post-eri	after:N:m pl	'coming generations'	N
post-eritas	after:N:f sg	'posterity'	N
prop-inquaare	near:V:inf	'hasten'	V
prop-inquitas	near:A:f sg	'nearness'	N
prop-inquus	near:A:m sg	'near'	A
subter-nus	beneath:A:m sg	'beneath'	A
super-bus	over:A:m sg	'arrogant'	A
super-nitas	over:N:f sg	'height'	N
super-nus	over:ADJ:m sg	'upper'	A
ulter-ior	beyond:COMP	'on the further side'	A
ult-imus	beyond:SUPER	'extreme'	A

Prepositional affixes

Prepositons may also serve as prefixes in these languages. This serves as additional evidence that prepositions are stored in the Lexicon, where they may be prefixed to content words.

(21) Prepositional affixes: English

after	afternóon	N
down	downcóurt	A
in	inwróught	A
off	offlímits	A
on	onbóard	A
out	outlást	V
over	overcónfident	A
	overéat	V

under	undergó	V
	underdóne	A
up	updáte	V
	upstánding	A

Prefixal prepostions may be distinguished by the stresslessness of the preposition: *updáte* (prefix) vs. *úpchuck* (compound).

Such a diagnostic is not available in Ancient Greek and Latin, but prefixal prepositions can often be distinguished by their non-locative semantics. Consider the following, where the contribution of the prefixed preposition is indicated by the italicized part of the gloss.

(22) Prepositional affixes: Ancient Greek

aná	'up'	ana-pnéin	'breathe *again*'
apó	'from'	ap-anaaliskein	'*utterly* consume'
diá	'through'	dia-phthefrein	'destroy *completely*'
en	'in'	en-galáan	'laugh *at*'
ek	'out'	ek-didáskein	'teach *thoroughly*'
epí	'on, by'	epi-krúptein	'hide well'
katá	'down'	kat-esthíein	'eat up'
pará	'by, near'	par-akoúein	'*mis*-understand'

Thus *aná* 'up' adds a repetitive meaning when prefixed to *pnéin* 'breathe', *apó* adds a completive meaning to *anaalískein* 'consume'. The same applies in Latin:

(23) Prepositional affixes: Latin

ab	'from'	ab-sorbere	'soak *up*'
ad	'toward'	ad-amaare	'love *deeply*'
ante	'before'	ante-ponere	'prefer' (lit: 'put before')
ex	'out'	ex-lex	'law*less*'
inter	'between'	inter-cludere	'shut *out*'
ob	'from'	ob-audire	'obey' (lit: 'listen from')
prae	'in front of'	prae-altus	'*very* high'
pro	'in front of'	pro-tegere	'protect'
sub	'under'	sub-acer	'*somewhat* sharp'
super	'over'	super-figere	'fasten *up*'

Affixed prepositions and prepositional affixes both indicate that at least some prepositions are susceptible to derivational affixation. On the model of grammar advocated here, this is seen as evidence that these prepositions are listed in the Lexicon (where they may receive derivational affixes) as well as in the Phrasicon (where they feed late lexical insertion, etc.). These prepositions are presumably listed as bound roots in the Lexicon, or as affixes. It is of course possible that all prepositions are doubly-listed, though the evidence does not favor this.

Some prepositions seem to be purely grammatial and they serve neither as affixes nor as roots for affixation. In English *of, for, with* and *to* are likely candidates: the fact that none undergoes affixation nor serves as an affix may be taken as evidence that these prepositions are *not* doubly-listed in the lexicon. Similarly for Ancient Greek *anéu* 'without', *ákhri* 'until', *mékhri* 'as far as', *héneka* 'on account of', and Latin *apud* 'among', *cum* 'with', *sine* 'without'.

Other prepositions seem to be purely lexical. English *concerning* and *regarding* are deverbal prepositions and thus likely to be in the lexicon rather than the phrasicon. Ancient Greek has denominal *khárin* 'for the sake of' (acc sg of *kháris* 'grace') and *díkeen* 'after the manner of' (acc sg of *díkee* 'right'). And Latin has *secundum* 'next to' (*sequor* 'follow'), *versus* 'towards' (*verso* 'turn'), *tenus* 'up to, as far as' (*teneo* 'hold'), and *pone* 'behind' (*ponere* 'place, put').

Affixation provides a preliminary partition of prepostions into purely lexical (*regarding*), purely grammatical (*of*) and doubly-listed (*under*). In the next section, compounding will be shown to make essentially the same partitioning in each language.

4.1.2 Affixation in the Phrasicon

The foregoing section has sought to show that derivational morphology is restricted to stems that are listed in the Lexicon. In the general case, i.e., omitting doubly-listed

words, this means that derivation attaches only to content words. I have tried to show that this follows straightforwardly from a 2 lexicon grammar in which the first serves as input to the second: words in the Phrasicon do not serve as input to the Lexicon, predicting that derivational affixes will not attach to function words.

But notice that nothing in the 2LH predicts that function words should not take inflectional affixes: indeed, *if* function words and inflectional affixes are stored in the same component of the grammar (like content words and derivational affixes), we expect inflected function words.

English is not a highly inflected language, but it does have some inflected pronouns (see below) and prepositions. Inflected prepostions are those that take the adverbial affixes *-wards* (*homewards, northwards*) and *-ly* (*quickly*).

(24) Inflected prepositons

after	+	wards	>	afterwards	'subsequently'
down	+	wards	>	downwards	'descending-ly'
in	+	wards	>	inwards	'toward the inside'
on	+	wards	>	onwards	'forward'
out	+	wards	>	outwards	'externally'
up	+	wards	>	upwards	'toward a higher place'
over	+	ly	>	overly	'to an excessive degree'

(Recall from Chapters 1 and 2 that adverbial affixes are treated as inflectional.) More highly inflected languages like Ancient Greek and Latin have a number of inflected function words.

Inflected pronouns

Consider the following pronouns from Ancient Greek and Latin.

(25) Inflected pronouns: Ancient Greek

'who, what' (interrogative)

		sg	*pl*
masc/fem	nom	tí-s	tín-es
	gen	tín-os	tín-oon
	dat	tín-i	tí-si
	acc	tín-a	tín-es
neuter	nom	tí	tín-a
	gen	tín-os	tín-oon
	dat	tín-i	tí-si
	acc	tí	tín-a

'who, what' (relative pronoun)

		sg	*pl*
masc:	nom	hó-s	ho-í
	gen	hó-u	hó-on
	dat	hó-oi	hó-is
	acc	hó-n	ho-ns[1]
fem:	nom	heé	ha-í
	gen	hée-s	h-óon
	dat	hée-i	há-is
	acc	heé-n	ha-ás
neut:	nom	hó	há
	gen	hó-u	hó-on
	dat	hó-oi	hó-is
	acc	hó	há

(26) Inflected pronouns: Latin

'who, what' (relative pronoun)

masc	nom	qui-i	qui-i
	gen	cui-us	quo-orum
	dat	cui	qui-bus
	acc	que-m	quo-s
	abl	quo-o	qui-bus
fem	nom	qua-e	qua-e
	gen	cui-us	qua-arum

[1] The -n- in underlying *ho-ns* is lost with compensatory lengthening to yield surface [hóus]; deletion of the nasal occurs in other forms as well (tís < tin-s, etc.).

	dat	cui	qui-bus
	acc	qua-m	qua-s
	abl	qua-a	qui-bus
neut	nom	quo-d	qua-e
	gen	cui-us	quo-orum
	dat	cui	qui-bus
	acc	qui-d	quo-s
	abl	quo-o	qui-bus

Inflected pronouns in English are less common than in Ancient Greek or Latin. Many case-forms of pronouns in English are suppletive (I, me, we, us). Inflectional affixes can still be seen, however, in *who* vs. *whose* vs. *whom* where the [z] and [m] mark possessor and object case, respectively; and perhaps in *you* vs. *your*.

Inflected determiners

Inflected articles and demonstrative adjectives have been lost in English but are familiar from languages like German, Spanish and French: e.g., German 'a' *ein* (m), *eine* (f), *ein* (n); 'this' *dieser* (m), *diese* (f), *dieses* (n). Ancient Greek and Latin inflected determiners include the following.

(27) Inflected articles: Ancient Greek

'the' (definite article)

		sg	*pl*
masc:	nom	ho	ho-î
	gen	tó-u	tó-on
	dat	tó-oi	tó-is
	acc	tó-n	to-ús
fem:	nom	hee	haí
	gen	tée-s	tó-on
	dat	tée-i	tá-is
	acc	teé-n	taá-s
neut:	nom	tó	tá
	gen	tó-u	tó-on
	dat	tó-oi	tó-is
	acc	tó	tá

(28) Inflected demonstratives: Latin

'that' (demonstrative adjective)

		sg	*pl*
masc:	nom	ille	illi-i
	gen	illii-us	illo-orum
	dat	ili-i	ili-is
	acc	illu-m	illo-os
	abl	illo-o	illi-is
fem:	nom	illa	illa-e
	gen	illii-us	illa-arum
	dat	ili-i	ili-is
	acc	illa-m	illa-as
	abl	illa-a	illi-is
neut:	nom	illu-d	illa
	gen	illii-us	illo-orum
	dat	ili-i	ilii-is
	acc	illu-d	illa
	abl	illo-o	illi-is

Inflected grammatical adjectives

Possessive pronominal adjectives in these languages take the same agreement affixes as other inflected adjectives.

(29) Inflected pronominal adjectives: Ancient Greek

'my' (pronominal adjective)

		sg	*pl*
masc:	nom	emó-s	emo-í
	gen	emó-u	emó-on
	dat	emó-o	emó-is
	acc	emó-n	emo-ús
fem:	nom	emeé	ema-í
	gen	emée-s	emó-on
	dat	emée	emá-is
	acc	emeé-n	ema-ás
neut:	nom	emó-n	emá
	gen	emó-u	emó-on

	dat	emó-o	emó-is
	acc	emó-n	emá

(30) Inflected pronominal adjectives: Latin

'my' (pronominal adjective)

		sg	*pl*
masc:	nom	meu-s	mei-i
	gen	mei-i	meo-orum
	dat	meo-o	me-iis
	acc	meu-m	meo-os
	abl	meo-o	m3-iis
fem:	nom	mea	mea-e
	gen	mea-e	mea-arum
	dat	mea-e	me-iis
	acc	mea-m	mea-as
	abl	mea-a	me-iis
neut:	nom	meu-m	mea
	gen	mei-i	meo-orum
	dat	meo-o	me-iis
	acc	meu-m	mea
	abl	meo-o	me-iis

Despite the limited number of inflected function words in English, more highly inflected languages like Latin and Ancient Greek are full of them. Thus the prediction made by the 2LH, that function words and inflectional affixes should combine, is borne out. The (near-)lack of inflected function words in English only reflects the impoverished inflection system in that language.

Summary

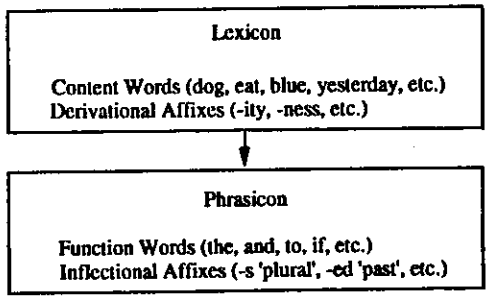
There are two kinds of words (content, function) and two kinds of affixes (derivation, inflection) in a language like English or Latin. We thus expect four kinds of affixed words, as follows:

(31) Word-Affix combinations

	Derivation	*Inflection*
Content Word	CW + der	CW + infl
Function Word	----------	FW + infl

In a number of languages, however, we find the bottom-left cell unfilled. Since this is a robust phenomenon, we would like something in the grammar to account for it naturally. The 2 Lexicon Hypothesis allows us to account for the empty cell in the following way. Content words and derivation are stored in the Lexicon, function words and inflection in the Phrasicon; the Lexicon feeds the Phrasicon.

(32) The 2 Lexicon Hypothesis



Content words and derivational affixes meet in the Lexicon (CW + der); function words and inflectional affixes meet in the Phrasicon (FW + infl); content words (more precisely, stems) and inflectional affixes meet when the output of the Lexicon is input to the Phrasicon (CW + infl). Since the Phrasicon does not feed the Lexicon, function words and derivational affixes never meet (*FW + der).

4.2 Compounding

In addition to affixation, most languages also form new words by compounding, adding a word to a word to form a new word. Compounding differs from affixation in that both items that undergo the process are words. This section will look into compounding as it

occurs in English, Ancient Greek and Latin. I will try to show that restrictions on compounding further support the separation of the vocabulary into 2 lexicons.

The evidence is of two types: the first involves the *input* to compounding, i.e., which words can be compounded with which words. The second involves the *output* of compounding, i.e., the properties of compounded words.

Input restrictions on compounding involve a general (though not absolute) prohibition on compounding CWs with FWs: compounds like **dog-the, *if-red, *of-eat,* etc. are not encountered in any of these languages. This is predicted by the 2LH if compounding only applies to words that occur within a single lexicon: e.g., *dog* and *the* are in different lexicons and therefore cannot undergo the same word formation process. In its most general formulation, the 2LH allows FW/FW and CW/CW compounds only; both of these will be shown to be quite common. More importantly, it will be shown that FW/FW compounds have a number of different features than CW/CW compounds: this will be taken as evidence that compounding-in-the-Lexicon and compounding-in-the-Phrasicon are different word-formation processes and thus as evidence for the 2LH.

The double-listing hypothesis (DLH) introduces an interesting twist: words that are doubly-listed in the Lexicon and Phrasicon (e.g., *after, down, in, off, on, out, over, up)* should be able to undergo compounding *both* with CWs that appear only in the Lexicon *and* with FWs that appear only in the Phrasicon; this prediction is generally borne out. Function words that were hypothesized to be doubly-listed because they took derivational affixes are shown to undergo compounding both with content words (*in-depth, in-lay, self-abuse, self-aware*) and with function words (*within, into, herself, myself*).

Output restrictions on compounds differ greatly depending on whether the compound in question has a content word in it or not. In English, for instance, FW/FW compounds generally have a [0 1] stress contour (*within, himsélf*), whereas CW/CW compounds generally have a [1 2] stress contour (*bâby-sit, móvie-ticket*). Similarly, compounding

with CWs is both recursive and productive, but compounding with only FWs is neither recursive nor productive. These and other considerations lead to clear distinction between compounding-in-the-Lexicon and compounding-in-the-Phrasicon, a distinction that adds considerable weight to the 2LH.

4.2.1 Compounding in the Lexicon

In English, Ancient Greek and Latin, content words enter into compounds only (a) with other content words and (b) with a small set of pronouns and prepositions. (a) follows directly from the 2LH if we assume that compounding occurs only within a single lexicon: if content words and function words are stored and processed in separate lexicons, a compounding rule which 'sees' one will necessarily not 'see' the other. (b) is compatible with the 2LH given the Double Listing Hypothesis: only those prepositions and pronouns that are listed both in the Lexicon and in the Phrasicon may undergo compounding with content words--FWs that are not doubly listed will not enter into compounds with CWs.

Input restrictions on Lexical compounding

It has often been claimed that both words in an English compound must be N, A, V or P. Thus, Selkirk states that

> Compounds in English are a type of word structure made up of two constituents, each belonging to one of the categories Noun, Adjective, Verb, or Preposition. The compound itself may belong to the category Noun, Verb, or Adjective. (1972:13)

But such a claim is too strong: it rules out compounds consisting of [adverb + CW] and [pronoun + CW] of the type shown below.

(33) [adverb + CW] compounds: English

well	well-bred, well-endowed, well-meaning, well-wisher, etc.
ill	ill-advised, ill-bred, ill-conceived, etc.
ever	ever-green, ever-lasting, everpresent, etc.
far	far-fetched, far-reaching, far-seeing, far-sighted, etc.
hard	hard-working, hard-hitting, hard-wired, hard-earned, etc.

(34) [pronoun + CW] compounds: English

all	all-American, all-consuming, all-seeing, all-important, all-seer, etc.
self	self-determination, self-destructive, self-love, self-made, self-seeker, etc.

Similar facts obtain for Ancient Greek and Latin:

(35) [adverb + CW] compounds: Ancient Greek

eu 'well'	eu-tukheés 'happy'	(lit: 'well-chanced')
dys- 'ill'	dus-tukheés 'unhappy'	(lit: 'ill-chanced')
aei 'ever'	aei-khrónios 'everlasting'	(lit: 'ever-timed')
pálai 'long ago'	palai-geneés 'born long ago'	

(36) [pronoun + CW] compounds: Ancient Greek

pan 'all'	pan-atheénaia 'all-Athenian'
	pán-agros 'all-catching'
auto 'self'	auto-któnos 'self-slaying'
	autó-noos 'self-willed'

(37) [adverb + CW] compounds: Latin

bene 'well'	bene-dictio 'praise'	(lit: 'well-speech')
male 'ill'	male-factor 'evil-doer'	
diu 'long'	diu-tinus 'long-lasting'	

(38) [pronoun + CW] compounds: Latin

omnis 'all'	omnipotens 'all-powerful'
	omni-parens 'all-producing'
se 'self'	se-cubo 'lie alone'
	se-cludo 'shuf off, seclude'

The constituents of a compound, then, may be one of {N, A, V, Adv, Prepostion, Pronoun}, at least for certain adverbs, prepostions and pronouns. Except for the pronouns, this is exactly the set of lexical items now contained within the Lexicon

according to the 2LH, augmented with the Double-Listing Hypothesis. If we assume that *all* and *self* are also doubly listed in these languages, we may state the restriction on compounding as follows:

(39) The input to compounding-in-the-Lexicon consists only of the words in the Lexicon.

(Some support for the double-listing of *self* comes from the fact that it may take derivational affixes, at least in English (selfish, selfless); *all* has no such additional evidence for double-listing.)

Note that (39) does not predict that *all* prepositons and pronouns may be compounded with CWs, but only those that are in the Lexicon. Of the 40 or so prepositions in English, half do not occur in compounds at all; another group occur in only one or two compounds:

(40) Prepositions that do not undergo compounding

across	beneath	for	till
against	beyond	from	to
around	but	of	towards
at	during	round	until
below	except	since	with

(41) Prepositions that non-productively undergo compounding

above	above-board, above-ground
about	about-face
along	alongshore
among	alongside
before	beforehand
behind	behindhand
between	between-times
like	like-minded
through	through street, through-way

Whether the prepositions above are listed in the Lexicon as well as the Phrasicon is an

open question; but it is significant that *none* of these prepositions occur with affixes.

Now consider the prepositions that do actively undergo compounding with CWs:

(42) Prepositions that actively undergo compounding with CWs

after	afterbirth, after-burner, aftershock, afterglow, after-hours, etc.
by	by-election, bygone, bylaw, byline, bypass, by-product, etc.
down	down-town, down-play, down-beat, down-stairs, down-grade, etc.
in	inbreed, income, inweave, input, inboard, incoming, inbred etc.
off	offspring, offset, offshoot, off-color, offhand, offshore, etc.
on	on-looker, on-line, oncoming, ongoing, on-stage, onshore, etc.
out	outhouse, outbreak, outrun, outset, outspoken, outright (adv), etc.
over	overabundance, overtime, overachieve, overactive, overlay, etc.
under	undercoat, underachieve, understaffed, under way, underframe, etc.
up	uproot, upstroke, upstage, upstream , upkeep, upbeat, upshot, etc.

This list is almost identical to the list of prepositions that take affixes (only *by* and *under* do not occur with affixes). Thus the double-listing of a large number of prepositions is supported both by affixation and by compounding.

Ancient Greek and Latin show a similar pattern. The prepositions that actively compound with content words (i) are essentially the same as those that are prefixed to content words and (ii) include those that take derivational affixes. Partial lists of these may be found above in (22) and (23). I have assumed here that a distinction may be drawn between prefixal prepositions and prepositonal compounds in these languages, as in English (*underpáid, óverkill*). The only diagnostic for such a distinction is the semantics of the preposition: prepositions retain their normal semantics in compounds, but take on different semantics as prefixes. Thus the claim in (i) is somewhat tenuous.

To review, the 2LH (with double-listing) predicts that the following types of compounds should be *possible* in the English Lexicon:

(43) Possible types of compounds predicted by 2LH/DLH

(+)	N	V	A	Adv	Pro	Prep
N	+	+	+	+	+	+
V	+	+	+	+	+	+
Adj	+	+	+	+	+	+
Adv	+	+	+	+	+	+
Pro	+	+	+	+	+	+
Prep	+	+	+	+	+	+

(Recall that 'Prep' here stands only for those prepositions postulated to be in the Lexicon; likewise 'Pro' stands only for *all* and *self*.) Only a subset of these are *actual* compounds in English:

(44) *Attested English compounds*

	Nouns	*Verbs*	*Adjectives*	*Adverbs*	*Pronouns*	*Prepositions*
N	book-store	sky-dive	sky-blue	ø	ø	runner-up
V	crybaby	ø	ø	ø	catch-all	drive-in
Adj	blackøberry	ø	grey-green	ø	ø	black-out
Adv	welløwisher	ready-made	well-founded	ø	ø	right-on
Pro	alløstar	self-actualize	all-merciful	alright	ø	ø
Prep	undergshirt	over-feed	in-grown	downright	ø	ø

Similarly for Ancient Greek and Latin: both languages fill out most, but not all of the predicted possibilites. The 2LH with double-listing thus allows slightly more than is actually attested in these languages.

But it is important to see as well what the 2LH with double-listing *does not* allow; for here its superiority to a single-Lexicon grammar is more evident. The 2LH and DLH, as given so far, are not compatible with *any* of the following types of compounds.

(45) [CW + FW] compounds

Noun-based	Verb-based	Adj-based	Adv-based
[N + Aux]	[V + Aux]	[A + Modal]	[Adv + Modal]
[N + Modal]	[V + Modal]	[A + Aux]	[Adv + Aux]
[N + Det]	[V + Det]	[A + Det]	[Adv + Det]
[N + Conj]	[V + Conj]	[A + Conj]	[Adv + Conj]
[N + Comp]	[V+ Comp]	[A + Comp]	[Adv + Comp]

[N + Neg]	[V+ Neg]	[A + Neg]	[Adv + Neg]

(46) [FW + CW] compounds

[Aux + N]	[Aux + V]	[Aux + Adj]	[Aux + Adv]
[Modal + N]	[Modal + V]	[Modal + Adj]	[Modal + Adv]
[Det + N]	[Det + V]	[Det + Adj]	[Det + Adv]
[Conj + N]	[Conj + V]	[Conj + Adj]	[Conj + Adv]
[Comp+ N]	[Comp + V]	[Comp + Adj]	[Comp + Adv]
[Neg+ N]	[Neg + V]	[Neg + Adj]	[Neg + Adv]

That is, if word-formation (affixation and compounding) occurs only among words within a particular lexicon (the Lexicon or the Phrasicon), compounds composed of a content word and a (singly-listed) function word are predicted not to occur. And this is the case for English, Ancient Greek and Latin, none of which allows these types of compounds.

Undoubtedly, most of the hypothetical compounds in (36) and (37) would not make useful words for naming situations and objects in everyday life; perhaps they may be ruled out on purely pragmatic grounds. Perhaps their non-existence isn't evidence for two lexicons, but merely for two types of words.

Perhaps. But this assumes that the semantic contribution of function words in compounds is essentially the same as the semantic contribution they make when they occur as free-standing words. We have already seen many cases in which this is not true: e.g., the out in *black-out* and *out-run* has little to do semantically with the out in *fell out of the the window*. Why should *to* and *of* be any different than *out*?

Output restrictions on Lexical compounding

The result of lexical compounding in English, Ancient Greek and Latin is subject to a number of restrictionss. That is, when we consider compounds that have at least one content word in them, a number of generalizations emerge. These are not the same generalizations that emerge from non-Lexical compounding (compounding-in-the-

Phrasicon, see below).

Properties of Lexical compounds: English

Stress. Stress in English lexical compounds (LCs) is usually on the first word, with secondary stress on the second: *drúg-stòre, champágne-glàss.* A smaller number of compounds have primary stress on both member: *réd-hót, sélf-adáptive.* Compounds whose first member is *out, over,* or *under* generally have primary stress on the second member, with secondary stress on the first. A few compounds have no secondary stress: *cráftsman, fíreman.*

(47) Stress-types in Lexical compounds

Preferred:	[1 2]			
	bírd-wàtcher			
Other:	[1 1]	[2 1]	[1 0]	
	grás-gréen	òver-éat	políce-man	

Stripping. The first member of a LC is generally stripped of inflectional affixes such as plural or tense: bird-watcher, *birds-watchers; run-away, *runs-away. Phrases too are stripped of function words when they are compounded. As Marchand puts it,

> In a composite the word occurs as a pure nondeterminate semanteme. Elements of syntactic reference are omitted. Flexional endings insofar as they express rection do not occur, i.e., we find no genitive, nor do we use prepositions which have a similar function. Cf. 'we go *to* the theater' and *theater goer*, 'we ride *on* a boat' and *boat ride*, 'we fight *for* freedom' and *freedom fighter*, 'we rely *on* him' and *unreliable*. It is only in the type **uncalled-for** that the particle is taken to be a lexical part of the verb rather than a syntactic element of rection.
> The position of the plural as a first element of compounds is not tied up with the question of rection, and flexional endings therefore would play a different part for the plural in that respect. However, the idea of plural as a mere grammatical expression of plurality contradicts the principle of the nondeterminateness of the first compound member. In English, plurals do not usually occur in compounds or derivatives: *toothbrush, man-eater,*

man-killing, moth-eaten, two-seater, toothless.
Linguistic elements that serve to place a statement in an actual speech situation... are omitted. No determiners (articles, pronouns) occur; 'we go to the theater', but *theater goer* without the article, 'they govern (them)selves', but *self-government.*
The semanteme does not include tense, which is an element added to place the verbal idea in an actual time frame. Mincemeat is from 'the meat has been minced', draw-bridge from 'the bridge will be, is to be drawn'.
Exceptions to the preceeceding rules are observed with certain exocentric combinations as sit-by-the-fire, stay-at-home, kiss-me-quick, reach-me-down...i.e., derivations from exclamatory sentences, quotations or the like.
(1969:38)

Stripping is straightforwardly predicted by the 2LH: function words and inflectionsal affixes are omitted in LCs because they are not stored where LCs are formed. Compounding-in-the-Lexicon has access *only* to content words and doubly-listed function words (*over-arching*). Note that derivational affixes on content words are generally not stripped: *divinity school* (**divine school*), *chewing gum* (**chew gum*), *Swedish-American* (**Swede-American*). Again, this is compatible with the claim that the Lexicon is the domain of the (uninflected) stem and it is this stem that undergoes lexical compounding.

Headedness. English compounds are generally right headed in two ways (see Selkirk 1972 for a fuller discussion). First, the right-most member usually determines the part-of-speech of the compound as a whole: *race-horse* (a type of horse) vs. *horse-race* (a type of race). (Exceptional cases include Adj where we would expect N (*underwater, undersea, underweight*), N where we would expect V (*nosebleed, sunset, earthquake*), N where we might expect P (*sit-in, dug-out*)). Second, the right-most member usually determines the sense of of the compound as a whole: a *cat-nap* is a type of nap, not a type of cat; *under-ripe* is a degree of ripeness not a type of under-ness, and so on.

Recursiveness. Compounds may generally be recursively embedded within each other in

English. From *finger* and *nail* we get *finger-nail*; then *finger-nail polish, finger-nail polish remover, finger-nail polish remover bottle, finger-nail polish remover bottle cap, finger-nail polish remover bottle cap cover,* etc.

Productivity. Not all types of LC are productive in English, but those that are are generally very productive. Aside from verb-adjective, noun-verb, adjective-verb and verb-verb compounds, most LC types are highly productive (cf. Recursiveness).

Lexicality. The output of English compounding is restricted to nouns, verbs and adjectives. Unlike affixation-in-the-Lexicon, compounding-in-the-Lexicon never produces prepositions or adverbs. This may be captured with an output filter on compounding-in-the-Lexicon (cf. Aronoff 1976; Selkirk 1982; Scalise 1984):

(48) The output of English compounding is N, A or V.

The major properties of English compounds may be summed up as folows; note that these are strong tendencies rather than exceptionless generalizations.

(49) Properties of lexical compounds: English

a. Preferred Stress:	*[1 2]*
b. Stripping:	*yes*
c. Headedness:	*right*
d. Recursiveness:	*yes*
e. Productivity:	*yes*
f. Lexicality:	*N, A,* or *V*

(a) - (f) hold of Ancient Greek and Latin as well, with a few exceptions. The following LCs are typical of these languages:

(50) Lexical Compounds: Ancient Greek

nau-makh-ia	ship:battle:N f sg	'naval battle'	N
glukú-pikro-s	sweet:bitter: A m sg	'sweetly bitter'	A
sum-mákh-omai	with:fight:1 sg mid	'fight along with'	V

(51) Lexical Compounds: Latin

su-ove-taur-ilia	pig:sheep:bull:N f sg	'type of sacrifice'	N
omni-poten-s	all:powerful:A m sg	'omnipotent'	A
carni-fico	flesh:make:V 1 sg	'execute, behead'	V

As the examples suggest, compounds in these languages are built on stripped stems (b), are generally rightheaded (c), recursive (d), productive (e) and result in nouns, verbs and adjectives (f). The main difference between English LCs and its classical counterparts is stress: English has a distinctive compound stress, where Ancient Greek and Latin stressed compounds essentially the same as other orthotonic words.

(52) Properties of lexical compounds: Ancient Greek and Latin

a. Preferred Stress: *n/a*
b. Stripping: *yes*
c. Headedness: *right*
d. Recursiveness: *yes*
e. Productivity: *yes*
f. Lexicality: *N, A,* or *V*

For the most part, then, English, Ancient Greek and Latin LCs share the same properties. But these languages have *non*-lexical compounds as well; and these compounds have traits essentially opposite to those of LCs.

4.2.2 *Compounding in the Phrasicon*

Whereas Lexical compounding has been dealt with extensively in the literature, much less work has been done on compounding involving function words. In those cases where function words are compounded with other words, we may distinguish two cases: first, those predicted by the Double Listing Hypothesis, whereby certain prepositions and pronouns are compounded with content words; these cases have already been dealt with. Second, there are cases of function words compounded with other function words--this does not require double-listing, of course, since the 2LH allows for compounding-in-the-Phrasicon. Such compounds may be called *non-lexical compounds* (NLCs) or, to look ahead a bit, *portmanteaux*..

Input and output restrictions on non-Lexical compounding

The literature on compounding has concentrated almost exclusively on compounds that *contain* at least one lexical (N, A, V) item and *produce* one lexical item (N, A, V): compounds that consist only of function words or are themselves function words are generally assumed not to exist. This has resulted in the following putative constraints on English compounding:

(53) Input Constraint: Only N, A, V, P may undergo compounding.

(54) Output Constraint: Only N, A, V may be produced by compounding.

Such constraints are implicit in the work of a number of authors:

> Compounds in English are a type of word structure made up of two constituents, each belonging to one of the categories Noun, Adjective, Verb, or Preposition. The compound may belong to the category Noun, Verb, or Adjective. (Selkirk 1972:13)

> All regular word-formation processes [including compounding--C.G.] are word-based. A new word is formed by applying a regular rule to a single already existing word. Both the new word and the existing one are members of major lexical categories. (Aronoff 1976:21)

> A more speculative possibility is that lexical rules apply to lexical categories only, that is, to such categories as Noun, Verb, Adjective, Adverb, but not to such categories as Determiner, Pronoun, Auxiliary, Complementizer, Conjunction, Interjection.... By exclukding nonlexical categories from the lexical system we account...for *their failure to enter into word-formation processes* [affixation and compounding--CG, emphasis mine] and... for their failure to undergo rules of lexical phonology. (Kiparsky 1983:4)

The statements above are not true of English. As Marchand points out:

> Compounding occurs in all word classes. There are compound substantives, verbs, pronouns, and particles (conjunctions and prepositons). The strongest group is that of substantives. Next come compound adjectives, then verbs. There is a small group of compound pronouns (the pronominal adverbs included), conjunctions and prepositions, which is naturally restricted. (1969:30)

Consider the NLCs below:

(55) Reflexive Pronouns

myself	yourself	herself	himself	itself
ourselves	yourselves	themselves		

(56) Indefinite Pronouns

whatever	whichever	whoever	whomever
whatsoever	whichsoever	whosoever	whomsoever

(57) Pronominal Adverbs

hereabout	hereto	thereagainst	thereon	whenever
hereafter	heretofore	thereat	thereto	whensoever
hereby	hereunto	thereby	theretofore	whereabouts
herein	hereupon	therefor	thereunder	wherefore
hereinafter	herewith	therefore	thereunto	wherein
hereinbefore	however	therefrom	thereupon	(whereon)
hereinto	howsoever	therein	therewith	whereto
hereof	thereabout	thereinafter	therewithal	whereunto
hereon	thereafter	thereof	whencesoever	wherewith

(58) Compound Conjunctions

however	whensoever	wherefrom	wheresoever	whereupon
whenas	whereas	wherein	wherethrough	wherewith
whencesoever	whereat	whereinto	whereto	wherewithal
whenever	whereby	whereof	whereunto	

(59) Compound Prepositions

into	throughout	within
onto	upon	without

NLCs include compounds composed of pronouns and adverbs, violating the claim that compounds consist only of N, A, V, or P. NLCs also include compounds that function as pronouns, adverbs, conjunctions and prepositions, violating the claim that compounds must be of the category N, A or V. Thus, although (53) and (54) are trivially true of LCs, they are not true of NLCs. The point here is that compounding-in-the-Lexicon and compounding-in-the-Phrasicon are different processes, subject to different input constraints.

NLCs are subject to different *output* constraints than LCs as well. These include stress, stripping, headedness, recursiveness, productivity and lexicality.

Stress. The most common compound stress for [FW + FW] compounds is [0 1]. All disyllabic reflexives, compound prepositions (except *ínto* and *ónto*) and *here-, there-* compounds are stressed this way:

(60) Main stress-pattern in non-lexical compounds: English

Reflexives	*Prepositions*
mysélf	withín
howéver	
yoursélf	withóut
hersélf	upón
himsélf	throughóut
itsélf	
oursélves	
yoursélves	
themsélves	

here-cpds	*there*-cpds	*where*-cpds	*when-cpds*	*ever*-cpds
hereáfter	thereáfter	whereát	whenás	soéver
here'by	thereát	where'by		whatéver
hereín	there'by	wherefróm		whenéver
hereóf	therefór	whereín		wheréver
hereínto	therefróm	whereínto		whoéver

herеón	therefn	whereóf	whomévеr
hereтó	thereóf	whereón	
hereúnto	thereón	whereтó	
herewíth	thereтó	whereúnto	
	thereúnder		
	therewíth		

Additional stress-patterns for English NLCs include the following.

(61) Other stress-patterns in non-lexical compounds: English

[1 0]	[1 2]	[2 1]	[1 0 2]	[2 0 1]
ínto	héreabòut	thèreuntó	héretofòre	hèreináfter
ónto	héreupòn	thèreabóut	théretofòre	hèreinbefóre
	thérefòre	thèreagáinst	whéncesoèver	hòwsoéver
	whéreabòuts	thèreupón	whénsoèver	thèreináfter
	whéreфòre	thèrewithál	whéresoèver	whàtsoéver
	whérewíth		whéreupòn	whìchsoéver
	whérethròugh		whérewithàl	whòsoéver
	whéretò			whòmsoéver
	whérewíth			

Stripping. Stripping is not found in Phrasicon-compounds. Reflexive pronouns, for instance, are inflected for number on *both* halves of the compound:

(62) Doubly-inflected Reflexives

myself	ourselves	*myselves
yourself	yourselves	(*your* is both sg and pl)
herself	themselves	*herselves
himself	themselves	*himselves
itself	themselves	*itselves

Similarly, indefinite pronouns are inflected on the left half of the compound, rather than stripped:

(63) Inflected Indefinite Pronouns

whoever	whomever	*whoeverm
whosoever	whomsoever	*whosoeverm

Headedness. English NLCs are generally left-headed: [pronoun + adverb] compounds are pronouns, [adverb + prepositions] compounds are adverbs, and [conjunction + adverb] compounds are conjunctions:

(64) Left-headed Phrasicon-compounds

what (pro)	+	ever (adv)	>	whatever (pro)
who		ever		whoever
which		ever		whichever
here (adv)	+	about (prep)	>	hereabout (adv)
there		by		thereby
when (conj)	+	ever (adv)	>	whenever (conj)
where		ever		wherever

Recursiveness. NLCs allow no more than the one level of embedding: This is no doubt related to productivity (below).

(65) Embedding in Phrasicon-compounds

[[here in] after]	[[there to] fore]
[[here in] before]	[[there in] after]
[[here to] fore]	

Productivity. The number of compounds in the Phrasicon is both small and fixed; non-lexical compounding is not productive (and therefore not recursive).

Lexicality. Not surprisingly, non-lexical compounding never results in N, V or A. The output of compounding two function words is always a function word; this is a result of the headedness of NLCs.

The properties of NLCs are summarized below:

(66) Properties of non-lexical compounds: English

a. Preferred Stress: *[0 1]*
b. Stripping: *no*
c. Headedness: *left*
d. Recursiveness: *no*
e. Productivity: *no*
f. Lexicality: *may not result in N, A, or V*

Again, some of the properties (such as [0 1] stress) are strong trends rather than exceptionless generalizations and some of the properties are interconnected (esp. recursiveness and productivity, headedness and lexicality).

NLCs in Ancient Greek share the same properties (except stress). The following are exemplary. AG reflexive and indefinite pronouns show that both halves of a NLC may be inflected (no stripping) and that NLCs are left-headed.

(67) Reflexive Pronouns: Ancient Greek

	'myself, yourself, himself'		
	1st (m)	*2nd (m)*	*3rd (m)*
Gen S	em-autóu	se-autóu	he-autóu
Dat	em-autóo	se-autóo	he-autóo
Acc	em-autón	se-autón	he-autón
Gen P	heemóon-autóon	humóon-autóon	he-autóon
Dat	heemíin autóis	humíin autóis	he-autóis
Acc	heemáas autoús	humáas autoús	he-autoús

(68) Indefinite Pronouns: Ancient Greek

	'who-ever'		
	masc	*fem*	*neut*
Nom S	hós-tis	heé-tis	hó-ti
Gen	hóu-tinos	hées-tinos	hóu-tinos
Dat	hóo-tini	héе-tini	hóo-tini
Acc	hón-tina	heén-tina	hó-ti
Nom P	hoí-tines	haí-tines	há-tina
Gen	hóon-tinoon	hóon-tinoon	hóon-tinoon
Dat	hóis-tisi	háis-tisi	hóis-tisi
Acc	hoús-tinas	haás-tinas	há-tina

(69) Compound Particles: Ancient Greek

toínun 'accordingly'	toi 'surely'	+	nun 'now'		
toigár 'so then'	toi	+	gár 'for'		
toigártoi 'therefore'	toi	+	gár	+	toi
toigaróun 'accordingly'	toi	+	gár	+	óun

(70) Compound Prepositions: Ancient Greek

amphiperí	amphi 'about'	+	perí 'around'
apék	apó 'from'	+	ek 'out'
apopró	apó 'from'	+	pró 'before'
diapró	diá 'through'	+	pró 'before'
diék	diá 'through'	+	ek 'out'
hupék	hupó 'under'	+	ek 'out'
paréks	pará 'alongside'	+	eks 'out'
peripró	perí 'around'	+	pró 'before'

Compound particles, of which those above are but a small fraction, and compound prepositions provide other cases of NLCs. None of the compounds above allows for recursion and none of them appear to be productive, though there are a great many compound particles.

Latin has similar NLCs, including those below. The demonstrative, indefinite and indefinite relative pronouns again do not strip compounded stems of their inflection. Note that the demonstrative may be analyzed either as a compound (ii + dem) or as an affixed pronoun (ii-dem): either way, the inflected forms inside of ***dem*** would be anomalous were these pronouns content words.

(71) Demonstrative Pronouns: Latin

	'the same'		
	masc	*fem*	*neut*
Nom S	ii-dem	ea-dem	i-dem
Gen	eius-dem	eius-dem	eius-dem
Dat	eii-dem	dii-dem	dii-dem
Acc	eun-dem	ean-dem	i-dem
Abl	eoo-dem	eaa-dem	eoo-dem
Nom P	ii-dem	eae-dem	ea-dem
Gen	eoorun-dem	eaarun-dem	eoorun-dem
Dat	eiis-dem	eiis-dem	eiis-dem
Acc	eoos-dem	eaas-dem	ea-dem
Abl	eiis-dem	eiis-dem	eiis-dem

(72) Indefinite Relatives: Latin

'who-ever, whatever'

	masc	*fem*	*neut*
Nom S	quis-quis	quae-que	quod-quod
Gen	(cuius-cuius)	(cuius-cuius)	(cuius-cuius)
Dat	cui-cui	cui-cui	cui-cui
Acc	quem-quem	(quam-quam)	quid-quid
Abl	quoo-quoo	(quaa-quaa)	quoo-quoo
Nom P	qui-qui	(quae-quae)	(quae-quae)
Gen	(quoorum-quoorum)	(quaarum-quaarum)	(quoorum-quoorum)
Dat	quibus-quibus	quibus-quibus	quibus-quibus
Acc	(quos-quos)	(quos-quos)	(quos-quos)
Abl	quibus-quibus	quibus-quibus	quibus-quibus

The forms in (71) and (73) show the common left-headedness of Latin NLCs:

(73) Indefinite Relatives: Latin

'who-ever, what-ever'

	masc	*fem*	*neut*
Nom Sg	quii-cum-que	quae-cum-que	quod-cum-que
Gen	cuius-cum-que	cuius-cum-que	cuius-cum-que
Dat	cui-cum-que	cui-cum-que	cui-cum-que
Acc	quem-cum-que	quam-cum-que	quid-cum-que
Abl	quoo-cum-que	quaa-cum-que	quoo-cum-que
Nom Pl	qui-cum-que	quae-cum-que	quae-cum-que
Gen	quoorum-cum-que	quaarum-cum-que	quoorum-cum-que
Dat	quibus-cum-que	quibus-cum-que	quibus-cum-que
Acc	quos-cum-que	quos-cum-que	quos-cum-que
Abl	quibus-cum-que	quibus-cum-que	quibus-cum-que

Indefinite correlatives (below), on the other hand, are right-headed, showing that left-headedness is only a strong tendency in Latin NLCs.

(74) Compound Correlatives: Latin

interrogative	*indefinite relative*	*indefinite*
quis	quis-quis	ali-quis
'who?'	'whoever'	'someone'
quantus	quantus-cum-que	ali-quantus
'how great?'	'however great'	'some'
ubi	ubi-ubi	alic-ubi
'where?'	'wherever'	'somewhere'
quandoo	quandoo-cum-que	ali-quandoo
'when?'	'whenever'	'at some time'
quotiens	quotiens-cum-que	ali-quotiens
'how often?'	'however often'	'at several times'

A comparison of the properties generally associated with NLCs in English, Ancient Greek and Latin and those associated with LCs in these languages is given below.

(75) Properties of compounds in the two lexicons

	Lexicon	Phrasicon
(Stress:	*[1 2]*	*[0 1])*
Stripping:	*yes*	*no*
Headedness:	*right*	*left*
Recursiveness:	*yes*	*no*
Productivity:	*yes*	*no*
Lexicality:	*N, A,* or *V*	*not N, A, or V*

As (75) makes clear, LCs are strikingly different than NLCs. This is to be expected, given the 2LH: compounding-in-the-Lexicon and compounding-in-the Phrasicon are different processes just as affixation-in-the-Lexicon (derivation) and affixation-in-the-Phrasicon are different processes.

4.3 Other word-formation processes

Most languages also use other word-formation processes that are less common than affixation and compounding. These fall into two groups: those that are used to form content words and those that are used to form function words. In this section I will try and show that these two sets do not overlap: as was the case with affixation and compounding, word-formation processes used to create new content words are not used to create new function words and vice versa.

4.3.1 Other word-formation processes in the Lexicon

English is a rich source for different types of word-formation. These include the formation of words through acronyms, blends, back-formations, clippings and proper names. All of these occur with content words, none with function words:

(76) Acronyms

Nouns:	CIA	'Central Intelligence Acenty'
Verbs:	F6	'delete'
Adjectives:	3D	'three dimensional'
Function words:	none	

(77) Blends

Nouns:	motel	'motor + hotel'
Verbs:	chortle	'chuckle + snort'
Adjectives:	slithy	'slimy + lithe'
Function words:	none	

(78) Back-Formations

Nouns:	?		
Verbs:	peddle 'sell'	<	peddler
Adjectives:	ept	<	inept
Function words:	none		

(79) Clipping

Nouns:	gym	<	gymnasium
Verbs:	canter	<	Canterbury
Adjectives:	comfy	<	comfortable
Function words:	none		

(80) Words from Proper Names

Nouns:	sandwich	<	Earl of Sandwich
Verbs:	pasteurize	<	Louis Pasteur
Adjectives:	Sig(Alert)[2]	<	Loyd Sigmon
Function words:	none		

2 A news bulletin indicating an unexpected blockage of traffic lanes that's expected to last at least half an hour. Used in Los Angeles radio announcements.

Again, these are word-formation processes that we may limit to the Lexicon: they derive content words from other content words. But function words are not formed by such processes, suggesting that these word-formation processes are not present in the Phrasicon.

4.3.2 Other word-formation processes in the Phrasicon

There is a class of function words, which we may refer to loosely as portmanteaux, that resembles blends. Consider the words in (51):

(81) Portmanteaux: English

won't	will + not
shan't	shall + not
don't	do + not
ain't	am/is/are + not

Note that *won't, shan't, don't* and *ain't* are not compounds, strictly speaking, because they are not derivable from *will + not, shall + not*, etc. via any phonological rules of English. If they were regular compounds (or formed with a clitic *not*) we would expect forms like [dunt] and [ænnt]. They are more similar to blends like *palimony* (*pal + alimony*) and *urinalysis* (*urine + analysis*). But blends have at least one property not shared by *won't, shan't, don't and ain't*: the former are always composed solely of segments found in the blended words: *p, a, l, i, m, o, n* and *y* all occur in either *pal* or *alimony*; *s, m, o* and *g* occur in either *smoke* or in *fog*; *b, r, u, n and ch* all occur in either *breakfast* or in *lunch*. Portmanteaux are not subject to this restriction: the vowel in *won't* [o] does not occur in either *will* or *not*; the vowel in *don't* [o] does not occur in *do* or *not*; and the diphthong in *ain't* [ey] does not occur in *am, is, are* or *not*.

Similar cases can be found in many languages. Portmanteaux consisting of a prepositon and a following article are common in French, German, and Portuguese.

(82) [preposition + article] portmanteaux

		German		
in 'in'	+	dem 'the' (m sg dat)	>	im
an 'at'	+	dem 'the' (m sg dat)	>	am
		French		
à 'to'	+	le 'the' (m sg)	>	au
	+	les 'the' (m pl)	>	aux
de 'of'	+	le 'the' (m sg)	>	du
	+	les 'the' (m pl)	>	des
		Portuguese		
em 'in, on'	+	o 'the' (m sg)	>	no
	+	a 'the' (f sg)	>	na
	+	os 'the' (m pl)	>	nos
	+	as 'the' (f pl)	>	nas
	+	um 'a' (m sg)	>	num
	+	uma 'a' (f sg)	>	numa
por 'for'	+	o 'the' (m sg)	>	pelo
	+	a 'the' (f sg)	>	pela
	+	os 'the' (m pl)	>	pelos
	+	as 'the' (f pl)	>	pelas

Other portmanteaux include [pronoun + pronoun] and [preposition + pronoun].

(83) [indirect pronoun + direct pronoun] portmanteaux

		Portuguese		
nos 'to us'	+	o 'it'	>	no-lo
	+	os 'them'	>	no-los
	+	a 'her'	>	no-la
	+	as 'them (f)'	>	no-las
vos 'to you'	+	o 'it'	>	vo-lo
	+	os 'them'	>	vo-los
	+	a 'her'	>	vo-la
	+	as 'them (f)'	>	vo-las
lhes 'to them'	+	o 'it'	>	lho
	+	os 'them'	>	lhos
	+	a 'her'	>	lha
	+	as 'them (f)'	>	lhas

(84) [preposition + pronoun] portmanteaux

		German		
mit 'with'	+	welchem 'which'	>	womit
nach 'after'	+	welchem	>	wonach
an 'by'	+	was 'what'	>	woran
auf 'on'	+	was	>	worauf
		French[3]		
de 'of'	+	la 'the' (f)	>	en
à 'to'	+	la 'the' (f)	>	y
		Portuguese		
a 'to'	+	me 'me'	>	a mim
	+	te 'you'	>	a ti
	+	lhe 'him/it'	>	a ele
com 'with'	+	me 'me'	>	comigo
	+	nos 'us'	>	conosco
	+	te 'you'	>	contigo
	+	vos 'you' (pl)	>	convosco

Each of these languages, then, has [function word + function word] portmanteaux. But none has portmanteaux with content words. The same applies for English, and may be summarized as follows:

(85) Portmanteau Types

[function word + function word]	YES
[content word + content word]	NO
[function word + content word]	NO
[content word + function word]	NO

It should not be thought, however, that any two function words may unite to form a portmanteau. Of the following 56 combinations, only 4 are attested in any of the languages discussed above.

3 For an analysis of *en* and y as pro-forms for PPs, see Kayne (1974:105-114).

(86) Unattested portmanteaux

*[Aux + Aux]	*[Modal + Aux]	*[Pro + Aux]	*[Prep + Aux]
*[Aux + Modal]	*[Modal + Modal]	*[Pro + Modal]	*[Prep + Modal]
*[Aux + Pro]	*[Modal + Pro]	[Pro + Pro]	[Prep + Pro]
*[Aux + Prep]	*[Modal + Prep]	*[Pro + Prep]	*[Prep + Prep]
*[Aux + Det]	*[Modal + Det]	*[Pro + Det]	[Prep + Det]
*[Aux + Conj]	*[Modal + Conj]	*[Pro + Conj]	*[Prep + Conj]
*[Aux + Comp]	*[Modal + Comp]	*[Pro + Comp]	*[Prep + Comp]
[Aux + Neg]	[Modal + Neg]	*[Pro + Neg]	*[Prep + Neg]

*[Det + Aux]	*[Conj + Aux]	*[Comp + Aux]
*[Det + Modal]	*[Conj + Modal]	*[Comp + Modal]
*[Det + Pro]	*[Conj + Pro]	*[Comp + Pro]
*[Det + Prep]	*[Conj + Prep]	*[Comp + Prep]
*[Det + Det]	*[Conj + Det]	*[Comp + Det]
*[Det + Conj]	*[Conj + Conj]	*[Comp + Conj]
*[Det + Comp]	*[Conj + Comp]	*[Comp + Comp]
*[Det + Neg]	*[Conj + Neg]	*[Comp + Neg]

The condition that must be met for two function words to be combined into a portmanteau is that they (typically) be string-adjacent at S-structure. This rules out the vast majority of forms in (86) but leaves those corresponding to (81)-(84).

Why should portmanteaux be limited to function words? The answer would seem to lie in the planning frames of the Garret production model (see Chapter 1). Commonly used collocations of function words that occur in adjacent positions in planning frames tend to merge over time, blurring the syntactic boundaries between formatives. Function words that do not commonly occur in adjacent positions in planning frames do not become portmanteaux. Since content words (not content words) are identified in terms of planning frames, they never develop into portmanteaux.

Summary

Portmanteaux, then, are a type of word-formation that occur in the Phrasicon but not in the Lexicon. The formation of words through acronyms, blends, back-formations, clippings and proper names, on the other hand, occurs in the Lexicon but not in the Phrasicon.

Conclusion

The resulting picture is given in (87). Word-formation processes that affect content words are distinct from those that affect function words. This may be modeled using the 2 Lexicon Hypothesis as follows:

(87) Word-Formation Processes in the 2 Lexicons

Lexicon	*Phrasicon*
Derivational Affixation	Inflectional Affixation
Lexical Compounding	Non-Lexical Compounding
Acronyms	Portmanteaux
Blends	
Back-Formations	
Clippings	
Words from Proper Names	

5. *Speech Errors*

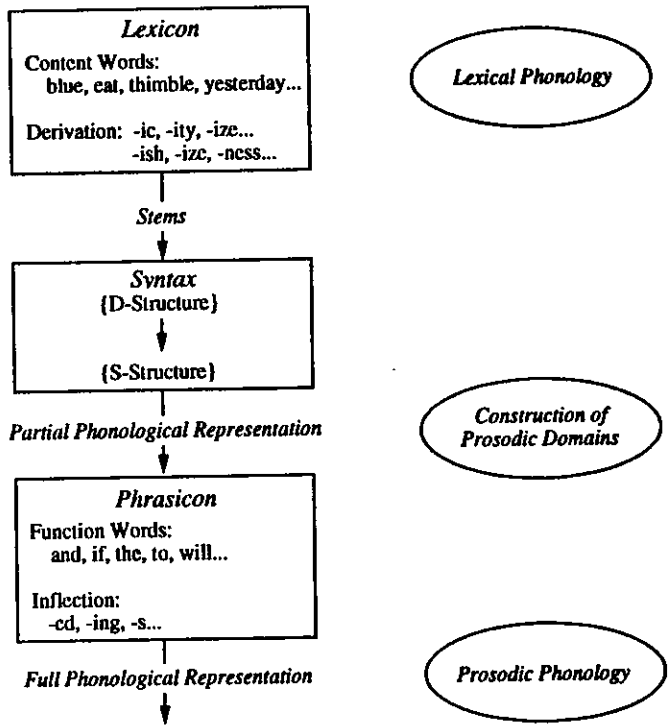
Chapters 2, 3 and 4 have presented grammatical evidence for two related claims: that the words and affixes of a language are organized into two lexicons rather than one and that the words and affixes of one lexicon are inserted into syntactic structure earlier than those of the other. Chapters 5 and 6 will provide evidence for these claims from outside the grammatical system. In this chapter I will present evidence from the study of English speech errors that supports the hypothesis that the insertion of words and affixes is level-ordered (LOLI).

LOLI and the 2LH, in addition to being purely grammatical models, may also be construed as (partial) models of language processing. This involves a rather strong assumption: viz., that the flow of information in a processing model parallels the flow of information in a formal grammar. To show that this assumption is correct would require a complete speech production processing model and grammatical model for English, neither of which are available nor the object of this dissertation. Thus, this chapter will fall short of fully supporting this assumption. What I will attempt to show is that a formal grammar incorporating the 2 Lexicon Hypothesis and Level-Ordered Lexical Insertion can contribute to a model of speech production.

Most linguists and psycholinguists would agree that the relation between a grammar and a processing model need not be isomorphic. Still, the strongest hypothesis is that the two models share the same structure. I will therefore hypothesize such a relationship and attempt to show that the kinds of speech errors which are produced and the constraints on those which occur can best be accounted for by the formal model of grammar outlined here.

If the outline of the grammar in (1) is viewed as part of a processing model (2), it predicts the kinds of speech error one will find as well as the kinds of speech error one will *not* find.

(1) 2 Lexicon Hypthesis model



(11) Function word fragments

a. Aux
|
was

b. Aux
|
had

(12) when everybody had left > when everybody was left

Additional cases of such grammatical errors are given below.

(13) Function word substitutions

I think your honor has really put your finger on it	> ...the finger on it
It looks as if	> I look as if...
how in the hell can you say that	> who in the hell can you...
the day when I was born	> ...where I was born

Note that the distinction between meaning- and sound- related substitutions is somewhat less clear in grammatical malapropisms than it is with content words: examples like *how* > *who*, for instance, might just as well be classified as meaning-related substitutions as sound-related substitutions.

5.4 Mis-insertion of lexical stems

Once (mis-)selected, the phonological forms of lexical stems are inserted into syntactic structure. Mis-insertion of content word stems results in three types of speech error: word-exchanges, sound-exchanges and stranding.

Word-Exchanges

Word-exchanges involve exchanging one word in a sentence with another; a sample is given below (from Fromkin 1973). Note that word-exchanges *usually* respect grammatical category: nouns tend to be exchanged with nouns, verbs with verbs, etc.

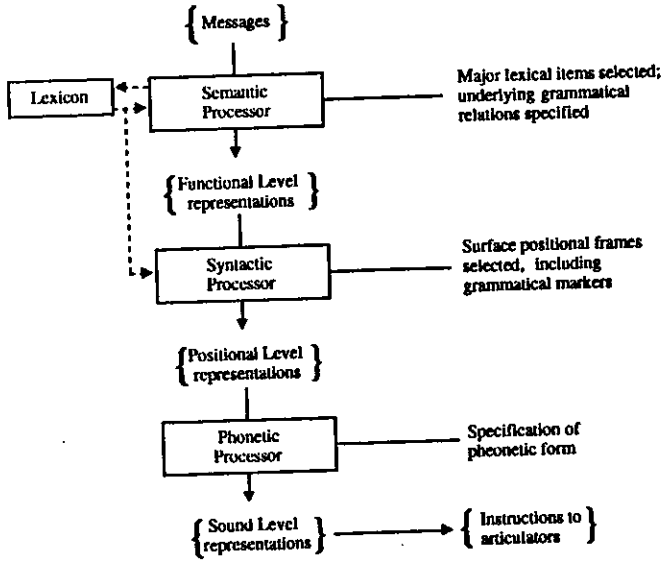
(14) Word-exchanges

the còst of the clèaning of the cárpet	> the clèaning of the còst of the cárpet
Sèymour sliced the salàmi with a knífe	> Sèymour sliced the knífe with a salámi

the...organization takes plàce in the plánt	> the...organization takes plànt in the pláce
threw the clock through the window	> threw the window through the clock

Garrett (1980:200) claims that "the elements which engage in exchanges are themselves almost exclusively major category items, or what we will call "open class vocabulary". His processing model accounts for this by positing that lexical and grammatical words are entered into speech production at different times:

(15) The Garrett Model



Word-exchanges occur prior to the function level of representation, before grammatical words are entered in; this excludes them from occurring in word-exchanges.

Garrett points out, however, that prepositions also undergo word-exchanges:

(16) Word-exchanges: Prepositions (Garrett 1980)

a. How much can I buy it *for* you *from*?
b. Write a request for tickets *at* two *for* the box office.

c. ...which was parallel, *to* a certain sense, *in* an experience of...

Garrett maintains that preposition-exchanges occur at the same processing level as other word-exchanges by claiming that they are treated as lexical categories *at this level*. At a later stage they are treated as non-lexical categories and at that stage they are involved in shifts (which otherwise affect only grammatical items).

We might add, however, that pronouns are often involved in word-exchanges as well:

(17) Word-exchanges: pronouns

he'll call to find out when I wanted him to pick me up	> ...when he wanted me to pick him up
your keys are in my purse	> my keys are in your purse
I didn't know Ravel had it in him	> ...had him in it
would you ask her to call me	> ...ask me to call her
he taught it to us	> ...us to it

This makes Garrett's analysis of prepositions less convincing: are *both* prepositions and pronouns to be construed by the processor as lexical at one level and grammatical at another?

A simpler solution is to let word-exchanges occur whenever words are inserted. Recall that the REG model advocated here (see (2) above) posits that function words and inflectional affixes also undergo lexical insertion, though at a later stage in processing than content word stems. This is one of the major ways in which the REG model differs from the Garrett model: according to the latter (see (15) above), function words and inflectional affixes come pre-inserted as parts of planning frames. This makes it difficult to model pronoun and prepositions word-exchanges.

On the REG model, however, we expect both content words and function words to undergo word-exchanges, though content word-exchanges are hypothesized to occur at a different stage than function word-exchanges. If both content and function words undergo word-exchanges, we might now ask the following question: Is there any evidence that content word-exchanges and function word-exchanges occur at *different* stages in speech production?

There is. The evidence concerns word-exchanges that do *not* respect grammatical category. Consider the following cases of cross-category word-exchanges:

(18)

a. Does Jàck smóke?	> Does smòke Jáck?
b. you got as far as typing page one	> ...as far as page typing one
c. cash a check	> check a cash

In the first case, an NP exchanges with a following V; in the second and third, a V exchanges with a following N. Despite such cross-category word-exchanges between content words, I am ware of *no* such cases involving an exchange of a preposition or a pronoun with a content word. Garret himself notes that "in the MIT corpus there are no clear examples of an exchange between a preposition and any [lexical--C.G.] element of either its own NP argument or any element of any other NP" (1980:191).

There are, however, many cases of prepositions exchanging with other non-preposition function words, namely with pronouns.

(19) Cross-category word-exchanges: function words

there's a crazy card I just sent to him	> there's a crazy card I just sent *him to*
how much do you want for this?	> how much do you want *this for?*

There seems to be an asymmetry, then, between verbs and prepositions with respect to the types of words they commonly exchange with: verbs exchange with other content words (nouns, verbs), prepositions exchange with other function words (pronouns, prepositions). This is expected if all content word-exchanges occur at one stage (usually, but not always respecting category) and all grammatical exchanges occur at another (usually, but not always respecting category).

Sentence-Level Stress: prosodic domains in sentence production

It has long been noted (Boomer & Laver 1968; Fromkin 1973; others) that sentence-stress is preserved under word-exchanges--i.e., that exchanged words do not take with them the level of stress appopriate to their target position. A phrase such as *a jòb for his wife* is realized as

a wife for his jób with the same [2....1] stress as its target (rather than hypothetical **a wife for his jòb* with [1....2] stress.

Fromkin (1973) and Garrett (1980) take this as evidence that sentence-stress is assigned *before* words are entered into a sentence: once the words are entered in, they get the degree of stress appropriate for the position they occupy. But this cannot be. To see why, consider the data below.

(20) a. Tom hoped [she]$_{NP}$ would study linguistics.
b. Tom hoped [Mary]$_{NP}$ would study linguistics.
c. Tom hoped [undergraduates]$_{NP}$ would study linguistics.
d. *Tòm hòped shè would stùdy linguístics. (where *shè* is unemphatic)
e. Tòm hòped Màry would stùdy linguístics.
f. Tòm hòped ùndergraduates would stùdy linguístics.

Before lexical insertion, the bracketed phrases in (a), (b) and (c) have identical syntactic features identifying them as NPs. If phrasal stress is assigned equally to anything labeled NP, a deviant stress pattern results in the inappropriately stressed utterance in (d) (cp. e, f). The Fromkin and Garrett models mistakenly predict that pronouns (*she*), proper names (*Mary*) and one-word noun phrases (*undergraduates*) will receive the same degree of sentence-stress. But phrasal stress is 'invisible' to pronouns and other function words. Similar arguments could be made for the stresslessness of non-lexical verbs (*have, be,* etc.) as opposed to lexical verbs (*hit, run, eat*).

The REG model, by incorporating LOLI, avoids these problems in a manner now familiar from Chapter 3. Derivations for *a jòb for his wife* and *a wife for his jób* are given below. LOLI correctly models the fact that stress is the same in both cases.

(21) Retention of sentence-stress under word-exchanges

LexIns	α	job	γ	δ	wife	α	wife	γ	δ	job
Pros. Constituents	α	[job]	γ	δ	[wife]	α	[wife]	γ	δ	[job]
GramIns	a	[job]	for	his	[wife]	a	[wife]	for	his	[job]
Stray-adjunction	[a	jòb]	[for	his	wife]	[a	wìfe]	[for	his	jób]
Stress	[a	jòb]	[for	his	wife]	[a	wìfe]	[for	his	jób]

Sound-Exchanges

Sound-exchanges involve the exchange of segments or groups of segments between words. Sound-exchanges are strongly constrained by syllable position: onsets tend to exchange with onsets, nuclei with nuclei, rhymes with rhymes, codas with codas (Nooteboom 1969; MacKay 1969, 1970; Fromkin 1968). By far the most common type of sound-exchange involves the onset of two content words.

(22) Sound-exchanges: initial consonants

left hemisphere	>	heft lemisphere
lawfully joined together	>	jawfully loined...
Katz and Fodor	>	fats and kodor
copy of my paper	>	poppy of my caper

Vowels may also exchange, though much less commonly than consonants (Shattuck-Hufnagel 1986):

(23) Sound-exchanges: vowels (from Shattuck-Hufnagel 1986)

a sudden death	>	a sedden duth
one of those T-group people	>	one of those tu greep...
see that movie	>	sue that meevie
or corn starch	>	or karn storch

Coda consonants exchange with one another, though this, again, is much less common.

Sound-exchanges are limited almost exlusively to lexical items--grammatical words *very* rarely undergo sound-exchanges (Garrett 1980). I know of only 3 cases in the UCLA corpus:

(24) Sound-exchanges involving function words

I must leave at five sharp	>	I just meave	FW-CW
that's not for me to say	>	for see to may	FW-CW
we shouldn't	>	she wouldn't	FW-FW

Garrett accounts for the fact that function words rarely undergo sound-exchanges by proposing that sound-exchanges occur at a processing level at which function words are not represented phonologically. The same account is inherited in the REG model in (2). Note that LOLI and the 2LH model in (1) provide a *competence* model for the Garrett (or REG)

performance model. The speech error facts are thereby not only modeled but *explained*: the grammatical model that explains them is independently motivated.

An alternative analysis

An alternative explanation for the absence of function words in sound-exchanges is that sound-exchanges involve only sounds in stressed syllables. This would account for the data above straightforwardly. This is because the vast majority of content words in English speech are monosyllabic: 'sounds in stressed syllables' and 'sounds in content words' make the same predictions for all but a small fraction of the data available.

Such a view, although initially plausible, suffers from a number of shortcomings. First, experimentally induced speech errors with polysyllabic words indicate that the primary condition for sound-exchanges involving initial consonants is not lexical stress but *word-initial position* (Shattuck-Hufnagel 1985). Consider the tongue-twisters below:

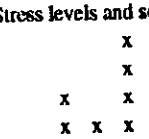
(25) a. paráde fád foót paróle
b. repeát fád foót repaír

Shattuck-Hufnagel found that the *f*s and *p*s in (a) were significantly more likely to exchange than the *f*s and *p*s in (b). If lexical stress were the primary determinant of sound-exchanges, the results would be reversed. Function words may also have onsets, of course, and yet these onsets generally fail to participate in sound-exchanges. Garrett makes the point by showing that [P N] structures rarely undergo sound-exchanges while [V N] structures commonly do (1980:191). If prepositions have not yet been inserted at the point at which exchange errors occur, they are not expected to undergo sound-exchanges with nominal complements; but quite the contrary for verbs, since they are inserted at the same time as nominals.

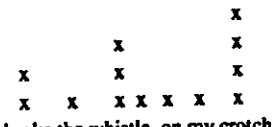
The second problem with a stress/stressless analysis is that stress is not a binary but an *n*-ary relation. The words of a sentence are not merely stressed or stressless. The notion of stress is inherently relational in a way that makes 'lexical stress' a difficult concept to apply in

modeling speech errors. Consider the data below (Fromkin 1973--stress levels indicated in original by means of integers: [3...1] and [3...2...1])[1].

(26) Stress levels and sound-exchanges



a. [parklz pijp]
(for *peoples park*)



b. broke the whistle on my crotch
(for *broke the crystal on my watch*)

Clearly, sounds may exchange from syllables with unequal ([3...1], [2...1]) stress. Indeed, since one of these words involved is usually the most-stressed word in the phrase (Boomer & Laver 1968) and since there is usually only one most-stressed word per phrase, sound-exchanges *almost always* involve words with different stress levels. Citing stress as the crucial factor in sound-exchanges fails to explain why *broke* and *the* are far less likely to be involved in a sound-exchange than *crystal* and *broke*: both pairs of words differ by the same degree of stress.

Stranding errors

Stranding involves leaving behind inflectional affixes when lexical stems exchange. Take the error *streeting sweeps* for *sweeping streets*. When street and sweep exchange, they strand their inflectional affixes in the intended positions: [___ing ___s]. Such errors provide concrete and surprising evidence for the claim that lexical stems are inserted prior to inflectional affixes. It is therefore important to see that all sorts of inflectional affixes are stranded in this way: plural, possessive and 3rd singular *-s*; *-ed*; *-ing*; comparative *-er*; adverbial *-ly*.

---

[1] I follow Selkirk (1986) in assuming that both metrical grids of the type shown in (26) and prosodic constituency of the type discussed in Chapter 3 are necessary in representing prosodic rule domains above the word.

(27) Stranded infectional affixes

Plural -s	
sweeping streets	> streeting sweeps
many people think he's the most underrated player in the nation	> many players think...
get your elbows off the table	> ...tables off the elbow
a floor full of holes	> a hole full of floors
she has pain in her legs	> ...leg in her pains
there are more ministers in the church	> ...churches in the minister
bunnies don't eat steak	> steaks don't eat bunny
a weekend for maniacs	> a maniac for weekends
one uncle and two aunts	> one aunt and two uncles
as I was putting the books on the shelf	> ...the shelfs on the book
seven runs in one inning	> seven innings in one run
a shelf full of cookbooks	> ...full of bookcooks
rules of word-formation	> words of rule-formation
three-subject degrees	> three-degree subjects
I can give you some top-of-the-head thoughts	> ...top-of-the-thought heads
Possessive -'s	
an aunt's money	> a money's aunt
the guy's name	> the name's guy
George's lab	> lab's George
I haven't got Leon's luo	> ...luo's Leon
3rd Singular -s	
he always keeps a pack	> ...packs a keep
it goes to show	> it shows to go
it pays to wait (G)	> it waits to pay
Past -ed	
the tie dropped out of the bag	> ...the drop tied out of the bag
I have baked a cake	> ...caked a bake
my check cashed	> my cash checked
a watched pot never boils	> a potted watch never boils
we've learned to love mountains	> we've loved to learn mountains
you're getting payed for playing	> ...played for paying
I have to get my check cashed	> I have to get my cash checked
I like to have my back scratched	> ...scratch backed
I passed the threshold	> I threshed the passhold
Participial -ing	
sweeping streets	> streeting sweeps
you're getting payed for playing	> ...played for paying

I have no way of knowing	> ...way know of waying
Comparative -er	
she has a sweeter flute than he has	> fluter sweet
Adverbial -ly	
linguistically significant	> significally linguisticant
obviously vocational	> vocationally obvious
steadily worse	> worsely steady
clearly enough	> enoughly clear
you have to face it squarely (G)	> you have to square it facely

The last case is especially interesting given that the homophonous adjectival suffix *-ly* *(friendly, neighborly)* is never stranded (to my knowledge at least).

Stranded derivational affixes, on the other hand, are very rare. The only clear cases I know of are those below (UCLA corpus).

(28) Stranded derivational affixes

a language learner needs	> a language needer learns
motherhood and apple pie	> applehood and mother pie
return your call	> recall your turn

The Garrett model and its descendents take stranding errors as evidence that lexical stems are inserted into planning frames that encode grammatical morphemes. When these stems are inserted into the wrong slots in a planning frame, they surface with the wrong grammatical morphemes attached to them. If inflected stems are inserted all at once (as in Lexical Phonology) stranding errors are difficult to model at all.

A derivation of a stranding error is given below.

(29) *A wàtched pòt nèver boíls* > *A pòtted wàtch nèver boíls*

LexIns	α	pot	γ	watch	never	boil	η
Pros. Constituents	α	[pot]	γ	[watch]	[never]	[boil]	η
GramIns	a	[pot]	ed	[watch]	[never]	[boil]	s
Stray-adjunction	[a	pot	ed]	[watch]	[never]	[boil	s]
Phrasal Stress	[a	pòt	ed]	[wàtch]	[nèver]	[boíl	s]

The error occurs at Lexical Insertion when pot and watch are mis-placed in each other's slot. From there prosodic constituents are built as usual, grammatical items are inserted and stray adjoined and phrasal stress is assigned.

Note that the inflectional affixes represented by α, γ and η represent features of Garrett's planning frame and are not yet inserted when lexical stems shift. Not so for derivational affixes, which are inserted at the same time as lexical words according to LOLI. Hence the fact that derivational affixes tend to move with the rest of the stem rather than be stranded with inflection. Below, the derived word *player* (play + er) moves:

(30) many people think he's the most underrated > many players think...
player in the nation

Had the derivation been stranded as well the result would have been ***many play-s think he's the most underrated people-er in the nation.***

5.5 Mis-insertion of function words and inflectional affixes

The REG model in (2) posits a separate processing level at which function words and inflectional affixes are inserted. In this it differs from Garrett's model and from the EG model of Lapointe and Dell: according to these models, the phonological forms of function words and inflectional affixes are the terminal elements of planning frames *ab initio*.

Grammatical motivation for this difference comes from the stray-adjunction analysis of stressless and reduced function words and the 'invisibility' of function words in the assignment of prosodic constituency (Chapter 3). Motivation for this difference in the production model comes from the analysis of *shifts*.

Shifts

Shifts involve misplaced function words or inflectional affixes, often at quite a distance. They provide striking evidence that function words and inflectional affixes work as a class in speech production; since content words and derivational affixes do not undergo shifts. Shifts involving various inflectional affixes are given below. Capital letters indicate a shifted element, blanks indicate the target position: e.g., *in transducerS array__* plural *-s* has shifted from *array* to *transducer*.

(31) Shifts: Inflectional Affixes

Plural -s
transducerS array__
we have several printS out__
for the rest of my talkS I'm going to do two thing__
all the phone__ rangS
EPL__ tendS to be

Possessive -'s
Jerry__ pancake'S house
this is Ralph__ and my'S article

3rd Singular -s
if she want__ to comeS here
she make__ sureS
add__ upS to
it come__ onS at
and Rachel come__ inS
he end__ upS
when someone come__ upS to me
what lies JoeS tell__

Participial -en
I'd forgot__ aboutEN that (Garrett 1980)

Participial -ing
I'm pày__ fòrING it all together
I should be shut__ upING
that would be the same as add__ tenING (Garrett 1980)
as I keepING suggest__

Verb-Particle
people tend UP to make-__ words
it must have been an UP put-__ job

Adverbial -ly
clear__ enoughLY
easy__ enoughLY
logical__ speakingLY
what does it mean to be high__ verbalLY

Again, I have found *no* cases of shifts involving adjectival *-ly* (e.g., ***a friend__ neighborLY*** for *a friendly neighbor*). Both in stranding and in shifts, then, adjectival *-ly* and adverbial *-ly*

pattern differently; the former patterns with derivational affixes, the latter with inflectional affixes.

Derivational affixes rarely shift; the two cases I know of are given below.

(32) Shifts: Derivational Affixes (rare)

sanitary inspector	> INsanitary __spectre
I over-heard you and Mommy talking	> I __heard you and Mommy OVER-talking

Function words are the other class of items that regularly undergo shifts. Shifts are attested for a large array of function words including pronouns, prepositions, articles, adverbials and auxiliary verbs. The most striking aspect of shifts is the degree to which they violate (syntactic) grammaticality: though they are phonologically well-formed, they violate very basic principles of phrase structure: * *the girl who taught I last year,* * *you want to send me somebody else to you,* * *an on unfortunate day,* etc.

(33) Shifts: Function Words

Pronominal

a...card I just sent *him* to __
what's *it* meant to be done with __
I had it __ planned *all* out
I would like to *all* remind you __
the girl who __ taught *I* last year
you want __ to send *me* somebody else to you
there's a crazy card I just sent *him* to __
Bever refers *it* to __ only in one form
I can *her* hear __
how *he* can __ get it done in time
but when *you* will __ leave?
why *you* do __ miss teaching when you're on leave
are __ going *you* to the Renaissance Faire this year?

Preposition

he's been __ a long *around* time
how much do you want *for* this __
the way we can characterize the situation would *by* be __
I don't see *in* any reason __
I don't want to part __ this book *with* too long

she was waiting __ her husband *for*

Article

he has spent *the* most of__ time on his synthesis
an on __ unfortunate day

Adverbial

what I'm working *then* on __
what I'm *there* saying __
Billie Jean is picking her *back* way __ up
I *once* think I could __ do this
you're in a more *better* position, you're __
it's *all* almost __finished
the *really* thing I __ like
I __ hate to *really* correct exams
but my English teacher __ doesn't know what she's *really* doing

Auxiliary

I __ say that because I *don't* want to go to Mexico
I__ glad I'*m* wasn't there
the meaning cannot __ a function *be* of truth
why do *have* we __

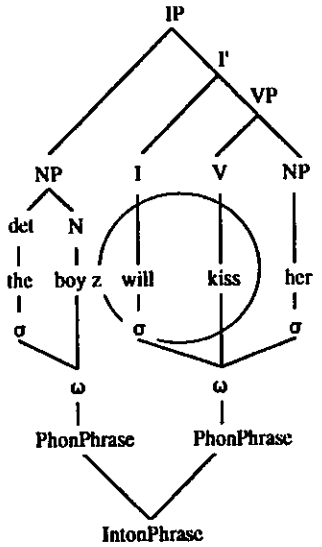
Note that multiple categorizations are possible: *a card I just sent him to* could be a result of a shifted pronoun (shifting *him* to the left), a shifted preposition (shifting *to* rightwards) or even a word-exchange (exchanging *to* and *him*). Cases such as *what's it meant to be done with* are more straightforward and allow only one categorization (shifting of *it* leftwards).

Garrett (1980) discusses shifts extenstively and uses them as additional evidence that grammatical items are inserted later than lexical items. But his model makes shifts somewhat mysterious: if they begin life as part of the 'planning frame', why don't they stay put? Garrett's model is well geared towards explaining stranding, where grammatical items remain in the correct frame position; but it fails to adequately explain shifts.

Here LOLI (as outlined in Chapter 3 above) provides a better model. According to this view, function words and inflectional affixes are stray adjoined to existing prosodic structure, whereas content words and derivational affixes have prosodic structure built upon them. Thus function words and inflectional affixes are both (1) the only morphemes that need to be stray adjoined and (2) the only morphemes that undergo shifts. With this in mind, we can

simply model shifts as (somewhat mis-guided) stray-adjunction. Grammatical morphemes 'shift' when they are stray-adjoined to the wrong words.

(34) Shifts as misguided stray-adjunction



Stray Adjunction

What allows such shifts is the *phonological* nature of stray-adjunction. Strays are adjoined to *prosodic* constituents, not to morphological or syntactic constituents. As far as the phonology is concerned, any morphological or syntactic host serves equally well: all that is required is that the host be of the right prosodic category, e.g., a ω or σ. In this sense, then, shifts result from a mis-match of prosodic and morpho-syntactic hosts (cf. Inkelas 1989): a stray function word or inflectional affix finds the *right* prosodic host (ω or σ) on the *wrong* morpho-syntactic host. The result is phonologically well-formed but syntactically and morphologically anomolous.

5.6 *Mis-specification of phonetic form*

The penultimate stage in sentence production according in Garrett-based models involves the specification of phonetic form. This applies at once to content word stems, function words and inflection since everything in the representation at this point is fully phonologized. Mistakes at this stage in the derivation should therefore apply to all words and affixes equally. A priori, we might expect three types of mistake: failing to produce some element (deletions), producing an element that isn't called for (additions) and mis-producing some element.

Each of these speech error types occur. Deletions are straightforward: e.g., *speech error > peach error*. Lexical stress is often adjusted to compensate for a lost syllable: *treméndously > trémendly*. Additions are slightly more suspect since the added element can often be accounted for by a competing plans hypothesis: e.g., *every time I want to go > every time I wantING to go* could be a result of a mis-selection of planning frames (see 9 and 10 above).

The last case, mis-production, has two main sub-cases: an element may be mis-pronounced like a following element (anticipation) or mis-pronounced like a preceding element (perseveration).

Anticipations and Perseverations

Anticipations occur when an upcoming feature, segment, cluster or syllable is realized too early in the utterance. Perseverations occur when an previeous feature, segment, cluster or syllable is realized again in an utterance.

(35) Anticipations

think through	>	thRink through
talking about	>	talking Babout
any harm	>	Hany harm
with a brush	>	wiSH a brush
be so sure	>	be SHow sure

falsify my claims > falsiMY my claims
instead of space food > inSPead of space food

(36) Perseverations

it doesn't strike me as funny > it doesn't strike me as Money
will pick up the 6th gold medal > will pick up the 6th gold Wedal
so that I can start the tape back up (G) > so that I can start the STape back up
I dreamt that he broke both arms (G) > I dreamt that he DRoke both arms
experiences become much more exportant than anything else (G) > experiences become much more EXportant than anything else
I don't understand the order at all (G) > I don't understand the order at OR

('G' from Garrett 1980.) As Garrett points out, a

> feature of the anticipation/perseveration errors that is worth mentioning is the occasional involvment of closed class vocabulary (i.e., function words and inflections) in these errors (42 cases of the 336). In this they contrast sharply with sound exchange errors which do not often involve those classes. Moreover, most of these cases are interations between an open and closed class vocabulary item, rather than interactions between two closed class items of the same minor category. What seems to be a very powerful constraint on exchange of words (correspondence of category) or exchange of sounds (confined to open class) seems of much lesser importance to the system which gives rise to errors of anticipation and perseveration. (1980:195)

Anticipations and perseverations, it should be noted, are not always transparent: some anticipations may be cases of corrected sound-exchanges:

(37) think through (target) > thrink though (sound-exchange) > thrink through (correction)

What is important about anticipations and perseverations, however, is that they seem to freely involve both content and function words, both derivation and inflection. In this way, they differ drastically from stranding errors, sound-exchanges and shifts. Just as these other speech error types argue for a level of representation at which only lexical items are phonologically realized, anticipations and preservations demand a stage at which all words and affixes are phonologically spelled out.

5.7 *Conclusion*

Different types of speech error affect lexical and grammatical morphemes in different ways. Some types of speech error affect only lexical stems (e.g., stranding errors), some affect only grammatical items (e.g., shifts), other affect both (e.g., anticipations).

The processing model in (2) predicts a number of types of speech error, namely, one or more type of error for each stage in production and no others. The predictions seem to be borne out, as folows.

(38) Predicted speech error types: REGM

1. Mis-selection of major lexical items	Substitutions, Blends
2. Mis-specification of underlying grammatical relations	Syntactic Errors
3. Mis-selection of surface positional frames	Tags, Substitutions
4. Mis-insertion of lexical stems	Stranding Errors, Word-Exchanges, Sound-Exchanges
5. Mis-insertion of function words and inflectional affixes	Shifts
6. Mis-specification of phonetic form	Deletions, Additions
	Anticipations, Perseverations

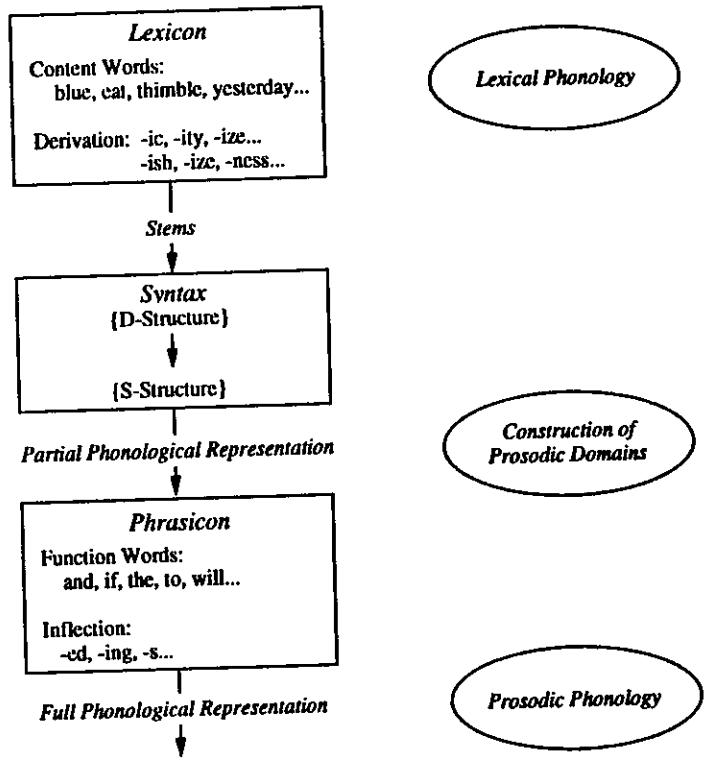
But this processing model is difficult to assess without some grammatical interpretation of the levels of representation it posits. The 2 Lexicon Hypothesis and Level-Ordered Lexical Insertion (1) provide such grammatical interpretation. In a very real sense, the grammar in (1) provides the *competence* behind the *performance* model in (2).

6. *Language Breakdown*

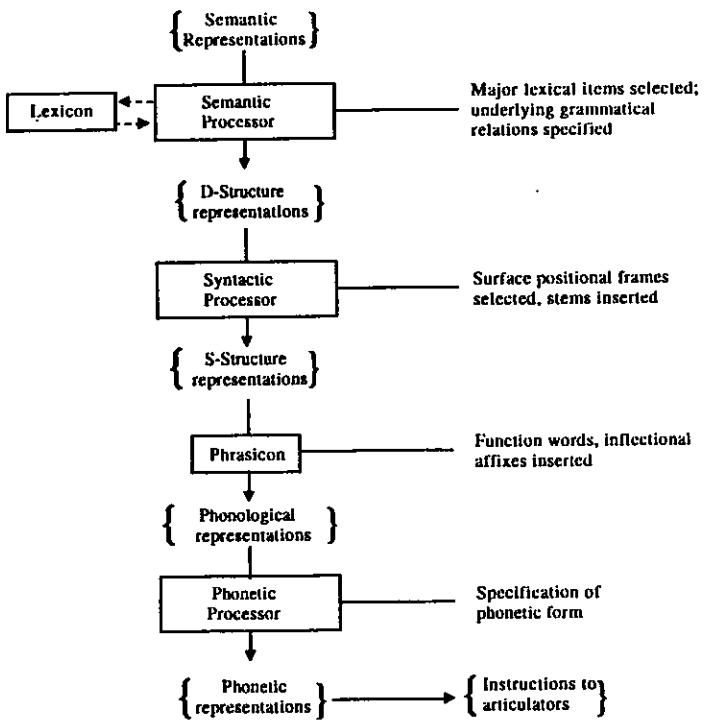
Aphasia offers evidence on how linguistic information is stored and processed in the brain. Linguistic aphasiology "deals primarily with the details of the *linguistic structures* that aphasic patients lose and retain, and with abnormalities in the processing of these structures" (Caplan 1987:143, emphasis mine). In this chapter I will review aspects of acquired aphasia that shed light on the organization and processing of words and affixes. Each of these areas has an extensive history and literature and I will not attempt to cover the range of analyses available. My main concern will be to discuss evidence for or against the models of sentence production in Chapter 1. For this reason, I will discuss aphasic production (speech) much more than aphasic comprehension, though the latter will be brought in to temper some of the claims made on the basis of the former.

As in Chapter 5, I will assume that the partial outline of the grammar in (1) may be construed as part of a processing model (2) that can be used to predict what elements of linguistic competence and performance can be differentially affected in aphasic speech.

(1) 2 Lexicon Hypthesis model



(2) Revised Extended Garrett Model



If these models are viable, we would expect each of components (boxes) on the right-hand side of the diagram in (2) to be selectively impaired in acquired aphasia. Each such impairment should result in particular aphasic symptoms and there should be no recurrent aphasic symptoms which cannot in principle be attributed to some such impairment. The grammatical and production models in (1) and (2), then, predict that the following types of selective impairment should occur.

(3) Predicted selective impairments in aphasic speech: REGM
1. Impaired (access to) Semantic Processor

2. Impaired (access to) Lexicon
3. Impaired (access to) Syntactic Processor
4. Impaired (access to) Phrasicon
5. Impaired (access to) Phonetic Processor

Types 1 and 5 fall outside the scope of this dissertation and will not be discussed. Emphasis here will be placed on types 2, 3 and 4 since these types crucially involve the hypothesis that the insertion of words and affixes is level ordered (LOLI) and the hypothesis that words and affixes are stored in two lexicons (2LH).

Note that whereas in discussing speech errors, the *processes* on the right of (2) were used to predict types of error, in discussing aphasic symptoms the *modules* on the left of (2) have been used. The boxes in (2) represent the component stores of information used in processing in both normal and aphasic speech. In modeling speech errors, these components must be intact and access to these components must be intact; speech errors in normals do not result from damaged components of the grammar or from restricted access to them but from mis-processing. Speech errors must be modeled on an intact and normal grammar/processor. In aphasia, on the other hand, either access to a component is disrupted or the component itself is disrupted; aphasia is taken to result from damaged components of the grammar or from restricted access to them. Aphasia is to be modeled with a *disrupted* grammar/processor.

Symptoms and syndromes

In discussing the linguistic analysis of aphasia, it is crucial to clearly distinguish *symptoms* from *syndromes*. The list in (3) is meant to broadly characterized types of symptom rather than types of syndrome. Syndromes are collections of symptoms that tend to recur in patient after patient. Wernicke's aphasia and Broca's aphasia are classical syndromes; each is characterized in terms of a number of symptoms: Wernicke's patients tend to have fluent, well-articulated and (apparently) syntactically complex speech marked by non-

sensical content and word-finding difficulties; Broca's patients, on the other hand, tend to have non-fluent and dysarthric speech that is syntactically impoverished but meaningful, often marked by a lack or misuse of function words and inflectional affixes. Syndromes are also often associated with different areas in the brain: Broca's aphasia is usually associated with insult to the anterior part of the temporal lobe of the left hemisphere, whereas Wernicke's aphasia is more associated with the posterior part of the temporal lobe.

Symptoms, on the other hand, may recur in any number of syndromes. They serve both to define syndromes and to sub-classify them. Symptoms are unitary phenomena that may be related to particular aspects of linguistic structure. Thus although Broca's aphasia as a syndrome may be difficult to characterize linguistically, the symptoms that characterize it are not: lack of prosody, problems with articulation, lack of syntactic structure, lack or misuse of grammatical morphemes, etc.

This chapter will necessarily focus, therefore, on symptoms that recur in aphasia, rather than on syndromes of those symptoms. In modeling these symptoms in terms of the models in (1) and (2) I will set aside the important issue of whether a symptom results (a) directly from a disrupted component of the grammar/processor or (b) indirectly from disrupted *access* to a (sub-)component of the grammar/processor. I will use four cross-hatches (////) over a component of the grammar/processor to represent either (a) or (b).

6.1 Impairment implicating the Semantic Processor

Impaired access to the Semantic processor in (2), with the rest of the processor intact would yield speech that is syntactically, morphologically, phonologically and phonetically well formed but *meaningless*. Such "paraphasic" speech is found with many patients diagnosed as Wernicke's aphasics. Consider the following samples:

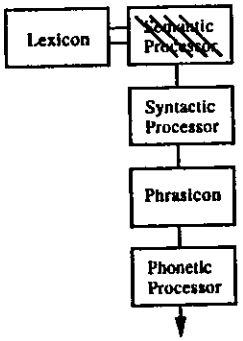
> But I figured that if I defective my my talking see my talking itself I I get my tongue back again to where I can talk from what they say why then its

> lliable to that will straighten me out again and bring me back to where I can hear something see and until I talk I under talk I got to do the interfering has got to act with me for a while see because it doesn't it won't interfere with me properly now now I hear them talking you know....(Howes 1964; cited in de Villiers 1978)

> I felt worse because I can no longer keep in mind from the mind of the minds to keep me from mind and up to the ear which can be to find among ourselves. (Goodglass 1973)

Paraphasic speech such as this may be modeled as selective impairment to the semantic procesor, as shown below.

(4) Paraphasic speech

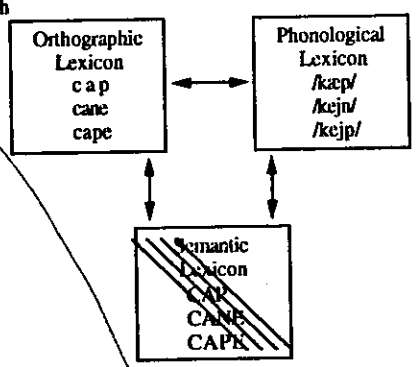


Access to the Lexicon is normal but fed by a dysfunctional semantic processor: patients produce actual words but they are not selected in a semantically coherent way. Syntax is also (apparently) relatively unaffected: phrase length is typically quite long, phrases are coordinated and subordinated, etc. Output from the syntax feeds the Phrasicon, which also functions normally to add function words and inflectional affixes. Finally, the Phonetic processor interprets prosodic domains in terms of intonational contours and correctly converts phonological strings into phonetic ones: the prosody and articulation of such patients is often totally normal. Their speech rate sometimes even exceeds that of normals, as if nonsensical but grammatical sentences were easier to form than meaningful ones.

Nothing seems to be amiss with any particular word or affix, with any (local) syntactic structure, or with pronunciation or intonation. As Blumstein (1978:6) puts it, the speech output of such aphasics is "fluent but empty".

With a sufficiently articulated model of the Lexicon, such as Fromkin's (1985) model, paraphasia may be modeled an impairment implicating part of the Lexicon.

(5) Paraphasic Speech



A semantic sub-lexicon such as this obviates the need for a semantic processor altogether. Fromkin (p.c.) points out that a semantic sub-lexicon provides a locus for paraphasias of this sort: if someone thinks color and says race it may be seen not so much as an error in semantic processing as an error in lexical selection in the proper semantic category.

6.2 *Impairment implicating the Lexicon*

At least two other types of impairment can be modeled as resulting from a deficit in (accessing) the Lexicon: neologistic jargon and anomia.

Neologistic jargon involves the creation of content word stems:

> He wife saw the wonting to wofin to a house with the umbledor. Then he left the wonding then he too to the womin and to the umbella upstairs...
> (cited in Goodglass and Kaplan 1972)

> Well all I know is somebody is clipping the krepes.... Now this here, I'm confoy here, because the have explained what I don't know.
> (cited in Goodglass 1978)

Here the Lexicon must be implicated in the impairment: the word-formation component produces non-normal output not attributable to deficient or abnormal semantic input or selection. Again, note that pronunciation, intonation, syntax and athe selection of function words and inflectional affixes may be almost totally spared in neologistic jargon.

Anomia is similar, except that patients use generic existing content words instead of creating new ones. A woman reported on in Saffran et al. (1980) is a typical example. Although this patient "constructs well-formed sentences, they are grossly deficient in specific lexical content" (1980:223). The patient (P) is being asked to describe a picutre in which a window has been broken by a baseball; a man is coming out of the house, pointing a finger at a little girl; a boy in baseball gear crouches behind a fence, out of sight.

> P: The guy did something, right there... He ran... and she's there like she didn't even know.
> E: Who broke it?
> P: She would never do it. She looks like a really nice kid. He's really getting mad (pointing to the man)... He did it (pointing to the boy). He broke it.
> E: How?
> P: I can't tell you, but I know what it is. It is just broken. 'Cause this kid did it.
> E: What kind of "kid" is that?
> P: Him.

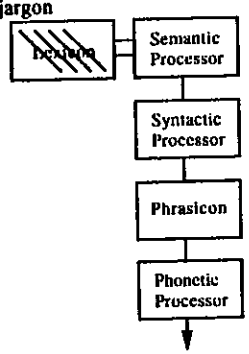
Paraphasic and anomic speech provide support for the claim that linguistic information is modularly represented in the brain.

> These findings are suggestive, since they point to the non-homogeneity of linguistic knowledge, the psychological distinctness of lexical and syntactic processing. For the meanings of individual lexical items must be represented differently from the structural patterns by which they are

> combined, if severe breakdown of the former can occur with no apparent disruption of the latter. (Lincharger 1989:204)

This isolability of lexical knowledge may be modeled as follows:

(6) Anomic and/or neologistic jargon



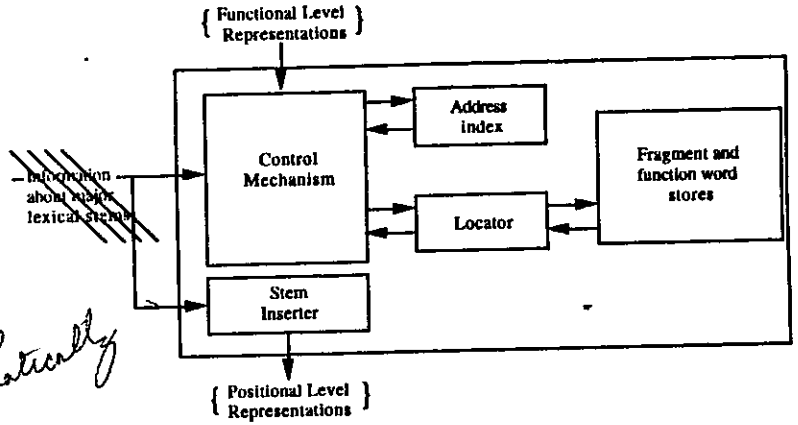
But note that these patients do not merely show a differential impairment of lexical items *vs.* structural patterns; they also show a differential impairment among lexical items. Content word stems are mis-selected (paraphasia), created (jargon) or may be reduced in number (anomia) but function words and inflectional affixes remain intact. This is represented above by the separation of the Lexicon and Phrasicon: impairment to the former does not necessarily imply impairement to the latter.

The Lexicon(s)

The selective problems that such patients have with content word stems and the relative ease with which they produce function words and inflectional affixes suggest that their access to the former has become impaired while their access to the latter has been relatively spared. This is most easily understood on a production model in which content word stems are stored and accessed differently than grammatical items. The REG model in (2) is not unique, of course, in having this feature.

Consider Lapointe & Dell's EG model, for instance. Recall that the syntactic processor includes a subcomponent, the fragment and function word stores, that contain phonological and syntactic information about function words and inflectional affixes.

(7) Syntactic Processor (Lapointe 1985)



Information about major lexical stems is stored in the lexicon, a separate component altogether. Damage to the lexicon or to the access route that links the lexicon to the syntactic processor would result in just the type of impairment seen in Wernicke's aphasia: problems producing and comprehending content words, without comparable problems in producing and comprehending function words and inflection.

Note that a double-listing model, such as that of Bradley, Garrett and Zurif (1980), can equally well describe this sort of selective impairment:

For the most part, anomalous sentences were also repeated verbatim. Semantically anomalous sentences, for instance, such as *You brush my teeth in the morning, The door opened the woman* were also usually repeated without change. And most sentences with near nonwords were repeated verbatim (*A yellow tencil*) unless the object referred to was presented to the patient visually: presented with a yellow pencil, she repeated *A yellow tencil* as *a yellow pencil*. Finally, nonsense words (*libl, shnart*) were usually echoed wihtout modification. In general, "HCEM's performance demonstrates that she had access to lexical storage in her grammar if and only if a verbal or a visual context was provided" (1976:52).

Remarkably, however, sentences that involved a violation of inflection or function word were repeated *with the violation corrected*:

HW: *One pencils.
HCEM: One pencil.
HW: *I talk to her yesterday.
HCEM: I talked to her yesterday.
HW: *She like to drink coffee.
HCEM: She likes to drink coffee.
HW: *You kidding me.
HCEM: You're kidding me.
HW: *Is the money him?
HCEM: Is the money his?
HW: *In a first place, read this.
HCEM: In the first place, read this.

Such corrections cannot be due to visual cues since the function words and inflectional affixes in general do not refer to anything in the visual field. Nevertheless, As Whitaker put it, "her echolalic mechanism was coupled with a 'grammatical filter'" that allowed her to correct ill-formed sentences that contained violations involving function words and inflectional affixes. (1976:40). The grammatical filter correctly supplies both function words and inflection but failed to filter out most lexical improprieties; again, access to

information about lexical items seems to be differentially disrupted: while content word improprieties were generally not corrected (unless visually prompted), function word and inflectional improprieties were.

A number of types of aphasia, then, can be modeled as a loss of access to components or sub-components of the Lexicon with no loss of access to other components of the grammar/processor, including the Syntactic prosessor or the Phrasicon. Impairements that single out these modules seem to occur as well.

6.3 Impairment implicating the Syntactic Processor

Impairment implicating only the syntactic processor but sparing the Lexicon is found in the speech of *agrammatics*. For most (but not all) agrammatic patients, this is accompanied by a morphological impairment in the use of function words and inflectional affixes. This may be called *general agrammatism*. We must also distinguish two other types of case, however, in which the syntactic and morphological (function word and inflection) deficits occur separately. Selective impairment of the syntax may be observed with relatively normal use of function words and inflection; and selective impairment of the use of function words and inflection may be observed with relatively normal syntax (see below, 6.4)

General Agrammatism

Agrammatism is a type of aphasia generally associated with Broca's aphasia. Unlike the speech of Wernicke's aphasics, which tends to be as fast or faster than normal speech, the speech of Broca's aphasics tends to be slow, labored, inarticulate and unmelodical. Agrammatic production is marked by a number of features, as summarized by Kean (1985, following Tissot *et al.* 1973):

1. The deletion[1] of function words in discourse, that is, the deletion of conjunctions, prepositions, articles, pronouns, auxiliary verbs, and copulas (Notable exceptions to this are the conjunctions and and because).
2. The predominance of nouns, at the expense of verbs, in some forms of agrammatic speech.
3. The loss of verb infleciton, with substitution of the infinitive for finite verb forms.
4. Loss of agreement of person, number , and gender, most notable in inflected languages. Jakobson (1963) points out that in languages with case declensions for nouns, nouns revert to the nominative form.

The result is "telegraphic speech". Characteristic samples bear this out:

> Cinderella...poor...um 'dopted her...scrubbed floor, um, tidy...poor, um, ...'dopted...Si-sisters and mother...ball. Ball, prince um, shoe.... Scrubbed and uh washed and uh...tidy, uh, sisters and mother, prince, no, prince, yes. Conderella hooked prince. [Laughs.] Um, um shoes, um, twelve o'clock ball... (Schwartz, Linebarger and Saffran 1985)

> Ah oui! Grève. Grève. Euh, marcher, drapeau rouge. Euh, matraque. Enfin, matraque, Faculté. Euh, ah oui: dix pour cent, salaire. Euh, bah! c'est tout.
> (Ah yes! Strike. Strike. Euh, walk, red flag. Euh, bludgeon. Well, bludgeon, Faculty. Euh, ah yes: ten per cent, wage. Euh, bah! that's all.)
> (Lecours and Rouillon 1976)

As Schwartz, Linebarger and Saffran (1985) stress, "telegraphic speech" should not be taken too literally: it is not simply the case that agrammatic speech *equals* normal speech *minus* infelction and function words. The syntactic structures used by agrammatics are

---

1 'Deletion' is generally determined by omissions involving items in obligatory contexts. In an agrammatic sentence like *The boy show a Valentine's day [card]*, a number of items may have been deleted:

The boy show(s)...
The boy show(ed)...
The boy (is) show(ing), etc.

Which of these elements has been deleted is an interesting question, but one that is often difficult to answer. What is not at issue, however, is that some grammatical morpheme or morphemes required by the context has not been supplied: this is an 'omission'.

quite limited. Schwartz et al. point out, for instance, that the dative causes some difficulty for a number of patients trying to describe a picture of a boy giving a valentine to a girl:

D.E.: The boy is gave... The boy is gave the card.
H.T.: The boy show a Valentine's day... The boy and the ;girl is valentine.
V.S.: The girl...the boy is giving a...giving his girlfriend. The boy valentine the girl. the boy givin' valentine to girl.
P.W.: The boy is valentine the girl. The boy is giving the valentine and the girl is pleased.
M.F.: Valentine's day and candy I think Valentine's day. Girl is Valentine's day...Boy is getting with the girl valentine's candy.

If agrammatics simply failed to produce function words and inflection, they argue, we would expect utterances like *Boy give valentine girl*.

Grodzinsky (1984, 1990) forcefully makes the point that agrammatics both omit and *substitute* grammatical markers. Omission only seems to take place in contexts where a possible word results (*dog for dogs, eat for eats*):

> Such is the English case: the singular form in the nominal system and the present tense form in the verbs system are both uninflected. Consequently, plurals, possessives, and verbal inflections are omitted in English agrammatic speech. But in languages such as Russian, Italian, and Hebrew, where omission of the nominal, adjectival, and verbal morphology would result in lexical ill-formedness, such elements are not omitted (although other closed-class items are). Rather, certain inflectional elements are substituted for others. In Hebrew this is due to the nonconcatenative nature of the morphology, and in the other languages it is due to the fact that in many instances there is no zero-inflection option. (Grodzinsky 1990:52-3)

Any anlaysis of agrammatism must take these qualifications into account.

A Phonological Account of General Agrammatism

Arguing that the class of items impaired in agrammatism cannot be characterized in syntactic or semantic terms, Kean (1977, 1980) proposes a model of the agrammatic

production deficit in terms of stress. Morphemes that are stressed (*dog, eat, blue*) or that affect the placement of stress within a word (*-ity*) are not impaired in agrammatic speech; those that are stress neutral (*to, the, -ed, -s*) are.

As Kean herself notes, "deficits are to be characterized in terms of the impairment to some component(s) of the language faculty" and not simply in terms of a level of representation(1980:260). That is, a characterization of agrammatism merely in terms of a phonological (or syntactic or semantic) level of representation does not go far in explaining how agrammatism comes about: a representational deficit is a symptom of agrammatism, not its cause. Kean therefore suggests that her phonological characterization of the deficit be linked with something like Bradley, Garrett and Zurif's (1980) model and with Garrett's processing model. In this way, her phonological characterization is grounded in an impaired *component* of a processing model.

Kean's account incorrectly predicts that stress neutral derivational affixes (*-ness, -ize*) will be affected to the same degree as stress neutral inflectional affixes, a claim which is not borne out by clinical observations. Indeed, her general claim, that "A Broca's aphasic tends to reduce the structure of a sentence to the minimal string of elements which can be lexically construed as phonological words in his language (1977:25) would seem to predict that *all* affixes would be lost in agrammatism. (See, among others, Grodzinsky 1984 for a critique of Kean's proposal.)

A Syntactic Analysis of General Agrammatism

A number of analyses have been proposed that characterize agrammatism as a pure deficit in syntactic processing (e.g., Berndt and Caramazza 1980, Grodzinsky 1984, 1988; see Schwartz, Linebarger and Saffran 1985 for a more complete discussion). As Schwartz *et al.* put it, these analyses all characterize agrammatism as [Lanugage System minus Syntactic Component]. The general idea is this:

> Without a planned syntactic frame to guide production, lexical items with a purely syntactic function would not be selected by the semantic interpreter. That is, patients' utterances would be expected to be agrammatic. In addition, wihtout adequately selected syntactic structures, we would expect other putput problems such as word order disturbances... The characteristic dysprosody of Broca's aphasics is also a predictable consequence of a failure to select a syntactic frame to guide production.
>
> (Berndt and Caramazza 1980:271)

Grodzinsky's (1990) syntactic account is perhaps the most elaborated. He claims that "the universal characterization of the grammatical representations for agrammatic speech production patterns" is as follows (1990:61):

> At S-structure the representation of agrammatic speech differs from the representation of normal speech in the following respects:
>
> a. Nonlexical terminals are deleted.
> b. Governed prepositions are deleted.

What are "nonlexical terminals?" Grodzinsky offers two definitions but says he is "aware of no data that distinguish between the two" (1990:59):

> A terminal element is lexical iff:
> Def. 1: It is dominated by a category defined by the features [±N, ±V].
> Def. 2: It contains lexical material at a given level.

Definition 1 runs into the problems discussed in Emonds (1985) and above in Chapter 1: pronouns, auxiliary verbs and most prepositions are also dominated by a category defined by the features [±N, ±V]--but they are not retained in agrammatic speech. Grodzinsky handles prepositions by dividing the class in two on syntactic grounds: those that are governed (dative *to*, for instance) and those that are not (lexical prepositions like *over, under*). But this account cannot be extended to pronouns and auxiliary verbs. Pronouns are especially difficult for Grodzinsky's account since he treats proper names like Mary as dominated by NP (1990:169): if proper names and pronouns are both dominated by NP, his definition 1 incorrectly predicts that both will be retained in agrammatic speech.

Grodzinsky does not discuss Definition 2 at any length but notes that it would include "determiners, inflection, auxiliaries, and case markers, as well as the empty categories trace and PRO and their assocaiated indices" (1990:59). If by containing 'lexical material at a given level' he means something like the first stage of level-ordered lexical insertion (Chapter 1 above), Definition 2 is non-distinct from "items in the Phrasicon". Since he does not treat this at any length I will not discuss it further here.

Grodzinsky's account suffers from two problems. The first concerns his (useful) notion of *breadkdown compatibility*. He proposes that an adequate grammar must be constrained by *learnability, parsability* and *breadkdown-compatibility*. The latter is defined as follows:

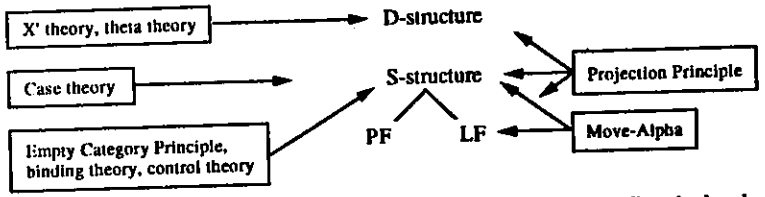
> the patterns of selectivity in the relevant domain observed after brain damage have to form natural classes in the theory [of grammar]. The internal structure of the theoretical account of a domain, then, effectively dictates which patterns of impairment are possible, and which are impossible. An examination of deficit descriptions can be used to evaluate the theory. If the predictions it makes are correct, and it is found to be compatible with breakdown patterns, we can conclude that it meets the neuropsychological constraint of *breakdown-compatibility*.

The problem is that the breakdown patterns he describes (the deletion of non-lexical terminals and governed prepositions) do not form a natural class in the theory of grammar he espouses, Goverment-Binding theory (Chomsky 1981, 1986a, b). GB theory makes available only the natural class {N, A, V, P} and, as Grodzinsky himself makes very clear (ps. 59-62), this is *not* the class of items that are impaired in agrammatism.

Another limitation of Grodzinsky's proposal is that it is characterized in terms of a level of representation (S-Structure) rather than a compenent of the grammar or processor. Recall Kean's claim that "deficits are to be characterized in terms of the impairment to some component(s) of the language faculty". Now the GB model of grammar Grodzinsky uses is *quite* modular, so we might expect the deficits he cites in agrammatism

to be statable in terms of one of these modules. Consider his characterization of the grammar and its modules (his figure 2.2):

(11) A GB Grammar



D-Structure, S-Structure, PF (phonetic form) and LF (logical form) define the levels of representation; modules of the grammar are given in boxes and their domains are indicated by arrows. Such a model of grammar predicts (i.e., is breakdown compatible with) a number of hypothetical language deficits: e.g., one in which case theory fails to apply, one in which nothing moves, one in which X' and theta theory fail to function properly, etc. But none of these modules defines the domain in which the agrammatic deficit is located. Nor does S-structure define it, though it does include it. Essentially, Grodzinsky's analysis boils down to the addition of two deletion rules to the grammar in (11): one deletes non-lexical terminals, the other deletes un-governed prepositions. These rules convert the normal representation, *S-structure*, into an agrammatic representation, *S-structure prime*. Though Grodzinsky's analysis is successful in *locating* the effects of agrammatism, it is unable to isolate the cause of these affects, i.e., the affected module of the grammar.

This is not to say that a syntactic analysis of general agrammatism is impossible. But there may well be something amiss in trying to define general agrammatic production purely in terms of a *syntactic* deficit:

> ...although considerable progress has been made in working out the details of the nature of agrammatism, there remain many unresolved issues. *Only two strong conclusions are possible regarding the structure of*

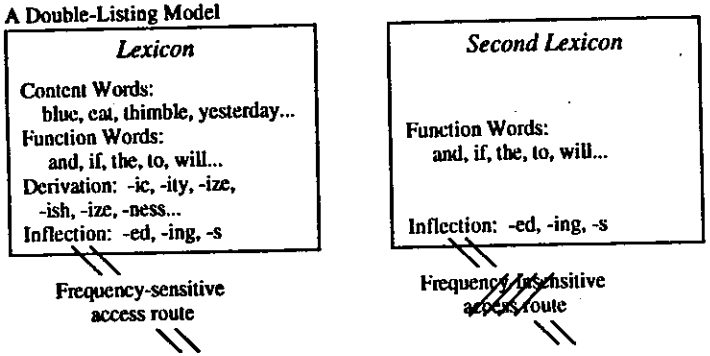
> *agrammatism; it is characterized by the relative omission of free-standing grammatical markers and the omission (in languages like English) or inappropriate selection (in languages like Italian) of inflectional morphology*. The other features of agrammatism...have a somewhat more tenuous status. Thus, it cannot be said with certainty that context has an affect on omission of grammatical markers, that phrase length should be severely restricted, that verbs are nominalized or omitted excessively, or that word-order problems necessarily co-occur with the omission of grammatical markers, despite some evidence for these features.
>
> (Caramazza and Berndt 1985:40--my emphasis)

Perhaps isolating function words and inflectional affixes by pruning syntactic representations upstream is not the answer. Another possibility lies in modeling the disruption of these items more directly.

The Closed Class Hypothesis.

Bradley, Garrett and Zurif (1980) introduced an analysis of agrammatism that singles out a specialized lexicon for *closed class* items, i.e. inflection and function words (see also Zurif 1980, Garrett 1982). As detailed above (Chapter 1), this second lexicon holds rapid-access copies of the function words and inflectional affixes that are stored in the regular lexicon. In agrammatism, access to the second lexicon is taken to be disrupted (*////*):

(12) A Double-Listing Model



On such a model, then, agrammatic speakers may be hypothesized to rely solely on the frequency-sensitive lexicon. Consequently, they must search for *the* and *-ed* in the same large bin in which they search for *blue* and *-ness*. The result is that the and -ed are not found as readily as they are by normals; this is taken to be the cause of telegraphic speech.

This approach has at least two advantages. First, it ties in an analysis of agrammatism with a model of the processor that is *independently* motivated by speech error data in normals, i.e. the Garrett model. Second, it models agrammatism on loss of access to a *component* in the process rather than merely locating the deficit somewhere in one or another *level of representation*.

One disadvantage of the model is its double-listing of closed-class items. The double-listing of grammatical items in the lexicon as well as in the second-lexicon was motivated by studies (Bradley and Garrett 1979; cf. also Bradley 1983) which showed that agrammatics retain the ability to recognize function words as existing words of their language. Moreover, they recognize function words *in the same way as* they recognize content words , namely, as a function of their frequency in everday speech. In this they differ from normals, whose recognition of content words is a function of their frequency but whose recognition of function words is not correlated to their frequency. If function words were not doubly represented in the (first) lexicon, it is reasoned, loss of the second lexicon would result in total loss of function word recognition.

The problem with this involves the disruption of function words and inflectional affixes in agrammatism. If agrammatics have the same speed of access to function words that they have to content words, why do they err primarily in the use of the former rather than the latter? Equal access should yield equal, not differential impairment. Indeed, if agrammatic access to both classes is frequency-sensitive, agrammatics should still have much faster reaction time for function words than content words, since they occur so much

more frequently in everyday speech; but the Bradley and Garrett results show about equal access times for the two classes of words.

A more intricate account of agrammatism is presented in work by Lapointe and Dell (Lapointe 1985; Lapointe and Dell 1989). Their Extended Garrett model includes fragment and function word stores that map syntactic features onto sets of grammatical morphemes. These "notion stores" are essentially highly articulated sub-Phrasicons: one sub-phrasicon for VP, one for NP, etc. Consider the notion store that maps features of VPs onto morphemes that spell out modality, voice and aspect (rows) in terms of person and number agreement (columns):

(13) Partial VP Notion Store for English (LaPointe and Dell 1989)

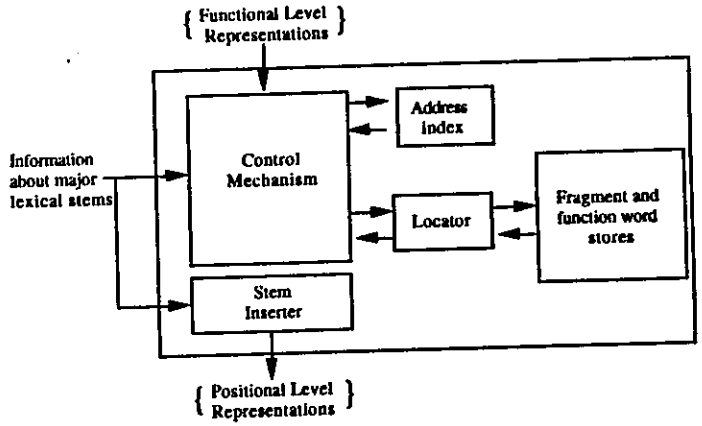
(indic, act, nonspec x	(indic, act, dur) x	(indic, pass, nonspec x	. .
(pres, sing-3) x	(pres, sing-3) x	(pres, sing-3) x x	. .
(pres, sing-2) x (pres, sing-1) (pres, pl)	(pres, sing-2) x x (pres, plur)	. . .	
(past) x	(pres, sing-1) x x		
(anter-pres, sing-3) x x	(past, sing) x x		
. . .	. . .		

The VP Notion Store is organized according to markedness scales based on traditional comparative and historical linguistic studies (ee Lapointe 1985a for discussion).

Lapointe acknowledges that agrammatics tend to produce *certain* function words and inflectional affixes, e.g., *and* and *the*, quite frequently compared to others. He notes that English speaking agrammatics, for instance, tend to produce verb forms consisting of *V, V + ing* and *is V + ing*. He accounts for these facts (and similar facts for Italian

agrammatics) by proposing that a Control mechanism in the processor scans rows and columns in the VP Notion Store by means of a Locator (see below--I have fused Lapointe's 1985 model with Lapointe & Dell's 1989 description of it; the latter substitutes notion stores for the address index and locator of the 1985 model).

(14) The Syntactic Processor (Lapointe 1985)



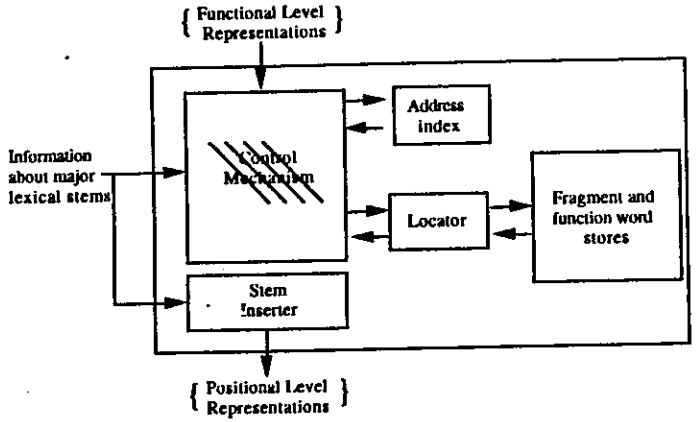
Scansion in the notion stores procedes from left to right and from top to bottom. The farther down and to the right a morpheme lies, the longer it takes to access. Lapointe proposes that agrammatic patients lack some of the resources necessary for deploying the locator: as a result, they tend to limit their search to the left-most and top-most corner of the VP notion store-- producing primarily V, V + *ing* and is V + *ing* in the case of English. This is extended to account for an asymmetry in function words vs. inflectional affixes: the former are produced less often than the latter. Lapointe models this by putting function words lower down in the VP notion store than inflectional affixes--again, agrammatics tend to stay up and to the left of the notion stores.

This model of the processor avoids the pit-falls of double-lising. Only major lexical items are stored in the mental lexicon; function words and inflectional affixes are stored

(only) in the syntactic processor as part of "prepackaged fragments of morphosyntactic structures" (1983:31). Agrammatics have problems accessing function words and inflectional affixes but the problem is stated in terms of the normal access route for retrieving these items. In the Bradley, Garrett and Zurif model, agrammatics have problems finding grammatical items in the first lexicon (where normals presumably never even look for them); in the EG model agrammatics have problems finding these items in the little lexicon (where normals also look for them, but with better results).

In general, the Lapointe and Dell model does an excellent job in modeling agrammatic speech. Note, in particular, that it models the deficit by isolating a single module of the processor, the Syntactic Processor. We might annotate the model to show the deficit as follows:

(15) The Agrammatic Syntactic Processor



Notice that the omission or misuse of inflection and function words on this model seems to be directly tied to the limited syntactic abilities (embedding, etc.) that many agrammatics show.

This is fine for general agrammatism, as defined here. But it is problematic when we consider dissociations within agrammatism. As Lapointe himself notes (1983:33), work by Micceli *et al.* (1983) and Saffran *et al.* (1980) suggests that the omissions and misuse of grammatical items is dissociable from the reduction in syntactic complexity and the well-formedness of utterances. The Lapointe model does not fare well in characterizing a syntactic deficit without a morphological deficit or vice versa.

Syntactic Agrammatism without Morphological Agrammatism

There are two well-studied cases of syntactic agrammatism without a deficit in the production of function words and inflection.

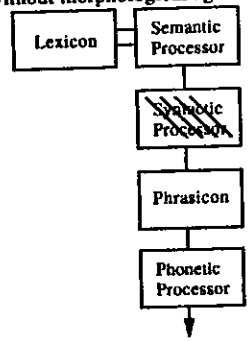
Saffran *et al.* (1980) hypothesize that "the constructional and morphological aspects of agrammatic production are dissociable" on the basis of the asyntactic but morphologically well-formed output of a patient they studied (p. 235). They note that

> while his output is in many ways similar to that of the classical agrammatics, the patient does not omit obligatory grammatical morphemes. He is even able to produce inflectional variants (such as the /s/ allomorph of the third person singular and the non-syllabic /t/ form of the past tense inflection) that are rarely achieved by the agrammatics: A picture of a *girl giving flowers to her teacher* elicits:
> Girl...wants to...flowers...flowers and wants to...The woman...wants to... The girl wants to...the flowers and the woman.
> *A truck towing a car*:
> The...man...uh automobile and truck...the man drives the truck and the automobile.
> *A woman kissing a man*:
> The kiss...the lady kissed...the lady is...the lady and the man and the lady...kissing.
> *A boy drying himself with a twoel*:
> Puts...the man puts on...on his...towel.
> *A woman putting clothes in a washing machine*:

> The lady...the lady launders the...the lady puts the washes...wash on...puts on the wash with the laundry.... (1980:234)

This seems to require a model of the grammar/processor which distinguishes the syntactic module from the module that spells out non-lexical features in syntactic representation (i.e., the Phrasicon). Given such a model, the deficit observed by Saffran et al. may be modeled as follows:

(16) Syntactic agrammatism without morphological agrammatism



Unlike the EG model, the REG model above is able to straightforwardly capture the generalization that "while the patient has a great deal of difficulty in putting a sentence together, the simple structures that he does elaborate are well formed morphologically (ibid, p. 235)

Miceli *et al.* (1983) report on G.G. and T. F., two Italian agrammatics. (T. F. will be discussed below, 6.4.) G. G. is similar to the Saffran *et al.* case. His speech is "slow, effortful, and dysarthric, with a flattened melody and a peculair 'staccato' character" (ibid. p. 68). Although he has only a mild morphological deficit (occasionally omits clitic pronouns, definite articles and prepositions and sometimes substitutes infinitives for inflected forms of the verb) he has a much greater syntactic deficit typical of agrammatics:

> his utterances are largely composed of disjoint phrases, and the main verb is omitted from about one-fifth of his clauses... He is apparently attempting to construct subordinate clauses at some points, but he fails to actually produce them. And much of his output cannot be segmented into sentences. (1983:73)

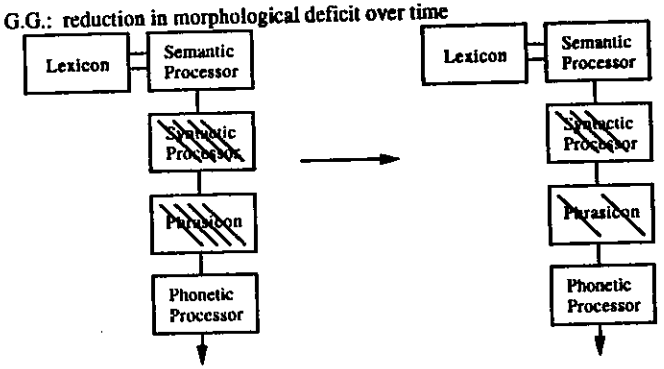
Despite these severe problems with production, his syntactic comprehension (including comprehension of passives and of center-embedded sentences) was within normal limits in all of the tests given.

G. G. is especially interesting because his morphological deficit improved over time. He was interviewed twice, once 3 months postonset and once 14 months postonset. Miceli et al. report that

> In the first interview, the patient showed considerable impairment in both morphology and syntax, as computed by number of omissions in obligatory context. In the second sample, morphological disturbances are clearly reduced, while syntactic deficiencies remain either virtually unchanged or only slightly reduced. (1983:76)

Thus G. G. shows not only a differential impairment for syntax and morphology, but a differential recovery rate of these aspects of agrammatism as well. His change over time might be modeled as follows:

(17) G.G.: reduction in morphological deficit over time



a. 3 months postonset b. 14 months postonset

Such an improvement is difficult to model unless the morphological and syntactic aspects of agrammatism are dissociable. (See Nespoulous 1973 for a similar report on a French agrammatic.)

Syntactic agrammatism without morphological agrammatism argues in favor of a model of the grammar/processor in which syntactic processes are carried out by a different module than the module responsible for the proper selection and insertion of function words and inflectional affixes. The same holds for the opposite type of agrammatism: morphological agrammatism wihtout syntactic agrammatism (see below).

6.4 Impairment implicating the Phrasicon

Impairment implicating only the Phrasicon but sparing the Lexicon and the Syntactic component may be found in cases of morphological agrammatism wihtout syntactic agrammatism. These cases show the same dissociation discussed in (6.3), but with the complementary set of facts. There, it seems that the syntactic module is impaired but the morphological module that selects and inserts function words and inflection is intact; here, the syntactic module is intact but the morphological module that selects and inserts function words and inflection is impaired.

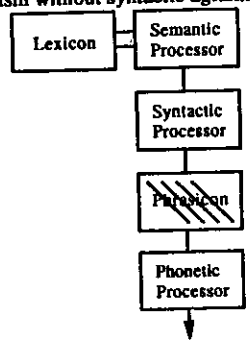
Morphological Agrammatism without Syntactic Agrammatism

Patient T. F. in the Miceli *et al.* study provides essentially complementary symptoms to G. G. He is "free of all dysarthria, dysfluency, and dysprosody; he is, again, a remarkable case of pure agrammatism of speech" (1983:71). T. F. omits inflections and function words in over half of the contexts in which they were required and replaced tensed verbs by infinitives 47% of the time, but has sentences of normal length and complexity. He

> omits no more than three main verbs in some 70-odd clauses. There are a few sequences of phrases which cannot be easily interpreted, but overall his sample of 600 words consists of some 30 well-formed compound and complex sentences; he uses only three simple sentences in his entire narrative.
> Finally, there are two more grammatical properties which distinguish the patients from one another: the use of the clitic pronoun system and the use of auxiliary verbs to form the perfect tense. both of these are very severely impaired for [T. F.] and very mildly impaired for [G. G.]. (1983:74)

Like G. G. and the patient studied by Saffran et al., T. F. points toward a separation between the syntactic processor and the morphological processes that write out non-lexical syntactic features. But whereas the predominantly syntactic agrammatism of G. G. is modeled as an impairment affecting the syntactic processor, T. F.'s purely morphological agrammatism must be modeled as impairment affecting the Phrasicon.:

(18) Morphological agrammatism without syntactic agrammatism



Kolk *et al.* (1985) report on a similar patient, K, an agrammatic Dutch woman in her fifties. K's articulation was only mildly impaired and she retained full control of sentence intonation; phrase length was fairly long and a number of comprehension tests (including comprehension of passives, center-embedded sentences) showed that "she understands the meaning of the grammatical morphemes that were employed" (1985:175). Nevertheless, her speech was marked by a relatively high percentage of omission of function words and inflectional affixes: "although the spontaneous speech of this patient does show the

required telegraphic quality [associated with agrammatics]...in other respects her language behavior is not typical for an agrammatic speaker" (1983:172).

Again, this is best modeled on a grammar/processor which isolates function words and inflectional affixes in a module separate from the modules that control syntax, pronunciation, etc.

At first blush, this would seem to overstate the case: after all, in *most* cases, agrammatic production involves *both* a deficit in grammatical items and one in syntactic complexity. How is this to be explained if the morphological and syntactic aspects of agrammatism come from disruption of separate components of the processor? The answer is probably obvious from the diagram of the grammar above: the syntax provides the input to the Phrasicon. Problems in the input to the Phrasicon would be very likely to cause problems with the correct selection and use of function words and inflectional affixes.

What is surprising, and what the T. F. and K cases show, is that problems with the correct selection and use of function words and inflectional affixes need not co-occur with a syntactic deficit: problems downstream need not be caused by problems upstream. What is more surprising, and what the Saffran *et al.* and G. G. cases show, is that problems in the syntax *need not* cause problems with the correct selection and use of function words and inflection: problems upstream need not cause problems downstream. In either case, then, the syntactic and morphological (function word- and inflection-related) aspects of agrammatism are dissociable and need to be modeled using separate modules of the grammar.

A word on asyntactic comprehension.

It is well-known that production often lags behind comprehension. A non-native speaker of French, for instance, can usually understand more (and better) French than she can

produce. Even for a native speaker, it would seem, sentences may be easier to understand than to produce. Garrett (1980:216) points out that "the production system must get the details of form 'right' in every instance, whether those details are germane to sentence meaning or not"; thus, "unlike comprehension mechanisms, which in principle can often succeed without taking account of grammatical features, the processes of creating an utterance are inextricably bound up with them" (Bock and Kroch 1989:158).

This is important in considering agrammatic production and agrammatic comprehension. Agrammatic listeners seem to be able to interpret far more information from grammatical items than they can use successfully in production. The point here is that we might expect an aphasic with limited access to the Phrasicon to do better at comprehending grammatical items than at producing them. And this seems to be the case. Linebarger (1989), for instance, reports on agrammatic patients who did well in correctly assessing the grammaticality of minimally different pairs such as

> Did the girl enjoy the show?
> *Was the girl enjoy the show?
>
> Which records are you going to give __ to Louise?
> *Which are you going to give __ records to Louise?
>
> The man sat on the new sofa.
> *The man sat the new sofa.

Detecting any of these violations requires access to information about the function words involved.

Still, asyntactic comprehension is well-established, though its limits are not well understood. Linebarger also notes that the patients she studied were unable to distinguish

> non-categorial feature mismatches between anaphorically linked elements. (In the relevant tests) the violations represent mismatches of *number*,

gender, person, animacy between elements which are anaphorically linked by grammatical principles. (1989:224)

She goes on to note that "it might be that these coindexations are represented, but that the mismatch in semantic features between the coindexed elements is not detected" (ibid). Agrammatic comprehension of (non-contentful) prepositions has been studied in German by Friederici (1985) and a recent study by Tyler and Cobb (1987) reveals agrammatic comprehension of inflectional morphology.

Work by a number of researchrs suggests that agrammatic production and asyntactic comprehension are also dissociable (Goodglass and Menn 1985; Caramazza and Berndt 1985; Miceli et al 1983; Kolk et al 1981).

Summary

Agrammatism is clearly not a simple disorder. Patients like G. G. and T. F. show that

> Italian agrammatism, at least in these cases, appears to be a complex of partially dissociable impairments to syntactic (sentence construction) and morphological (infelction and function word) procesing. Derivational morphology appears to be spared in the speech of both patients... (Micelli et al. 1983:82)

Patients in the Saffran et al. (1983) and Kolk et al. (1985) studies suggest that this characterization of agrammatism may hold for English and Dutch as well.

The implications of agrammatism for the grammatical and processing models discussed in Chapter 1, then, should be clear. Impairment that selectively targets function words and inflection but spares content words and derivation argues for a distinction along the lines of what I have called the Lexicon and the Phrasicon. Of the processing models discussed in Chapter 1, this leaves only two: the Extended Garrett (EG) model , with its separate function word and inflectional affix stores and the Revised Extended Garrett (REG) model, with its separate Phrasicon.

These two models differ primarily in whether function words and inflectional affixes are stored in a module other than the syntactic module. According to the EG model, function word and inflection stores are part of the syntactic processor; according to the REG model, they form a separate module, the Phrasicon. What have been called ***syntactic agrammatism without morphological agrammatism*** and ***morphological agrammatism without syntactic agrammatism*** strongly suggest that the module that controls construction of syntactic structure and the module that controls the selection and insertion of function words and inflectional affixes are indeed separate modules. Additional evidence for this separation comes from patients whose syntactic and morphological deficits improve or worsen at different rates. In this way, agrammatism adds invaluable support toone of the central hypotheses of the REG model: that the Lexicon, Syntax and Phrasicon are distinct modules of the grammar/processor.

Word Recognition in Agrammatics

Freiederici (1985) provides experimental evidence from German that lexical prepositions (*steht auf dem Stuhl* 'stands on the chair') and grammatical prepositions (*hofft auf den Sommer* 'hopes for the summer') are processed differently, even when they are homophonous (*auf* vs. *auf*). This provides interesting support for the double-listing of prepositions proposed in Chapter 4 on the basis of prepositions that undergo affixation and compounding.

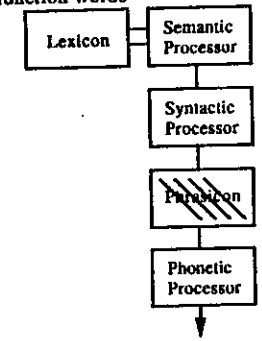
Normal subjects and agrammatics differed in their recognition rates for open and closed class words:

> The agrammatic subjects in our study differed from normals in their use of form class information. Like the normals, they showed facilitation due to semantic context on open, but not on closed class items. Unlike normal listeners who reacted faster to closed class than to open class elements, agrammatic patients recognized open class items faster than...closed class

> items.... So, this study provides some evidence for the position that agrammatics cannot use a special closed class retrieval system to retrieve the functional information of such items. (1985:155)

Whatever it is that enables normals to access grammatical prepositions faster than lexical prepositions is lost to the agrammatics in the study, who recognize content words more quickly than function words. Again, this may be taken as evidence that content and function words are stored in different components of the grammar: such patients display accessing problems with the Phrasicon, with (relatively) unimpaired access to the Lexicon.

(19) Slowed access time for function words



Experiments such as these provide "evidence for a computational distinction of different vocabulary types, and consequently, their attribution to different levels of sentence processing." (Friederici 1985:133)

Deep Dyslexia

There is a close connection between agrammatism and a form of acquired reading disorder called deep dyslexia. Not all agrammatics are deep dyslexics, however, so the two symptoms are dissociable (Carramazza, Berndt and Hart 1981; Coltheart, Patterson and Marshall 1980; Martin, Caramazza and Berndt 1982). In many ways, deep dyslexia is the mirror image of surface dyslexia, discussed above.

Deep dyslexics typically cannot read non-words: they are essentially limited to reading words they already know. According to Newcombe & Marshall, "The salient characteristics [of deep dyslexia] are the predominance of single-word semantic errors and the inability to read function words and to pronounce nonsense words" (1984:187).

Single-word semantic errors give rise to the term 'deep dyslexia'. The cases below are typical: asked to read bun, G.R. responds 'cake', etc.

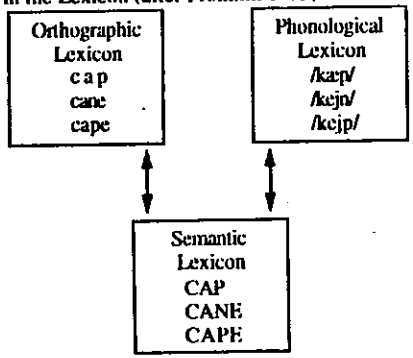
(20) Patient G.R.: Single-word semantic errors

Stimulus	*Response*
BUN	cake
GNOME	pixie
CRAFT	sculpture
AUDIENCE	clap
LITTLE	short
ANTIQUE	ornament
KILL	murder
DEVELOP	camera
OZ	pound
XII	BC

As shown here, the response is often a ('deeply') semantically related word without any phonological or visual (written) similarity to the word presented.

Fromkin (1987) proposes an analysis of semantic errors in terms of her modular model of the Lexicon.

(21) Subcomponents in the Lexicon (after Fromkin 1985)



Note that the connection between the orthographic and phonological lexicons is disrupted in deep dyslexia. Correct and incorrect reading responses arise as follows:

> When these patients read words correctly, there appears to be a direct route from visual orthography to orthographic listing; this either includes a phonological representation or is somehow connected with one. When they err, producing semantically-related word substitutions, this suggest that the semantic representation is separate from the orthographic and phonological representation; the orthographic listing is first mapped onto a semantic listing, which then connects to the phonology. if the wrong semantic listing is selected, then the wrong phonological representation is also selected and produced. (Fromkin 1987:11)

Whereas the surface dyslexic is better able to read function words than content words, the deep dyslexic tends to read only content words. Compare the performance of deep dyslexic G. R. with that of surface dyslexic J.C. above (9).

(22) Patient G. R.: Reading responses to Content and Function words

Stimulus	*Response*	*Stimulus*	*Response*
WITCH	√	BEAN	oh, I know...well...soup...in the soup
WHICH	no!	BEEN	no!
EYE	eyes	BEE	tiny...er...small...bee...bees
I	no!	BE	those...these...not quite sure
HYMN	bible	HOUR	time
HIM	a...a boy?...no!	OUR	no!

KNOT	ankle?	WOOD	√
NOT	no!	WOULD	no!
MOOR	mist...fog (Q) mist	FOUR	√
MORE	no!	FOR	no!

(Newcombe &Marshall 1984--again, √ indicates a correct reading. The response 'No!' is G. R.'s indication that he cannot read the word, not a misreading of the word as 'no'.)

This condition is quite serious:

> The patient can read almost no examples of form-classes apart from the major ones we have described [nouns, adjectives, verbs--C.G.]. No prepositions (n=20), adverbs (n=20), or words from the determiner syustem (n=18) were correctly read. Of the seven personal pronouns (in subject form) the patient could read only one ("I"). No question markers ("where", "when", "why", for example) were read correctly (n=8). Of the simple and correlative conjunctions only one ("and") was read correctly (n=8). (Marshall &Newcombe 1966:172)

Similar low levels of success in reading function words are found in other deep dyslexics.

(23) Percent of FWs read correctly by various patients (from Morton & Patterson 1980)

G.R.	2%	(Marshall and Newcombe 1973)
P.W.	8%	(Patterson 1978, 1979; Patterson and Marcel 1977)
K.F.	11%	(Shallice and Warrington 1975)
V.S.	29%	(Saffran and Marin 1977)

A particularly interesting type of mis-reading of function words is the 'substitution'. Whereas content words are generally misread as visually or semantically similar content words , function words are often misread as totally unrelated function words. Consider the following substitutions by P. W., a patient reported in Patterson and Marcel (1977).

(24) P. W.: Function word substitutions

Stimulus	*Response*	*Stimulus*	*Response*
where	because	and	of
had	of	such	whither
to	which	from	with
his	in	had	and
the	is	had	must

to	it	had	it
just	it	has	because
by	of	has	and
or	with	is	why

(from Appendix 2, Coltheart *et al.* 1980) There are, then, at least two notable differences in the ways content and function words are (mis-)read by deep dyslexics: (i) function words are read less often (correctly or incorrectly) than content words; (ii) function words are often read as semantically and orthographically unrelated function words (*the* > *is*; *had* > *it*; *is* > *why*).

Fromkin takes this as "strong evidence to support the claim that lexical and grammatical morphemes are listed in separate sub-lexicons" (1987:10), i.e., as evidence for a Lexicon vs. Phrasicon distinction.

An alternative analysis

But a one-lexicon analysis is also possible: function words and content words could be listed in the same lexicon but differ in that only content words have semantic addresses. A disconnection between the orthographic and phonological sub-lexicons would result in a reliance on the semantic pathway--words without semantic address (function words) could then not be read at all. This provides a good account of why function words are read less often than content words and it does not require a second lexicon. It is further supported by the fact that many abstract content words (*truth, sleep*) are read less well by deep dyslexics than concrete content words (*girl, fish*). A concreteness hierarchy seems to determine which words are most easily read:

nouns.....adjectives.....verbs.....function words
(easier to read) (harder to read)

This requirs no recourse to a second lexicon.

But such a one-lexicon analysis is less successful in explaining function word *substitutions*. If function words cannot be read by the orthographic -> phonological route nor by the more circuitous orthographic -> semantic -> phonological route (see (21)), how can they be *misread* as other function words?

Thus, a one-lexicon analysis fails to capture the fact that function words are recognized as members of the same class. This is implicit in substitutions, and is made explicit by patients such as G. R. : asked to read *be*, for instance, he responds "Small words are the worst"; asked to read *some*, "One of them horrid words again" (quoted in Fromkin 1987:10). P. W. once responded to an invitation to read by saying "Big words--Yes! Litle words--No!" (Morton and Patterson 1980:270). Responses like these indicate that the patient "does often recognize the functors as "those little words," belonging to a class that is impossible to dicipher" (Newcombe & Marshall 1984:186-7).

Another difficulty with the one-lexicon account sketched above is that deep dyslexics seem to know quite a lot about the semantics of function words; this makes it unlikely that their troubles reading function words may all be traced to semantics. Morton & Patterson report on a number of tests for P. W.'s comprehension of semantic features for function words. On tests that bypassed *production* of function words (i.e., accessing the phonological sub-lexicon), P. W. did very well. Consider his performance in triad tests (examples below) in which he needed to match function words in terms of number and gender, for instance:

(25) Triad tests

(a)	*number*	me	this
		we	that
		us	these
(b)	gender	him	he
		he	hers
		her	she

The test involves picking the word on the left (*me* or *us*) that best goes with the word on the right (*we*): here the correct answers are *us* and *this* (a), *him* and *she* (b). P. W.'s high scores on these and other tests led Morton & Patterson (1980:283) to conclude that "In spite of his very impoverished ability to read function words aloud, P.W. apparently has a great deal of lexical/semantic information about them".

One final difficulty with a one-lexicon analysis of deep dyslexia. Some responses to function word stimuli yield a semantically related content word. The corpus for P.W., for instance, includes the following (Appendix 2, Coltheart et al.1980):

(26) P.W.: function word -> content word errors

Stimulus		*Response*
BENEATH	->	downstairs
FEW	->	little
MORE	->	little
MOST	->	big
SHE	->	girls
HER	->	girl
HE	->	man
IF	->	query

These errors strongly suggest that P. W. has access to the semantics of function words. They equally suggest that P. W. has problems accessing the phonological representa;tions of function words. His corpus shows no errors in the other direction, i.e., content words misread as semantically related function words. Thus it seems that it is not the semanctic representations of function words that gives deep dyslexics problems, but their phonological representations

How may this be represented with two lexicons? P.W. has semantic access to orthographically represented words in the Phrasicon, but he is unable to find anything pronouncable there; consequently, he leaves the Phrasicon and searches the Lexicon for a word with the same semantic feature(s) and pronounces that word instead of the function

word. This explains why content words are not read aloud as semantically similar function words (*GIRL -> she): even if a semantically related function word is found, it will be more difficult to access the phonological form that corresponds to it than it is to access the phonological form of a semantically related content word.

Fromkin's suggestion that function words and content words are stored in separate sub-lexicons is thus well-motivated: the orthographic representation of a function word gets a deep dyslexic into the Phrasicon where the phonological representations of words are not accessible through orthography. This leaves a patient the following possibilities:

(i) pick a semantically related content word that you can pronounce
(ii) pick something that is pronouncable and in the Phrasicon (substitutions)
(iii) just say 'No!' (omissions)
(iv) point out that all you know is what set of words the stimulus belongs to

The orthographic representation of a content word, on the other hand, gets a deep dyslexic into the Lexicon where the phonological representations of words are not accessible through the orthography. This leaves a patient the following possibilities:

(i) pick a semantically related content word that you can pronounce

Since phonological representation in the Lexicon are better preserved than in the Phrasicon, a deep dyslexic has no need to search the latter for a word whose orthographic and semantic representations are in the former.

I have tried to argue, following Fromkin 1985, that a one-lexicon analysis of deep dyslexia fails to account for the problems patients have in reading function words. But the evidence is clearly weakened by the fact that in reading content words, deep dyslexics show more difficulty in reading more abstract, less picturable words (noun > adjective > verb). It is only natural to try and extend this difficulty to function words.

More convincing would be patients who showed a clear deficit with function words, but no deficit with content words, regardless of how abstract they are. This would provide

a clearer case of dissociattion between the putative elements of the Lexicon and Phrasicon. Such cases have been reported; the disorder is often called 'phonological dyslexia'.

Phonological Dyslesia

Patterson (1982:97) reports on a patient, A. M., whom she describes as a 'phonological' dyslexic, following Beauvois and Dérouesné (1979a, b), who report on a similar patient. She describes a number of tests which "demonstrate that A.M., with no syntactic deficit in speech production and no (or only the most minor) syntactic deficit in auditory comprehension, was consistently slower and less accurate in oral reading of function words than of content words." A. M. shows many symptoms of deep dyslexia, as a comparison of his symptoms (below, right) and those of typical deep dyslexics (left) show:

(27) Comparison of a Phonological Dyslexic with typical Deep Dyslexic symptoms

Deep Dyslexia	*Phonological Dyslexia* (A. M.)
Deficit in assembling phonology from print (e.g. reading nonwords)	yes
Deficit in reading abstract words (relatitive to imageable/concrete words)	no
Deficit (relative to nouns) in reading:	
Adjectives	no
Verbs	no
Function words	yes
Occurrence, in reading, of:	
Semantic paralexias	no
Visual paralexias	no (?)
Derivational paralexias	yes
Omissions	no (?)

In particular, A. M. shows a deficit in assembling phonological representations from orthographic ones and a clear (if somewhat mild) deficit in reading function words.

> On the rather loosely defined class of words known as function words (which includes prepositions, conjunctions, pronouns, articleas, auxiliary verbs, and certain adverbs and adjectives), A.M. showed a reading deficit. This was not, as it is for most deep dyslexic patients, a deficit of huge proportions,; but it is worthy of attention because it was (a) consistent; (b) specific to reading (that is, A.M.'s spontaneous speech showed normal use of function words); and (c) his only word-class deficit.
>
> (Patterson 1982:96)

He does not, however, show a deficit in reading abstract words or words that are hard to image (*decree, phase*) and shows no preference for reading nouns over adjectives or verbs:

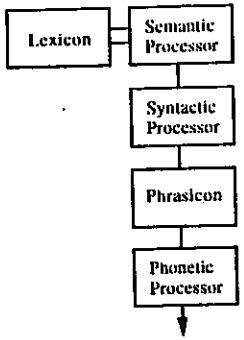
> in terms of oral reading of single content words, A.M.'s accuracy was virtually normal. Furthermore, his performance was robust, showing relative insensitivity to manipulations such as reduced exposure duration or unusual format. (Patterson 1982:83)

A. M.'s condition clearly points to a dissociation between function words and content words that is not tracable to a deficit in reading abstract words or words which have low imagability. This selective deficit in reading strongly suggests that the orthographic representations of content words and function words are represented in such a way that one may be affected without the other.

6.5 Conclusion

I have tried to show that different types of acquired aphasia offer support for the two main hypotheses argued for here: the 2 Lexicon hypothesis and level-ordered lexical insertion. In particular, I have argued that a number of central symptoms observable in different aphasic syndromes can be modeled with the type of grammar/processor in (1) and (2), simplified below:

(28) The grammar/processor



Damage or restricted access to each of the modules in the grammar/processor may be used to model a different set of aphasic symptoms.

The strength of the model, however, does *not* lie in its ability to characterize, predict, or model symptoms observed in different types of aphasia. Rather, its strength lies in the ability to characterize, predict and model aphasic symptoms *linguistically*, in terms of the sub-components of an independently motivated grammatical model.

Closing Remarks

I have argued for two hypotheses concerning both formal grammars and formal models of speech production. The first is that grammars and the processors that implement them contain two lexicons rather than one; the traditional lexicon is thus replaced by two separate modules situtated on opposite ends of a syntactic module. The Lexicon contains content words and derivational affixes; it provides the lexical input for the syntactic module. The Phrasicon contains function words and inflectional affixes; it takes the output of the syntax and annotates it with the phonological representations of purely grammatical items. I have called this hypothesis about the modular organization of lexical storage the 2 Lexicon hypothesis.

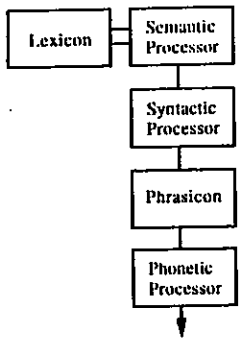
The second hypothesis pursued here is that the selection and insertion of the phonological forms of words and affixes is level-ordered. Phonological forms are not inserted all at once, but in two stages, one taking material from the Lexicon, the other from the Phrasicon. In the first stage of lexical insertion, the phonological forms of content words and derivational affixes are realized; only at a later stage are function words and inflectional affixes converted into phonological strings. I have called this hypothesis Level-Ordered Lexical Insertion.

Evidence for these two hypotheses was drawn from two major areas. Three types of grammatical evidence were presented: (i) minimal word and affix requirements on content words and derivational affixes that fail to hold for function words and inflection, which seems to require something like the 2 Lexicon Hypothesis, (ii) the construction of prosodic constituents, which seems to require something like Level-Ordered Lexical Insertion, and (iii) restrictions on affixation and compounding, which also seem to require something like the 2 Lexicon Hypothesis.

Two broad types of psycholinguistic evidence were also presented: (i) a number of types of speech error were claimed to support Level-Ordered Lexical Insertion in speech production and (ii) a number of types of acquired aphasic and dyslexic orders were claimed to support the 2 Lexicon Hypothesis in the actual storage of lexical information in the mind.

A broad array of evidence, then, supports the hypothesis that lexical information is stored modularly in grammars and in the production models that utilize them. Content words and derivational affixes form a natural class that is distinct from the natural class composed of function words and inflectional affixes. This seems to be true not only for a number of distinct areas of the grammar, but for speech errors and aphasia as well. The 2 Lexicon Hypothesis and Level-Ordered Lexical Insertion help to bridge some of the distance between formal grammars and production models. Hopefully, bridging this gap

(28) The grammar/processor



Damage or restricted access to each of the modules in the grammar/processor may be used to model a different set of aphasic symptoms.

The strength of the model, however, does *not* lie in its ability to characterize, predict, or model symptoms observed in different types of aphasia. Rather, its strength lies in the ability to characterize, predict and model aphasic symptoms *linguistically*, in terms of the sub-components of an independently motivated grammatical model.

Closing Remarks

I have argued for two hypotheses concerning both formal grammars and formal models of speech production. The first is that grammars and the processors that implement them contain two lexicons rather than one; the traditional lexicon is thus replaced by two separate modules situtated on opposite ends of a syntactic module. The Lexicon contains content words and derivational affixes; it provides the lexical input for the syntactic module. The Phrasicon contains function words and inflectional affixes; it takes the output of the syntax and annotates it with the phonological representations of purely grammatical items. I have called this hypothesis about the modular organization of lexical storage the 2 Lexicon hypothesis.

The second hypothesis pursued here is that the selection and insertion of the phonological forms of words and affixes is level-ordered. Phonological forms are not inserted all at once, but in two stages, one taking material from the Lexicon, the other from the Phrasicon. In the first stage of lexical insertion, the phonological forms of content words and derivational affixes are realized; only at a later stage are function words and inflectional affixes converted into phonological strings. I have called this hypothesis Level-Ordered Lexical Insertion.

Evidence for these two hypotheses was drawn from two major areas. Three types of grammatical evidence were presented: (i) minimal word and affix requirements on content words and derivational affixes that fail to hold for function words and inflection, which seems to require something like the 2 Lexicon Hypothesis, (ii) the construction of prosodic constituents, which seems to require something like Level-Ordered Lexical Insertion, and iii) restrictions on affixation and compounding, which also seem to require something like the 2 Lexicon Hypothesis.

Two broad types of psycholinguistic evidence were also presented: (i) a number of types of speech error were claimed to support Level-Ordered Lexical Insertion in speech production and (ii) a number of types of acquired aphasic and dyslexic orders were claimed to support the 2 Lexicon Hypothesis in the actual storage of lexical information in the mind.

A broad array of evidence, then, supports the hypothesis that lexical information is stored modularly in grammars and in the production models that utilize them. Content words and derivational affixes form a natural class that is distinct from the natural class composed of function words and inflectional affixes. This seems to be true not only for a number of distinct areas of the grammar, but for speech errors and aphasia as well. The 2 Lexicon Hypothesis and Level-Ordered Lexical Insertion help to bridge some of the distance between formal grammars and production models. Hopefully, bridging this gap

will allow increased linguistic analysis of psycholinguistic phenomena as well as increased psycholinguistic testing of linguistic theories.

References

Allen, M. R., 1978. *Morphological Investigations.* Unpublished doctoral dissertation, University of Connecticut.

Allen, W. S., 1973. *Accent and Rhythm.* Cambridge University Press.

Anderson, S.R., 1977. On the formal description of inflection. *CLS* 13, 15-44.

Anderson, S.R., 1982. Where's morphology? *Linguistic Inquiry*, 13, 571-612.

Anderson, S.R., 1988. Inflection. In M. Hammond and M. Noonan (eds.) *Theoretical Morphology.* San Diego: Academic Press.

Aronoff, M., 1976. *Word Formation in Generative Grammar.* Linguistics Inquiry Monograph 1. Cambridge: MIT Press.

Bally, C., 1945. Manuel d'accentuation grecque. Berne, France.

Beauvois, M.F. and J. Dérouesné, 1979a. Phonological alexia: Three dissociations. *Journal of Neurology, Neurosurgery and Psychiatry 42*, 1115-24.

Beauvois, M.F. and J. Dérouesné, 1979b. Phonological processing in reading: data from alexia. *Journal of Neurology, Neurosurgery and Psychiatry 42*, 1125-32.

Bloomfield, L., 1933. *Language.* New York: Holt.

Blumstein, S. E., 1978. The perception of speech in pathology and ontogeny. In A. Caramazza and E. B. Zurif (eds.), *Language Acquisition and Language Breakdown.* Baltimore: The Johns Hopkins Univeversity Press.

Bock, J. K. and A. S. Kroch, 1989. The isolability of syntactic processing. In Carlson and Tannenhaus (eds.), *Linguistic Structure in Language Processing*, 157-196. Dordrecht: Kluwer Academic Publishers.

Boomer, D.S. and J.D.M. Laver, 1968. Slips of the Tongue. In Fromkin 1973.

Borer, H., 1986. Introduction. In *The Syntax of Pronominal Clitics, Syntax and Semantics* vol. 19. Academic Press, Orlando.

Bradley, D. C., 1978. *Computational distinctions of vocabulary type.* Doctoral Dissertation, MIT. Distributed by Indiana University Linguistics Club (1983).

Bradley, D.C., and M. F. Garrett, 1979. Effects of vocabulary type in word recognition. *Cognition.*

Bradley, D.C., M. F. Garrett and E. Zurif, 1980. Syntactic deficits in Broca's aphasia. In D. Caplan (ed.), *Biological Studies of Mental Processes.* Cambridge, MA: MIT Press.

Butterworth, B. L.,1980. Some constraints on models of language production. In B. L. Butterworth (ed.) *Language production:* Vol. 1, *Speech and Talk.* London: Academic Press.

Butterworth, B. L., 1989. Lexical access in speech production. In W. Marslen-Wilson (ed.) *Lexical Representation and Process.* Cambridge, MA: MIT Press.

Bybee, J., 1985. *Morphology.* Amsterdam: Benjamins.

Caramazza, A., and R. Berndt, 1985. A multi-component deficit view of agrammatic Broca's aphasia. In M.-L. Kean (ed.), *Agrammatism.* New York: Academic Press.

Chomsky, N., 1957. *Syntactic Sturctures.* The Hague: Mouton.

Chomsky, N., 1965. *Aspects of the Theory of Syntax.* Cambridge, Mass.: MIT Press.

Chomsky, N., 1970. Remarks on nominalization. In *Readings in English Transformational Grammar,* ed. by R. Jacobs and P. Rosenbaum. Waltham, MA: Ginn.

Coltheart, M., 1980. Deep dyslexia: a review of the syndrome. In Coltheart, Patterson and Marshall (eds.), *Deep Dyslexia.* London: Routledge & Kegan Paul.

Coltheart, M., K. E. Patterson & J. C. Marshall (eds.), 1980. *Deep Dyslexia.* Boston: Routledge & Kegan Paul.

Davis, L., N. Foldi, H. Gardner and E. Zurif, 1978. Repetition in the transcortical aphasias. *Brain and Language* **6**, 226-238.

de Villiers, J. G., 1978. Fourteen grammatical morphemes in acquisition and aphasia. In A. Caramazza and E. B. Zurif (eds.), *Language Acquisition and Language Breakdown.* Baltimore: The Johns Hopkins Univeversity Press.

Devine, A. M., and L. Stephens. 1978. The Greek appositives: toward a linguistically adequate definition of caesura and bridge. *Classical Philology,* Vol. 73, No. 4.

Devine, A. M., and L. Stephens, 1981. Bridges in the Iambographers. *Greek, Roman and Byzantine Studies,* Vol. 22, No. 4.

Devine, A. M., and L. Stephens, 1983. Semantics, syntax, and phonological organization in Greek: aspect of the theory of metrical bridges. *Classical Philology,* Vol. 78, No. 1.

Di Sciullo, A.-M. and E. Williams, 1987. *On the Definition of Word.* Cambridge, MA: MIT Press.

Ellis, A.W., (ed.), 1982. *Normality and Pathology in Cognitive Functions.* London: Academic Press.

Elvin, M. and F.M. Hatfield, 1978. Comprehension by agrammatic patients of prepositions and prepositional phrases. *I.A.L.P. Congress Proceedings,* B. Egg (ed.), vol. 2. Copenhagen: Speciale-Paedagogisk Forlag.

Emmorey, K. and V. A. Fromkin, 1988. The mental lexicon. In F. J. Newmeyer (ed.) *Linguistics: the Cambridge Survey,* Vol. 3, *Language: Psychological and Biological Aspects.* Cambridge: Cambridge University Press.

Emonds, J.E., 1985. *A unified theory of syntactic categories.* Dordrecht: Foris.

Ernout, A. and A. Meillet, 1959. Dictionnaire Étymologique de la Langue Latine. Paris: Librairie C. Klincksieck.

Friederici, A. D., 1982. Syntactic and semantic processes in aphasic deficits: The availability of prepositions. *Brain and Language, 15,* 249-258.

Friederici, A. D., 1985. Levels of processing and vocabulary types: Evidence from on-line comprehension in normals and agrammatics. *Cognition* **19**, 133-166.

Fromkin, V. A., 1968. Speculations on performance models. *Journal of Linguistics* **4**: 47-68.

Fromkin, V. A., 1971. The non-anomalous nature of anomalous utterances. *Language* **47**, 27-52.

Fromkin, V. A., 1973 (ed.). *Speech Errors as Linguistic Evidence*. The Hague: Mouton.

Fromkin, V. A., 1985. Evidence in Linguistics. In R. H. Robbins and V. A. Fromkin (eds.) *Linguistics and linguistic evidence*. Newcastle upon Tyne: Grevatt & Grevatt.

Fromkin, V. A., 1987. The lexicon: evidence from acquired dyslexia. *Language* **63**, No. 1.

Fromkin, V. A. 1988. Grammatical aspects of speech errors. In F. J. Newmeyer (ed.) *Linguistics: the Cambridge Survey*, Vol. 3, *Language: Psychological and Biological Aspects*. Cambridge: Cambridge University Press.

Fromkin, V. A., in preparation. Speech Production. In J. B. Gleason and N. B. Ratner (eds.) *Psycholinguistics Today*.

Garnsey, S. M., 1985. Function words and content words: reaction time and evoked potential measures of word recognition. University of Rochester Cognitive Science Technical Report URCS-29.

Garrett, M. F., 1975. The analysis of sentence production. In G. Bower (ed.) *Psychology of Learning and Motivation*, Vol. 9 . New York: Academic Press.

Garrett, M. F., 1976. Syntactic processes in sentence production. In E. Walker and P. Wales (eds.) *New Approaches to Language Mechanisms*. North Holland Publishers: Amsterdam.

Garrett, M. F., 1978. Word and sentence perception. In R. Held, H.W. Leibowitz and H.-L. Teubner (eds.), *Handbook of Sensory Physiology, Vol. III: Perception*. Berlin, Springer.

Garrett, M. F., 1980. Levels of processing in sentence production. In B. L. Butterworth (ed.), *Language Production*, Volume 1: *Speech and Talk*. London: Academic Press.

Garrett, M. F., 1982. Production of speech: Observation from normal and pathological language use. In A. Ellis (ed.), *Normality and pathology in cognitive functions*. London: Academic Press.

Garrett, M. F., 1984. The organization of processing structure for language production: application to aphasic speech. In D. Caplan (ed.) *Biological perspectives on language*. Cambridge, MA: MIT Press.

Garrett, M. F., 1988. Processes in language production. In F. J. Newmeyer (ed.) LInguistics: the Cambridge Survey, Vol. 3, *Language: Psychological and Biological Aspects*. Cambridge: Cambridge University Press.

Garrett, M. F., and M.-L. Kean, 1981. Levels of representation and the analysis of speech errors. In M. Aronoff and M.-L. Kean (eds.) *On Juncture*. San Francisco: Amni Libri.

Gazdar, G., E. Klein, G. K. Pullum and I. A. Sag, 1985. *Generalized Phrase Structure Grammar*. Cambridge, MA: MIT Press.

Golston, C., 1990a. Minimal word, minimal affix. NELS 1990. (to appear in proceedings)

Golston, C., 1990b. Level-Ordered Lexical Insertion. WECOL 1990. (to appear in proceedings)

Golston, C., 1991. Level-Ordered Lexical Insertion in Ancient Greek. Paper presented at the 1991 LSA meeting.

Goodglass, H., 1973. Studies on the grammar of aphasics. In H. Goodglass and S. Blumstein (eds.), *Psycholinguistics and Aphasia*. Baltimore: Johns Hopkins University Press.

Goodglass, H., 1978. Acquisition and dissolution of language. In A. Caramazza and E. B. Zurif (eds.), *Language Acquisition and Language Breakdown*. Baltimore: The Johns Hopkins Univeversity Press.

Goodglass, H., and E. Kaplan, 1972. *The Assessment of Aphasia and Related Disorders*. Philadelphia: Lea & Febiger.

Goodglass, H., and L. Menn, 1985. Is agrammatism a unitary phenomenon? In M.-L. Kean (ed.), *Agrammatism*. New York: Academic Press.

Gordon, B. and A. Caramazza, 1982. Lexical decision for open and closed class items: Failure to replicate differential frequency sensitivity. *Brain and Language 15*, 143-180.

Grodzinsky, Y., 1984. The syntactic characterization of agrammatism. *Cognition* **16**, 99-120.

Grodzinsky, Y., 1990. *Theoretical Perspectives on Language Deficits*. Cambridge, MA: MIT Press.

Halle, M. & K. P. Mohanan, 1985. Segmental phonology of Modern English. *LI* **16**. 57-116.

Halle, M., 1973. Prolegomena to a theory of word formation. *Linguistic Inquiry* 4, 3-16.

Halle, M., 1977. Tenseness, vowel shift, and the phonology of the back vowels in Modern English. *LI* 8. 611-625.

Hamilton, E. and H. Cairns, eds., 1961. *Plato: the Collected Dialogues.* Princeton University Press.

Hammond, M. and M. Noonan, 1989. Morphology in the Generative Paradigm. In Hammond & Noonan (eds.), *Theoretical Morphology.* San Diego: Academic Press.

Hayes, B., 1989. The Prosodic Hierarchy in Meter. In P. Kiparsky and G. Youmans, (eds.), Rhythm and Meter. Orlando: Academic Press.

Hockett, C. F., 1954. Two models of grammatical description. *Word* 10, 210-231.

Howes, D., 1964. Application of the word-frequency concept to aphasia. In A. V. S. de Reuck and M. O'Connor (eds.), *Disorders of Language.* Baltimore: Williams and Wilkins.

Huddleston, R. 1984. *Introduction to the Grammar of English.* Cambridge: University Press.

Hulst, H. van der and N. Smith, 1982. *The structure of phonological representations,* I, II. Dordrecht: Foris.

Inkelas, S., 1987. Tone feature geometry. NELS 17.

Inkelas, S., and D. Zec, 1990. *The Phonology-Syntax Connection.* Chicago: University of Chicago Press.

Jackendoff, R., 1972. *Semantic Interpretation in Generative Grammar.* Cambridge, MA: MIT Press.

Jaeger, J. J., 1986. On the acquisition of abstract representations. *Phonology Yearbook* 3, 71-97.

Jakobson, R., 1963. *Essais de linguistique générale.* Paris: Ed. de Minuit.

Kayne, R. 1975. *French Syntax: The Transformational Cycle.* MIT Press, Cambridge, Massachusetts.

Kean, M.-L., 1977. The linguistic interpretation of aphasic syndromes: agrammatism in Broca's aphrais, an example. *Cognition* 5:9-46.

Kean, M.-L., 1980. Grammatical representations and the description of language processes. In D. Caplan (ed.), *Biological studies of mental processes.* Cambridge, MA: MIT Press.

Keating, P. A., 1988. The phonology-phonetics interface. In F. J. Newmeyer (ed.) Linguistics: the Cambridge Survey, Vol. 1, *Linguistic Theory: Foundations.* Cambridge: Cambridge University Press.

Kertesz, A., 1979. *Aphasia and Related Disorders.* New York: Grune and Stratton.

Kertesz, A., 1981. The anatomy of jargon. In J. Brown (ed.), *Jargonaphasia.* New York: Academic Press.

Kiparsky, P., 1982. From cyclic phonology to lexical phonology. In van der Hulst and Smit (1982).

Kiparsky, P., 1983. Word-formation and the Lexicon. In F. Ingermann (ed.) *Linguistics in the morning calm.* Seoul: Hanshin. 3-91.

Kiparsky, P., 1985. Some consequences of lexical phonology. *Phonology Yearbook* 2, 85-138.

Kolk, H.H. and Blomert, 1982. Lexical access to function words. Manuscript, University o Nijmegen.

Kolk, H. H. J., M. J F. van Grunsven and A. Keyser, 1985. On parallelism between production and comprehension in Agrammatism. In M.-L. Kean (ed.), *Agrammatism.* Orlando: Academic Press.

Lapointe, Steven, 1985a. A theory of verb form use in the speech of agrammatic aphasics. *Brain and Language* 24, 100-155.

Lapointe, Steven, 1985b. A model of syntactic phrase combination during speech production. In S. Berman *et al.* (eds.), *Proceedings of the Fifteenth Northeast Linguistic Society Meeting.* Distributed by GLSA, University of Massachusetts, Amherst, MA.

Lapointe, S. G. and G. S. Dell, 1989. A synthesis of some recent work in sentence production. In G. N. Carlson and M. K. Tanenhaus (eds.), *Linguistic Structure in Language Processing*. Dordrecht: Kluwer Academic Publishers.

Laum, B., 1928. Das Alexandrinische Akzentuationssystem. Paderborn, Schöning.

Lecours, A. R. and F. Rouillon, 1976. Neurolinguistic analysis of jargonaphasia and jargonagraphia. In H. Whitaker and H. A. Whitaker (eds.), *Studies in Neurolinguistics*, Vol. 2.

Lees, R., 1960. *The Grammar of English Nominalizations*. The Hague: Mouton.

Linebarger, M., M. Schwartz, and E. S. Saffran, 1983a. Sensitivity to grammatical structure in so-called agrammatic aphasics. *Cognition* 13, 361-392.

Linebarger, M., M. Schwartz, and E. S. Saffran, 1983b. Syntactic processing in agrammatism: A reply to Zurif and Grodzinsky. *Cognition* 15, 207-213.

Linebarger, M., M. Schwartz, and E. S. Saffran, 1983c. Patterns of syntactic processing in agrammatic aphasics. Paper presented to the annual meeting of the Academy of Aphasia, Minneapolis, October 1983.

Linebarger, Marcia C., 1989. Neuropsychological evidence for linguistic modularity. In Carlson and Tannenhaus (eds.), *Linguistic Structure in Language Processing*, 197-238. Dordrecht: Kluwer Academic Publishers.

Ljung, M., 1974. *A Frequency Dictionary of English Morphemes*. AWE/Gebers, Stockholm.

Low, A.A., 1931. A case of agrammatism in the English language. *Archives of Neurology and Psychiatry, 25*, 556-97.

MacKay, D. G., 1969. Forward and backward masking in motor systems. *Kybernetik* 6:57-64. In Fromkin (ed.) (1973), *Speech Errors as Linguistic Evidence*. The Hague: Mouton.

MacKay, D. G., 1970. Spoonerisms: the structure of errors in the serial order of speech. In Fromkin (ed.) (1973), *Speech Errors as Linguistic Evidence*. The Hague: Mouton.

Marchand, H. 1969. *The Categories and Types of Present-Day English Word-Formation*. München: C. H. Beck'sche Verlagsbuchhandlung.

Marlett, S., and J. Stemberger, 1983. Empty consonants in Seri. LI 14: 617-39.

Marshall, J.C., and F. Newcombe, 1966. Syntactic and semantic errors in paralexia. *Neuropsychologia, 4*, 169-176.

Marshall, J. C., and , F. Newcombe, 1973. Patterns of paralexia: a psycholinguistic approach. *Journal of Psycholinguistic Research*, 2,175-199.

Matthews, P. H., 1972. *Inflectional Morphology*. Cambridge: Cambridge University Press.

McCarthy, J. J., and A. S. Prince, 1986. Prosodic Morphology. ms, UMass and Brandeis.

McCawley, J. D., 1986. Today the world, tomorrow phonology. *Phonology Yearbook* 3, 27-43.

Michelson, K., 1985. Ghost R's in Onondaga: an autosegmental analysis of *R-stems. In L. Wetzels and E. Sezer (eds.), *Studies in Compensatory Lengthening*. Dordrecht: Foris.

Miller, G. A., and P. N. Johnson-Laird, 1976. *Language and Perception*. Cambridge, MA: Harvard University Press.

Mohanan, K. P., 1982. *Lexical phonology*. Bloomington: Indiana University Linguistics Club.

Morton, J., and K. Patterson, 1980. Little words-No!. In Coltheart, Patterson and Marshall (eds.), *Deep Dyslexia*. London: Routledge & Kegan Paul.

Nespor, M. and I. Vogel, 1982. Prosodic domains of external sandhi rules. In H. van der Hulst and N. Smith (eds.), *The Structure of Phonological Representations*. Part I. Dordrecht: Foris, 225-255.

Nespor, M., and I. Vogel, 1986. *Prosodic Phonology*. Foris.

Nespoulous, J. -L., 1973. Approche linguistique de divers phenomènes d'agrammatisme– Étude comparative. Unpublished doctoral dissertation, Universitè de Toulouse-le-Mirail.

Newcombe, F., and J. C. Marshall, 1980a. Response monitoring and response blocking in deep dyslexia. In M. Coltheart, K.E. Patterson & J.C. Marshall (eds.) *Deep Dyslexia*. Boston: Routledge & Kegan Paul.

Newcombe, F., and J. C. Marshall, 1980b. Transcoding and lexical stabilization in deep dyslexia. In M. Coltheart *et al.*, *Deep Dyslexia*. Boston: Routledge & Kegan Paul.

Newcombe, F., and J.C. Marshall, 1981. On psycholinguistic classifications of the acquired dyslexias. *Bulletin of the Orton Society 31*, 29-46.

Newcombe, F., and J. C. Marshall, 1984. Varieties of acquired dyslexia: A linguistic approach. Seminars in Neurology **4:2**.181-195.

Newcombe, F., and J. C. Marshall,1985. Reading and writing by letter sounds. In Patterson *et al.*, 35-52.

Nooteboom, S. G., 1969. The tongue slips into patterns. In A. G. Sciarone, A. J. van Essen & A. A. van Raad (eds.) *Nomen Society, Leyden studies in linguistics and phonetics*. The Hague: Mouton.

Oatley, K., 1978. *Perceptions and Representations: The Theoretical Bases of Brain Research and Psychology*. London: Methuen.

Ohlander, S., 1976. *Phonology, Meaning, Morphology*. Acta Universitatis Gothoburgensis, Göteborg.

Otero, C., 1976. The dictionary in generative grammar. Unpublished paper.

Patterson, K.E., 1978. Phonemic dyslexia: errors of meaning and the meaning of errors. *Quarterly Journal of Experimental Psychology, 30*, 587-601.

Patterson, K.E., 1979. What is right with 'deep' dyslexic patients? *Brain and Language, 8*, 111-29.

Patterson, K.E., 1982. The relation between reading and phonological coding: further neurophychological observations. In Ellis (ed.), 77-111.

Patterson, K. E., and A. J. Marcel, 1977. Aphasia, dyslexia and phonological coding of written words. *Quarterly Journal of Experimental Psychology*, **29**, 307-318.

Patterson, K. E., J. C. Marshall and M. Coltheart, 1985 (eds.). *Surface dyslexia*. London: Routledge & Kegan Paul.

Perlmutter, D. M., 1988. The split morphology hypothesis: evidence from Yiddish. In M. Hammand and M. Noonan (eds.), *Theoretical Morphology*. San Diegeo: Academic Press.

Postgate, J.P., 1924. A Short Guide to the Accentuation of Ancient Greek. University Press of Liverpool.

Saffran, E. M., and O. S. M. Marin, 1977. Reading without phonology: evidence from aphasia. *Quarterly Journal of Experimental Psychology, 29*, 515-25.

Sauzet, P., 1989. L'accent du grec ancien et les relations entre structure métrique et représentation autosegmentale. In *Langages*, **95**, 81-113.

Scalise, S., 1984. *Generative Morphology*. Dordrecht: Foris.

Schwartz, M. F, M. C. Linebarger and E. M. Saffran, 1985. The status of the syntactic deficit theory of Agrammatism. In M.-L. Kean (ed.), *Agrammatism*. Orlando: Academic Press.

Segui, J., J. Mehler, U. Frauenfelder and J. Morton, 1982. Word frequency effect and lexical access. *Neuropsychologia, 20*, 615-627.

Selkirk, E. O., 1978. On prosodic structure and its relation to syntactic structure. Paper presented at the Conference on Mental Representation in Phonology. IULC, 1980. Published in T. Fretheim (ed.) (1981) *Nordic Prosody II*. Trondheim: TAPIR.

Selkirk, E. O., 1980. Prosodic domains in phonology: Sanskrit revisited. In M. Aronoff and M.-L. Kean (eds.), *Juncture*. Saratoga, CA: Anma Libri.

Selkirk, E. O., 1982. *The Syntax of Words*. Cambridge: MIT Press.

Selkirk, E. O., 1984. *Phonology and Syntax*. Cambridge: MIT Press.

Shallice, T., and E. K. Warrington, 1975. Word recognition in a phonemic dyslexic patient. *Quarterly Journal of Experimental Psychology, 27*, 187-99.

Shallice, T., and E. K. Warrington, 1980. Single and multiple component central dyslexic syndromes. In Coltheart, Patterson Marshall (eds.), *Deep Dyslexia*.

Shattuck-Hufnagel, S., 1983. Sublexical units and suprasegmental structure in speech production planning. In P. F. MacNeilage (ed.) *The production of speech*. New York: Springer.

Shattuck-Hufnagel, S., 1986. The representation of phonological information. *Phonology Yearbook* 3.

Siegel, D., 1974. *Topics in English morphology.* MIT doctoral dissertation.

Smyth, H. W., 1920. *Greek Grammar.* Harvard Press

Sommerstein, A. H., 1973. *The Sound Pattern of Ancient Greek.* Publication of the Philological Society, Oxford.

Stemberger, Joseph Paul, 1982a. *The Lexicon in a Model of Language Production.* Dissertation, University of California, San Diego.

Stemberger, Joseph Paul, 1982b. The nature of segments in the lexicon: evidence from speech errors. *Lingua* 56.235-59.

Steriade, D., 1988. Greek accent: a case for preserving structure. *LI*, Vol 19, #2.

Stockwell, R. P., P. Schachter and B. H. Partee, 1972. *The Major Syntactic Sturctures of English.* New York: Holt, Rinehart and Winston, Inc.

Swinney, D.A., E.B. Zurif and A. Cutler, 1980. Effects of sentential stress and word class upon comprehension in Broca's aphasics. *Brain and Language, 10,* 132-144.

Tissot, R. J., G. Mounin and F. Lehrmitte, 1973. *L'Agrammatisme.* Brussels: Dessart.

Tyler, L.K., and H. Cobb, 1987. Processing bound grammatical morphemes in context: The case of an aphasic patient. *Language and Cognitive Processes,* Vol. 2, Nos. 3/4, 245-262.

Vendryes, J., 1929. *Traité d'accentuation grecque.* Paris.

Wang, H. S., and B. L. Derwing, 1986. More on English vowel shift. *Phonology Yearbook* 3, 99-116.

Whitaker, H., 1976. A case of the isolation of the language function. In H. Whitaker and H. A. Whitaker (eds.), *Studies in Neurolinguistics,* Vol. 2. New York: Academic Press.

Zurif, E., 1980. Language mechanisms. A neuropsychological perspective. *American Scientist,* 68, 305-311.

Zurif, E., and Y. Grodzinsky, 1983. Sensitivity to grammatical structure in agrammatic aphasia: A reply to Linebarger, Schwartz, and Saffran. *Cognition* 15, 207-213.