

NEW PARADIGMS:
A RULE-AND-FEATURE BASED, MORPHOLEXICAL ANALYSIS
OF THE SPANISH VERBAL SYSTEM

by

KAREN WILLIAMS BURDETTE

(Under the direction of Dr. Hildebrando Ruiz-M.)

ABSTRACT

This analysis of the Spanish verbal system is based in form on Stephen Anderson's Extended-Word-and-Paradigm model of morphology, but is theoretically based on Joan Bybee's analogical model of morphology. The purpose of the proposed model is to generalize, condense, and describe as elegantly as possible the patterns of Spanish inflectional verbal morphology, so as to produce all attested forms without producing any unattested forms, and without any unnecessary redundancy or superfluous rules. The model is based on linking the two parts of the Saussurean linguistic sign, that is, the *signifié*, represented by feature sets, and the *signifiant*, represented by rules that transform the base lexeme into the fully inflected word. The rules in the model are seen as patterns of organization in the mental lexicon and are not regarded as generative in themselves. Rules and representations are viewed as opposite ends of the same continuum rather than as separate entities in the mental grammar. Fully inflected words are generated by analogical connections among forms in the lexicon, rather than by rules that are independent of the lexical representations. The phonological strings in the *signifiants* in the model are not true morphemes in the traditional sense, but rather, consistent with Anderson's 'a-morphous' morphology, represent a process whereby a base lexeme becomes a fully inflected word. The model is morpholexical, since these word formations occur in the lexicon, and the segments traditionally viewed as grammatical morphemes are not considered significantly different from the segments representing the 'content' part of the word. The model is analogical in nature, since the form-meaning connection can only be made through consistent, sufficient exposure to the phonological forms (or their analogical counterparts) within the context specified by the given feature set, so that the listener/speaker makes connections to forms with similar features, similar physical representations, and similar collocations. In addition to providing insight into the mental organization and processing of language in general and of Spanish verbs in particular, this model has strong implications for the importance of comprehensible input in second language acquisition.

INDEX WORDS: A-morphous morphology, Analogical Modeling of Language, Cohort model, Extended Word and Paradigm, Inflectional morphology, Spanish verbal system, Morpholexical, Word-Formation Rules

NEW PARADIGMS:
A RULE-AND-FEATURE BASED, MORPHOLEXICAL ANALYSIS
OF THE SPANISH VERBAL SYSTEM

by

KAREN WILLIAMS BURDETTE

B.S., The University of Virginia, 1969

M.A.T., Middle Tennessee State University, 1996

A Dissertation Submitted to the Graduate Faculty of The University of Georgia
in Partial Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

ATHENS, GEORGIA

2002

© 2002

Karen Williams Burdette

All Rights Reserved

NEW PARADIGMS:
A RULE-AND-FEATURE BASED, MORPHOLEXICAL ANALYSIS
OF THE SPANISH VERBAL SYSTEM

by

KAREN WILLIAMS BURDETTE

Major Professor: Hildebrando Ruiz-M.

Committee: Don R. McCreary
Marlyse Baptista

Electronic Version Approved:

Maureen Grasso
Dean of the Graduate School
The University of Georgia
December 2002

DEDICATION

For Ricky and Hannah,
without whose love, patience, and support this would not have been possible,
and for Sherry,
who sang to me about dreams and fairy tales.

ACKNOWLEDGMENTS

I would sincerely like to acknowledge and thank Dr. Hildebrando Ruiz-M., my major professor and director of my dissertation committee, along with the other committee members, Dr. Marlyse Baptista and Dr. Don R. McCreary, for their invaluable guidance, linguistic expertise, encouragement, and patience throughout this project. Special thanks also go to Dr. Keith Langston, who provided guidance in the early stages, as well as the inspiration for the topic, which evolved out of his Morphological Theory course. Dr. Sarah Blackwell provided help in the early stages, as did Dr. Diana Ranson, who also graciously allowed me to use her comparison of Spanish and French verbs. Thanks also go to Dr. William Kretzschmar, who might recognize his own influence in this work in its view of language as a dynamic organism with fuzzy boundaries.

I would especially like to thank Dr. David Eddington, currently of the University of New Mexico, who has been my inspiration and mentor since 1994. In the spring of that year I saw a flyer at Middle Tennessee State University about a new course Dr. Eddington was offering on the History of the Spanish Language. I signed up for that course on a lark and then decided to enter MTSU's graduate program full time, although it had been many years since my undergraduate studies in Spanish. All of Dr. Eddington's linguistics courses at Middle Tennessee State were so fun and fascinating that I then decided to further pursue my interest in linguistics at the University of Georgia. All through my years at MTSU and UGA, for the past eight years in fact, Dr.

Eddington has advised, encouraged, and supported me to an invaluable degree through numerous e-mails and phone conversations.

In addition to the above, I would like to thank Tze Lau, Jean Plumer, and all of my other friends and family who kept me going when the going got tough. I would also like to thank all of my former teachers and professors who nurtured my mind and inspired me to follow my dreams. Special thanks also go to Dr. Patrick Conley, my former professor at MTSU who encouraged me to pursue a doctoral degree, as well as to Dr. Judith Rusciolelli, Dr. Sonja Hedgepeth, Dr. Diane Harper, and Dr. Patrice Caux, all of whom taught me how to teach.

Finally, but certainly not least in importance, I would like to thank my parents, Chambers Williams and the late Ruth Cunningham Williams, for giving me a love for language, letters, words, and life; an appreciation for ‘learning for learning’s sake’; and the belief that I could accomplish anything I set out to do.

To all of the above, I am eternally grateful.

TABLE OF CONTENTS

	Page
ACKNOWLEDGMENTS	v
LIST OF ABBREVIATIONS.....	viii
PART I: THEORETICAL BACKGROUND TO THE PROPOSED MODEL	2
CHAPTER	
1 INTRODUCTION	3
2 MORPHOLOGY AND THE LINGUISTIC SIGN	29
3 RULES, REPRESENTATIONS, AND MORPHOLOGY	
MODELS: REVIEW OF LITERATURE.....	54
4 FORM VS. FUNCTION.....	131
PART II: THE PROPOSED MODEL	155
5 INTRODUCTION TO THE MODEL	156
6 FURTHER CONSIDERATIONS	225
PART III: CONCLUSIONS AND PEDAGOGICAL IMPLICATIONS	267
7 ORGANIZATION AND LABELING OF VERB PARADIGMS	268
8 BINDING-ACCESS, CONNECTIONS, AND INPUT	300
APPENDIX: VERB STEM VARIANTS AND FEATURE SETS	343
REFERENCES	347

ABBREVIATIONS USED IN THE PRESENT ANALYSIS

TMA	Tense/Mood/Aspect
PN	Person/Number
Pres	Present
Pret	Preterite
Imperf	Imperfect
Fut	Future
Perf	Perfect/Perfective
Prog	Progressive
Inf	Infinitive
Part	Participle
Indic	Indicative
Subj	Subjunctive
Imprtv	Imperative
Cond	Conditional
Sg	Singular
Pl	Plural
1, 2, 3	First, Second, Third Person

“I believe that those who envision a future world speaking only one tongue, whether English, German, Russian, or any other, hold a misguided ideal and would do the evolution of the human mind the greatest disservice. Western culture has made, through language, a provisional analysis of reality and, without correctives, holds resolutely to that analysis as final. The only correctives lie in all those other tongues which by aeons of independent evolution have arrived at different, but equally logical, provisional analyses”
(*Language, Thought, and Reality*, Benjamin Whorf 1948: 244).

PART I
THEORETICAL BACKGROUND
TO THE PROPOSED MODEL

CHAPTER 1

INTRODUCTION

1.1. What's in a word?

The title 'New paradigms: A rule-and-feature-based, morpholexical analysis of the Spanish verbal system' is a deliberate play on words. 'Paradigm' and 'paradigm shift' seemed to be the buzzwords of the 1990's in pop psychology, science, education, business and management, and society in general. People were encouraged from all sides to adopt some 'new paradigm' as a better framework for managing a business, achieving success, improving human relationships, or just viewing the world around us. In his popular book *The Seven Habits of Highly Effective People*, Stephen Covey (1989) devotes twelve pages to the discussion of paradigms and paradigm shifts. He points out that the word 'paradigm', from the Greek, was "originally a scientific term, and is more commonly used today to mean a model, theory, perception, assumption, or frame of reference." He goes on to say that "In the more general sense, it's the way we 'see' the world – not in terms of our visual sense of sight, but in terms of perceiving, understanding, interpreting" (23).

Swiss linguist Ferdinand de Saussure brought about a paradigm shift in language and linguistics in the early twentieth century when he began studying language synchronically, as it functions as a system at a particular point in time, rather than diachronically, meaning studying how languages change and develop over time. Saussure made a distinction between *langue*, the subpart of spoken language that is systematic and abstract and "exclusively psychological" (Saussure, trans. 1959: 18), and

parole, the secondary “psychophysical” part that is “the individual side of speech, i.e., speaking, including phonation” (18). For Saussure, language is like a symphony, in which “*Langue* represents the unvarying score, *parole* the actual performance, no two of which are exactly alike” (Newmeyer 1986: 32). The paradigm shift initiated by Saussure became known as structuralism, since, in the Saussurean view, “*Langue* represents the abstract system of structural relationships inherent in language, ... a coherent structural system” (Newmeyer 1986: 32). Henceforth, “any approach to language devoted to explicating the internal workings of this system has come to be known as ‘structural linguistics’ or simply as ‘structuralism’” (32). Structuralism dominated the field of linguistics until Noam Chomsky brought about another paradigm shift in the late 1950’s and 1960’s with the advent of generative linguistics. Although Chomsky himself is “firmly within the ‘structuralist’ tradition in linguistics” (71) in that he also sees the grammar of a language as an abstract “autonomous structural system” (71), “Chomsky and his followers began to apply the label ‘structuralist’ just for the earlier autonomous approaches, reserving the term ‘generativists’ for themselves” (72).

Edward Sapir and his pupil Benjamin Whorf were working within the structuralist paradigm when they formulated a theory about language itself as a paradigm. Their hypothesis of ‘linguistic determinism’ claims that language determines the way we think. In the structuralist/deterministic view, language is the paradigm through which we humans view and interpret the world around us. Or, in the milder form of the Sapir-Whorf hypothesis, language at the very least influences, if not determines, the paradigm through which we view the world.

Stephen Bonnycastle (1997) comments that what the structuralists were suggesting is that “language modifies what we see, because it creates the categories we use when we describe things,” and although language to some extent refers to external reality, structuralists “would insist that we can only see the world through the categories and relations that our language emphasizes” (89). So in the structuralist view, “a language (such as English) is very much like a paradigm in science: you have to use a paradigm to see the world, but the particular paradigm you are using determines what you see.” (89).

In linguistics, the word ‘paradigm’ has a different, more specific, meaning. Bybee (1985) discusses paradigms as “groups of related surface forms” (77), referring to verb conjugations and noun declensions. Bonnycastle (1997), although actually referring more to social paradigms, defines ‘paradigm’ as “a kind of model or pattern,” adding that “a number of things which have the same basic form or pattern, or are made on the same model, have a paradigmatic relation to one another” (100). Paradigms in the linguistic sense are groups of inflected forms of a given lexical item, or “groups of closely related forms, like all forms of a given verb in its various tenses” (Ruiz 1986: 104). Anderson (1995) explains this kind of paradigm as follows:

We can now define the paradigm of a lexical item in terms of its lexical stem set: an item’s paradigm is the complete set of surface forms that can be projected from the members of its stem set by means of inflectional Word Formation Rules of the language (134).

This latter kind of paradigm – the linguistic sort – is the kind of paradigm that the present study is focused on. It might be useful, however, to keep in mind that it may be our world-view paradigm in the larger sense of the word that determines the organization of our language paradigms. In other words, our paradigms shape our paradigms.

Without debating extensively the validity of the structuralist and deterministic views of language, we will simply keep these views in mind when we discuss the organization and labeling of verb paradigms and the reality of time and tense. We will see that how a language organizes and labels its paradigms is language-specific to a large degree, and that the organization of the grammar of a language may be influenced by the paradigm through which that language sees the world. Thus the title “New paradigms: A rule-and-feature-based, morpholexical analysis of the Spanish verbal system” refers to at least a threefold meaning of the word ‘paradigm’: 1) the paradigms through which we humans see the world in a broad sense, 2) the language paradigm through which we see the world, and 3) verb paradigms, and more specifically, Spanish inflectional verb paradigms.

The proposed model is rule-and-feature based. The rules in the model describe an operation that transforms a stem or some other basic part of a word into a fully inflected word form. These word formation rules are linked to the semantic and grammatical features that are associated with the word represented by the given rule. The word ‘rule’ is used somewhat loosely here, however, because the rules in the model are really just descriptive patterns showing how the different kinds of verb forms are related to each other in the mental lexicon of the language. Skousen (1989) comments that “An optimal rule description serves as a kind of metalanguage that efficiently describes past behavior

and allows us to talk about that behavior,” but that “if we wish to predict language behavior rather than just describe it, we must abandon rule approaches” (139).

As in Bybee’s view, rules and representations in the present analysis are not regarded as independent from each other and are not “two distinct and discrete components of the description” (1988: 121). Bybee argues that “the best exemplar of a rule and the best exemplar of a representation are two poles of a continuum” (121). She proposes “a model which can account for the rule-like nature of human language without forcing unnatural dichotomies” (121), that is, dichotomies that separate the rules from the representations in descriptions of language behavior. Rather than regarding rules as independent processes that act upon the representations in the lexicon, Bybee and the present analysis regard rules as descriptions of the most generalized representations in the lexicon. Completely suppletive forms, on the other hand, are the most specific kinds of lexical representations and are at the other end of the continuum. Hence, a rule is simply a description of a representation, both in Bybee’s model and in the present proposed model.

The changes that build up the fully inflected words in the language (in other words, the morphology of the language) are viewed as taking place in the lexicon, the mental dictionary, rather than in some other component of the grammar, such as the syntax or the phonological component. Thus, the model is morpholexical. Anderson (1988b) defines the lexicon in a traditional way, assuming that “the lexicon is primarily the locus of what is arbitrary and unpredictable about the words of a language” (184). In this view, it may seem contradictory to have predictable morphological processes occurring in the lexicon, as do Bybee and Skousen. However, Anderson himself goes on

to say that “The lexicon must thus contain ways of relating existing words, in so far as these are (at least partially) systematic, and also for describing the formation of new words” (184). Bybee says that “Previously the lexicon has been conceived of as the mental counterpart of a dictionary, a list of forms set down once and for all (1988: 131).

However, she proposes

a more dynamic representation [of the lexicon] in which not all forms have the same status, but rather in which forms are affected by use or disuse. Frequently used forms gain in lexical strength and forms that are not used lose lexical strength. Lexical strength, then, is an index of word frequency (131).

Rather than asking whether a given word is in the lexicon or not, Bybee proposes that we should ask instead “what is its lexical strength?” and “what are its lexical connections?” (116). This view of the lexicon as a dynamic, constantly changing component rather than a dictionary-like listing of words, forms, or morphemes, is important and relevant in that it can “account for the psychological, historical, and cross-linguistic effect of frequency on morphology” (131).

The proposed model is an analysis and description of the inflectional morphology of Spanish verbs, meaning the changes and relationships based on person, number, tense, mood, and aspect. The model does not deal with derivational morphology, which is the aspect of word formation that adds segments to the word to change the word’s basic meaning or word class. The fact that the present analysis deals only with inflectional, rather than derivational, morphology is partly an arbitrary choice on the part of the present author. Bybee’s analogical model of morphology is entirely capable of handling derivational morphology, as will be seen more clearly when Bybee’s model is fully

explained in Chapter 3. The present proposed model of Spanish verbal inflectional morphology is based in form on the kind of Word-Formation Rules that Anderson (1982, 1986, 1995) uses, and this representational style is especially suited for inflectional, rather than derivational, morphology. Even though the lexicon is seen in the present analysis as a dynamic, rather than static, component of language, inflectional morphology is more of a closed system than is derivational morphology. The constants in the proposed model of inflectional morphology are relatively fixed compared to the variables, since the initial variables in the model are the true ‘content-bearing’ parts of word forms. The variables, which in effect are the inflectional affixes, represent an essentially closed class, at least in a synchronic sense for a given dialect the language, but derivational morphology is subject to more expansion and more arbitrary combinations. Historical and regional variations exist in verbal inflections, of course, but it is nevertheless possible, and fairly easy, to list all the inflectional variations of a given verb. For this reason, it is more feasible to formulate a set of Word-Formation Rules that describe the totality of the Spanish inflectional verbal system than to formulate a set of rules that would describe all of the derivational combinations and connections of forms in the language. However, the decision to focus solely on inflectional morphology in the present analysis is largely an arbitrary choice of the author.

1.2 Purpose and essence of the proposed model.

Anderson (1995) contends that “the task of a theory of morphology is to bring order and coherence to our understanding of the way words are composed and related to one another” (2). The model proposed in the present study attempts to “bring order and

coherence” to the understanding of the complex system of inflectional verbal morphology in Spanish. The proposed model does this by generalizing and condensing the patterns of Spanish inflectional verbal morphology into a set of descriptive rules, so as to produce all attested forms without producing any unattested forms, and without any unnecessary redundancy or superfluous rules. The essence of the proposed model is the link between the two parts of the linguistic sign as defined by Saussure. The two-part Saussurean linguistic sign consists of the *signifié*, which is the mental concept, the meaning, or the semantic content, and the *signifiant*, which is the mental imprint of the physical sound that represents the given meaning. In the model of Spanish verbal morphology proposed in the present study, the *signifié* is represented by semantic/grammatical feature sets, and the *signifiant* is represented by descriptive rules that transform the base lexeme into the fully inflected word.

The model proposed in the present study corroborates Anderson’s Extended Word-and-Paradigm model of morphology, as well as his a-morphous morphological theory, and it also relates Anderson’s model to Bybee’s analogical model of morphology and Marslen-Wilson’s cohort model of lexical processing. The proposed model is based in form primarily on Stephen Anderson’s (1982) Extended Word-and-Paradigm model of morphology, as will be seen in Chapters 3, 5, and 6. It is also theoretically based on Anderson’s (1995) a-morphous morphology, since the present proposed model is also a ‘morpheme-less morphology’. However, unlike Anderson’s model, the proposed model concurs with Bybee (1985, 1988) in regarding morphology as taking place primarily in the lexicon; thus the present analysis is morpholexical. The proposed model is also theoretically based on Joan Bybee’s analogical model of morphology, which emphasizes

the role of semantic and phonological connections between words that are similar in form or meaning. William Marslen-Wilson's (Warren and Marslen-Wilson 1987, Marslen-Wilson 1987) cohort model of lexical processing, although not a morphology model, also provides a theoretical base for how words are built up in the lexicon in either receptive or productive tasks.

Chapter 2 of the present study, 'Morphology and the Linguistic Sign', discusses the Saussurean 'linguistic sign' from a morphological viewpoint, since the proposed model is based in form on the link between the two parts of the linguistic sign. The concept of 'word' is considered, along with a brief discussion of how and where words are composed, organized, stored, and accessed in the mental grammar. In other words, this chapter deals with how and where the morphological component works and whether it is a separate component of the grammar of a language. Problems with traditional views of morphemes and morphology are addressed, such as the problem of incongruities in the function-to-form relationship. This chapter provides an explanation of the notion of 'semantic feature sets' and discusses why grammatical features might be considered as part of the overall semantic feature set of a given word. Finally, a brief explanation of Anderson's view of words and morphology leads to the more detailed explanation and examination of Anderson's, Bybee's, and others' theories presented in the next chapter, all of which provide a theoretical background for the proposed verb model presented in subsequent chapters.

Chapter 3, 'Rules, Representations, and Morphology Models', first presents a general background synopsis of morphology and morphemes, including structuralist and generativist views, before discussing Anderson's 'a-morphous morphology' (1995) and

his arguments against traditional treatments of morphology. The differences between inflectional and derivational morphology are then considered. An explanation of Aronoff's (1976, 1992) Word-Formation Rules and Anderson's (1982, 1986, 1988b) Extended-Word-and-Paradigm model of morphology is presented, leading to the subsequent explanation of the present model's Word-Formation Rules. This chapter also explains the ordering of Word-Formation Rules, including conjunctive and disjunctive ordering and the 'Elsewhere principle', as well as discussing lexical stems, stem alternants, lexical stem sets, and paradigms. Skousen's (1989, 1995) analogical model of language is then discussed as it relates to Bybee's model of morphology. Bybee's "morphology as lexical organization" (1988) and her view of morphology as "a study of the relation between morphology and form" (1985) are then discussed, relating Bybee's model to the present morpholexical analysis of Spanish verbs, which also regards morphology as 'lexical organization'. Marslen-Wilson's (1987, 1989, 1999) cohort theory of lexical processing is discussed as it relates to lexical processing and production. All of the discussion in this chapter provides a formal and theoretical basis for the model of Spanish inflectional verbal morphology proposed by the present study.

Chapter 4, "Form and Function," considers the discrepancies between grammatical labels and the reality they represent, since grammatical labels are an important part of the model proposed in the present study. Such labels are a necessary part of the descriptive Spanish verb model proposed in the present study, since the model depends upon the link between the *signifié* and the *signifiant* for a given verb form. The grammatical labels used in the proposed model, and, in fact, wherever and whenever such linguistic labels are used, are nothing more than a convenient shorthand notation used by

linguists for specific grammatical forms of a language, just as a word is a shorthand notation for the concept it represents. This chapter explores the reality of time and tense, and then explains the various semantic values that these grammatical labels may represent for Spanish verbs in the proposed model.

The proposed model presented in full in Chapters 5 and 6 is essentially a set of morpholexical Word-Formation Rules that attempt to condense, generalize, and describe the patterns of Spanish inflectional verbal morphology as concisely and elegantly as possible. The proposed model is introduced in small increments throughout the early chapters, beginning with the brief synopsis presented in the next section of this Introductory chapter. Chapter 5 presents a broader view, introducing the Basic Model for categorical/regular verbs in the Present Indicative, Present Subjunctive, Imperfect Indicative, and Imperfect Subjunctive. Then the proposed model's treatment of irregular verbs is discussed, including the categories designated as phonologically variant, exceptional/regular, and idiosyncratic/suppletive.

Chapter 6 continues the presentation and discussion of the proposed model, dealing first with lexical gaps and defective verbs. The models for nonfinite forms (the Infinitive, along with Present and Past Participles) and Future and Conditional are then presented, including phonologically variant Past Participle, Future, and Conditional forms. The Basic Model for categorical/regular Preterite is then presented, followed by the exceptional/regular Preterites and idiosyncratic/suppletive Preterites. The irregular categories referred to here as 'allo-class' and 'j-allo-class' are introduced and explained here as well, and compound forms are then considered.

Chapter 7, ‘Organization and Labeling of Verb Paradigms’, presents an overview of the traditional ways of organizing and labeling verb paradigms for analytical and pedagogical purposes, comparing Spanish with several other languages in this regard. This overview serves as background and justification for the grammatical notations used in the model of Spanish inflectional verbal morphology proposed in the present analysis. This discussion also points out the variability, arbitrariness, and language specificity of such paradigms and their labels and organization. Constants and variables in different kinds of verb paradigms are also considered, including Bybee’s analogical model of morphology, thus establishing a basis for the constants and variables used in the proposed model.

Chapter 8, ‘Binding-Access, Connections, and Input’, first looks at why verbs are important in second language acquisition, and then discusses Terrell’s binding-access model of language acquisition and how it relates to Bybee’s analogical model of morphology and to the model proposed in the present analysis. Paradigms in language acquisition are then considered, followed by a discussion of the lexicon, lexical access, and matching procedures in the lexicon. Successive sections consider the role of input and interconnections in language acquisition and the relevance of frequency of occurrence to lexical strength, and hence to the regularity and irregularity of forms. Asher’s observation about the contrast between child and adult language acquisition leads into the discussion of explicit grammar instruction and the role of input and output in the proposed model. This chapter takes another look at the nature of rules and relates the concept of rules as descriptive patterns to language systematicity and variation, showing how the proposed model reconciles Anderson’s rule-based model with Bybee’s analogy-

based model. Finally, this chapter considers the extremely complex, dynamic nature of the lexicon, the human brain, and language acquisition, and sums up the pedagogical implications of the present analysis. The model supports VanPatten's Input Processing model (Lee and VanPatten 1995), Krashen's input hypothesis (1982), Krashen's emphasis on leisure reading for comprehensible input at the intermediate level (1995), as well as Communicative Language Teaching's emphasis on oral input at all levels. In addition, the model does support some degree of explicit grammar instruction in the classroom.

The Appendix following the final chapter presents a list of some of the lexical stem sets ('allomorphs' in traditional terminology) that might feed into the proposed model, along with the respective feature sets that determine which alternant is chosen for a given inflected verb.

1.3. How the proposed model works: A brief synopsis.

As in Anderson's Extended Word-and-Paradigm model (1995), the rules in the proposed model are based on constants and variables. Traditional analyses and traditional verb paradigms regard the true semantic part of the word (that is, the root or the stem) as the constant and the grammatical morphemes as the variables. However, the present analysis and Anderson's model take an opposite view. The constants are the phonological material that many words have in common – the phonological strings commonly referred to as grammatical morphemes, which in Spanish are affixes attached to the end of a given lexical stem. Furthermore, these constants belong to a closed class in the sense that they are essentially fixed, immutable phonological strings, except for

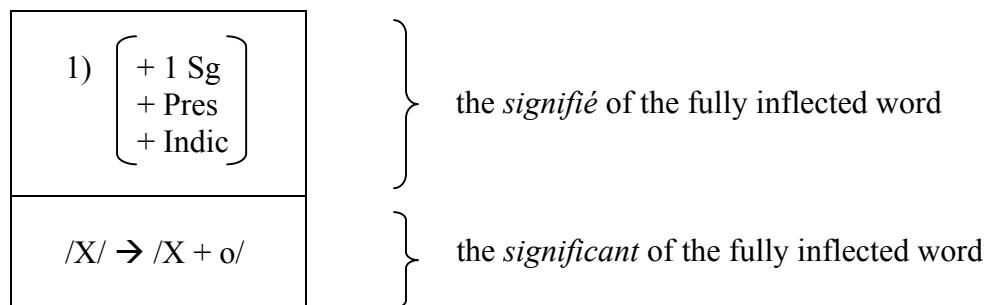
possible diachronic or dialectal variation. The variables that enter the model, on the other hand, are the phonological strings that carry the semantic content of the word. Unlike the constants in the model, these variables belong to an open class, since this class of semantic variables can be augmented by adding new words to the language. This alternate view of constants and variables is important to the way the proposed model functions, since the variables add more phonological material as they progress through the blocks of the model, as explained below.

At the outset of the proposed model, the variables are essentially the segments traditionally regarded as ‘lexemes’ or ‘stems’, that is, the content-bearing part of a given word. However, as the given variable progresses through the blocks in the model, it adds phonological material so that it no longer qualifies as a lexeme or stem in the traditional sense. For example, for the 2 Singular Imperfect Subjunctive form of the *e*-class verb *comer* ‘eat’, the first X variable entering the model at the beginning is /kom/, which is a root or lexeme in the traditional sense. However, one of the rules in the first block in the model adds /ye/, as per the rule /X/ → /X + ye/, yielding the form /komye/. This form (which is not a whole word in itself, nor is it a lexeme in the traditional sense) is the new X variable (thus the second variable) that enters the next block of the model, where it adds /ra/ as per the rule /X/ → /X + ra/, yielding /komyera/, or *comiera* in conventional orthography. Although this form is, in fact, a whole word in itself (the 1 Singular and 3 Singular Imperfect Subjunctive form), it is not the final form for 2 Singular Imperfect Subjunctive. Therefore, /komyera/ exits the second block and then enters the third block as the new X variable (thus the third variable), adding /s/ as per the rule /X/ → /X + s/ and thus exiting as the final correct form /komyeras/, or *comieras*.

The proposed model of Spanish inflectional verbal morphology is a ‘feature-based’ model of grammar, as is Chomsky’s model of syntax (Radford 1997: 59-69) and Anderson’s Extended Word-and-Paradigm theory (1982, 1986, 1995). In a feature-based grammar, words are specified for grammatical features in binary fashion, as having [+] or [-] values for given grammatical features, rather than simply being specified as members of certain categories. For example, the English word *news* is N (noun) [+ Common, – Count, – Plural], while the word *scissors* is [+ Common, – Count, + Plural] (Radford 1997: 61). Anderson uses the notation [± me] and [± you], since “The feature specifications [+ me} and [+ you] indicate 1st and 2nd person forms, while 3rd person forms are [– me, – you]” (Spencer 1997: 216). The proposed model uses binary morphological notations that are slightly different from those used by Anderson, preferring instead to use more traditional notations, as in [+ 3 Sg] or [+ 1 Pl]. A notation of [+] for any given morphological feature or features, such as [+ 3Sg], implies a negative value for all opposing values, which in the case of [+ 3Sg] would be all persons in the plural and first and second in the singular. For the most part, in fact, whenever possible, in the interest of simplicity and economy, the proposed model uses a [+] notation. It is necessary to use this [+] notation, however, because there is at least one instance in the present model where it is necessary to specify a [– Preterite] value. This particular rule is in the last block of the Preterite model, and is stated as /X/ → /X + s/, with the feature set [+ 2 Sg, – Pret]. A minus notation is needed in this rule in order to block application of this rule for 2 Singular Preterite Indicative forms, thus blocking ungrammatical forms such as **hablastes*, **comistes*, **hicistes*.

Figure 1.1 shows one rule-and-feature set from one of the blocks of Word Formation Rules in the proposed model. This particular rule would yield a form such as *compro* or *hablo*, both first person present indicative forms.

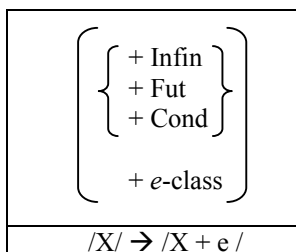
FIGURE 1.1. Sample rule and feature set from the blocks of Word Formation Rules in the proposed model.



The top section (the *signifié*) in Figure 1.1, as in all of the rule-and-feature sets in the model, represents a feature set, with individual features in the feature sets represented by the notation [+ Grammatical Function] enclosed in brackets, as is customary in syntactic and phonological theory. Within these brackets in the top half of the rule-and-feature set, all of the features must apply. For example, in the rule in Figure 1.1, all three of the features in the top of the model must apply, since they are enclosed in regular brackets. Thus, the resulting form is a first person singular present indicative verb form. Square brackets [], within which all enclosed features must apply, are used whenever possible in the proposed model, since the model attempts to seek the lowest common denominators and generalize as much as possible. Occasionally, however, it is necessary to include a rule with mutually exclusive features. For this reason, curly brackets { } are

used in the proposed model when the features in a given rule-and-feature set are mutually exclusive features, meaning that only one of the features (or sets of features) inside the curly brackets can apply at any given time. Figure 1.2 below shows an example of a rule-and-feature set that includes curly brackets with mutually exclusive features. In this particular rule, the form produced can only be an infinitive form, a future form, or a conditional form, not all three at once as in the case of the [1 Sg Pres Indic] rule in Figure 1.1. It should be noted, however, that the rule in Figure 1.2 also includes regular brackets in the notation. Thus, the rule means that the resulting form must be [+ e-class] and also be one and only one of the forms indicated within the curly brackets. The rule below would produce a form such as /kome/, which would enter the next consecutive block as the new X variable and would continue through the blocks to produce such forms as *comer* (Infinitive), *comería* (Conditional), *comeré* (Future), with or without further person/number markers on the latter two of these forms. Further phonological material added to the latter two would yield forms such as *comerías*, *comeríamos*, *comerían*, *comeremos*, *comeréis*.

FIGURE 1.2. Sample rule-and-feature set with mutually exclusive features enclosed in curly brackets.



Each of the feature sets given in the blocks is part of the semantic feature set conceptually associated with the *signifié* of a whole inflected word. In the present morphological analysis, only the grammatical features are listed in the semantic feature sets of the blocks, since the grammatical features are the constant part of the rules, while the true semantic part of a given form is the variable that enters the blocks in the model. However, a feature set for a specific word, with a specific value for /X/, would actually include all of the true semantic features associated with the given word as well as the grammatical functions, meaning that it would include all of the various classes, categories, and descriptions that comprise the basic meaning of the word. It should be noted that in the model proposed here the grammatical features of a word are seen as part of the whole set of semantic features, not as something separate. Chialant and Caramazza (1995) explain that for the whole word representation hypothesis,

Morphological information associated with a given lexical form (e.g., *walked*) would then be part of its semantic information (e.g., the meaning of *walked* is represented by the same features that represent the meaning of *walk* plus the feature of past tense) (59).

Similarly, the examples in the preceding paragraph show morphological information as essentially no different from the so-called ‘content-bearing’ semantic information represented by a given linguistic form. The features traditionally called ‘grammatical’ or ‘morphological’ are regarded in this view as part of the overall semantics of a given linguistic form. It is for this reason that the model proposed in the present analysis views semantic and grammatical features as essentially the same thing. They are both part of the overall meaning of a given word. The overall meaning of the English word *dogs*, for

example, in addition to the semantic notion of ‘furry, four-legged domesticated canine’, includes the semantic notion of ‘more than one’, which linguists label ‘Plural’ for the sake of convenience and for notational purposes.

The bottom half of each block represents the *signifiant* part of the linguistic sign. It is the phonological realization of the semantic and grammatical features of the *signifié* represented in the block above. The variable /X/ as it enters the model, enclosed between slashes as phonemes are conventionally notated, represents the content-bearing stem, or base lexical variant. The notation in each of the lower sections of the blocks represents one of the morphological rules that transform the stem into the fully inflected phonological form of the word; thus the symbol \rightarrow means ‘becomes’ or ‘is realized as’.

The rules in the proposed model are disjunctively ordered within each block. This means that they do not necessarily apply in the order in which they appear in the given block, but rather, they apply from most specific to least specific. For example, a rule whose feature set specifies a specific person, number, tense, and mood, such as [+ 1 Sg, + Pres, + Indic], is more specific than a rule that simply specifies [+ *a*-class]. So, if the rule for the feature set [+ 1 Sg, + Pres, + Indic] is /X/ \rightarrow /X + o/, and if there is in the same block another rule /X/ \rightarrow /X + a/ with a feature set that just specifies [+ *a*-class], then the more specific rule applies first, blocking application of the more general one, even though the two rules are in the same block. A 1 Singular Present Indicative form such as *hablo* ‘speak’), as per the rule /X/ \rightarrow /X + o/, for instance, blocks application of the more general rule /X/ \rightarrow /X + a/, even though this latter rule is in the same block, and even though the verb *hablar* is indeed an *a*-class verb.

The Elsewhere Principle, meaning that “Application of a more specific rule blocks that of a later more general one” (Anderson 1992: 132), is the governing principle regarding this disjunctive ordering within the blocks in the proposed model. The rules are mutually exclusive within a given block of rule-and-feature sets, thus stipulating that only one rule within a block is possible. Therefore, only one form can exit a given block. Once a rule applies in a given block, no other rule in that block can apply, and the resulting form then exits that block. Another way of stating the mutual exclusivity nature of the rules within a block is to say that the rules within any given block are in complementary distribution. For example, for the feature set [+ 1 Sg, + Pres, + Indic], as discussed above, the applicable rule is /X/ → /X + o/, so that this rule must apply and therefore blocks the more general rule /X/ → /X + a/ for the feature set [+ *a*-class].

It should be noted, however, for the sake of clarification, that a specific rule does not block the application of an earlier more general one (in a previous block, that is), but rather just the application of a later more general one. For example, the rule /X/ → /X + a/ (for the feature set [+ *a*-class] in the first block is more general than a later, more specific rule with the feature set [+ Imperf, + Indic, + *a*-class] in the second block. Because of the stipulation of conjunctive ordering of the blocks themselves, this more specific rule in Block II cannot cycle back and block the earlier, more general rule /X/ → /X + a/ (for the feature set [+ *a*-class]). The rule for the [+ Imperf, + Indic, + *a*-class] feature set would indeed block a later more general rule, if there were one. It should be noted, however, that this point is made for clarification purposes only, since the stipulation on conjunctive ordering of the blocks naturally prevents cycling back to an earlier block.

The form that exits a given block is the new X variable that proceeds and enters the next block. For example, for the rule $/X/ \rightarrow /X + a/$, a form such as $/a\beta la/$ would exit the block. This form then enters the next consecutive block as the new X variable. Then in the next block this form may pick up more phonological material, such as $/\beta a/$ if the context specifies [+ Imperfect, + Indicative, + *a*-class], as in the rule $/X/ \rightarrow /X + \beta a/$ for the feature set [+ Imperf, + Indic, + *a*-class]. In the third block, the new X variable $/a\beta la\beta a/$ could pick up further material, yielding either *hablabas*, *hablábamos*, *hablabais*, or *hablaban*. However, if there is no feature set in any given block that maps onto the feature set for the given form, then the form exits the block just as it entered, with no changes at all. For example, in the case of the above cited examples, the third block only needs to include rules for the feature sets [+2 Sg], [+1 Pl], [+ 2 Pl], and [+3 Pl], with no specifications for [+ 1 Singular] or [+ 3 Singular]. For the feature set [+1 Sg] or [+3 Sg], the form $/a\beta la\beta a/$ (*hablaba*) exits this final block as it entered, with no changes, as a sort of default form. The more general case (*hablaba*) prevailed, since there was not a more specific rule to take precedence over the more general rule or form. In traditional terms, we could say that the form *hablaba* has a ‘zero morph’ for person and number.

Although the rules within each block are *disjunctively* ordered, as explained above, the blocks (the groups of rule-and-feature sets) are *conjunctively* ordered by stipulation at the outset of the model. Conjunctive ordering means that if a rule can apply in a given block of rules, then it must apply, and the resulting form then must proceed consecutively to the next sequentially ordered block, and so on through all the blocks for the model until there are no more rules that can apply.

Although all of the feature sets in a given block are mutually exclusive, meaning only one can apply, it is not necessary that one of the sets apply in a given block. If none of the rules in a given block applies, then the phonological form exits the block exactly as it entered the block. When one of the rules does apply, then the form that exits that block enters the next block as the new variable /X/. When there are no more rules that apply, the form that exits the block is the final form.

At the outset, upon entering the blocks, /X/ = the base lexical variant, which for most verbs is the infinitive form minus the /ar/, /er/ or /ir/ ending, although for some verbs there may be other stems, that is, stem alternants other than the one based on the infinitive form. For example, the verb *hablar* only has one stem, /aβl/, but the verb *poner* has several: /pon/, /pong/ /pus/, /pondr/. It may be pedagogically useful to consider the infinitive the basic verb form, since the infinitive is the form listed in standard dictionaries, but there is no reason to claim that the infinitive is the main form of a given verb or that the infinitive stem is actually the main lexeme of a given verb. The infinitive itself is a stem, plus a theme vowel, plus the affix /r/. One of the other stem alternants could conceivably be more basic, or it could be that none takes precedence over any other, except perhaps in frequency. The choice of stem alternant for any given form is determined by the feature set for the context of the form. Each stem alternant is specified – or subcategorized, perhaps – for specific contexts. For example, the alternant /pus/ for the verb *poner* is specified for {[+ Preterite, + Indicative], [+ Imperfect, + Subjunctive]}, while /pong/ is specified for {[+ 1 Singular, + Present, + Indicative], [+ Present, + Subjunctive]}.

1.4. Basic verb model.

For all three Spanish verb classes (*a*-class, *e*-class, *i*-class), the basic model in Figure 1.3 is capable of producing all of the following forms for all of what are traditionally considered regular verbs: Present Indicative, Present Subjunctive, Imperfect Indicative, and Imperfect Subjunctive. For the purpose of verifying the rules in the Basic Model presented in Figure 1.3, several traditional verb paradigms of the above tenses and moods for the verbs *hablar* (*a*-class), *beber* (*e*-class), and *vivir* (*i*-class) are listed in Figure 1.4. In addition, for illustrative purposes, below each rule in the model presented below are one or more sample forms for that particular rule. It should be noted, however, that the sample forms given only represent a few of the many possible forms that the rule could yield. It is also important to remember that only one rule in each block can actually apply at any given time, that is, for any given representation as specified by its corresponding feature set. The sample forms are given in conventional orthography.

FIGURE 1.3. Basic verb model: Word-Formation Rules for regular forms of the Present Indicative, Present Subjunctive, Imperfect Indicative, and Imperfect Subjunctive.

Block I

1) $\left[\begin{array}{l} + 1 \text{ Sg} \\ + \text{Pres} \\ + \text{Indic} \end{array} \right]$	2) $\left[\begin{array}{l} + \text{Pres} \\ + \text{Subj} \\ + a\text{-class} \end{array} \right]$	3) $\left[\begin{array}{l} + \text{Pres} \\ + \text{Subj} \\ + e/i\text{-class} \end{array} \right]$	4) $\left[\begin{array}{l} + \text{Imperf} \\ + \text{Indic} \\ + e/i\text{-class} \end{array} \right]$
/X/ → /X + o/	/X/ → /X + e/	/X/ → /X + a/	/X/ → /X + ía/
hablo, bebo, vivo	hable	beba, viva	bebía, vivía

5) $\left[\begin{array}{l} + \text{Imperf} \\ + \text{Subj} \\ + e/i\text{-class} \end{array} \right]$	6) $\left[\begin{array}{l} + \{ 1 \text{ Pl} / 2 \text{ Pl} \} \\ + i\text{-class} \end{array} \right]$	7) $\left[+ a\text{-class} \right]$	8) $\left[+ e/i\text{-class} \right]$
/X/ → /X + ye/	/X/ → /X + i/	/X/ → /X + a/	/X/ → /X + e/
bebie, vivie	vivi	habla	bebe, vive

Block II

1) $\left[\begin{array}{l} + \text{Imperf} \\ + \text{Indic} \\ + a\text{-class} \end{array} \right]$	2) $\left[\begin{array}{l} + \text{Imperf} \\ + \text{Subj} \end{array} \right]$
/X/ → /X + βa/	/X/ → /X + ra/
hablaba	hablara, bebiera, viviera

Block III

1) $\left[+ 2 \text{ Sg} \right]$	2) $\left[+ 1 \text{ Pl} \right]$	3) $\left[+ 2 \text{ Pl} \right]$	4) $\left[+ 3 \text{ Pl} \right]$
/X/ → /X + s/	/X/ → /X + mos/	/X/ → /X + is/	/X/ → /X + n/
hablas, hables, hablabas, hablaras, bebes, bebas, bebías, bebieras, vives, vivas, vivías, vivieras	hablamos, hablemos, hablábamos, habláramos, bebemos, bebamos, bebíamos, bebiéramos, vivimos, vivimos, vivíamos, viviéramos	habláis, habléis, hablabais, hablarais, bebéis, bebáis, bebiáis, bebieerais, vivís, viváis, vivíais, vivierais	hablan, hablen, hablaban, hablaran, beben, beban, bebían, bebieran, viven, vivan, vivían, vivieran

FIGURE 1.4. Three sample verb conjugations for verifying the model in Figure 1.3.

<i>Hablar</i> ('speak', <i>a</i> -class)							
Pres Indic		Pres Subj		Imperf Indic		Imperf Subj	
hablo	hablamos	hable	hablemos	hablaba	hablábamos	hablara	habláramos
hablas	hablais	hables	habléis	hablabas	hablabais	hablaras	hablarais
habla	hablan	hable	hablen	hablaba	hablaban	hablara	hablaran

<i>Beber</i> ('drink', <i>e</i> -class)							
Pres Indic		Pres Subj		Imperf Indic		Imperf Subj	
bebo	bebemos	beba	bebamos	bebía	bebíamos	bebiera	bebiéramos
bebes	bebeis	bebas	bebáis	bebías	bebíais	bebieras	bebierais
bebe	beben	beba	beban	bebía	bebían	bebiera	bebieran

<i>Vivir</i> ('live', <i>i</i> -class)							
Pres Indic		Pres Subj		Imperf Indic		Imperf Subj	
vivo	vivimos	viva	vivamos	vivía	vivíamos	viviera	viviéramos
vives	vivís	vivas	viváis	vivías	vivíais	vivieras	vivierais
vive	viven	viva	vivan	vivía	vivían	viviera	vivieran

A Spanish speaker might notice an apparent problem presented by the 2 Plural Present Indicative *i*-class form generated by the model in Figure 1.3. Strictly speaking, for the verb *vivir*, for example, the model yields the form *viviis*. This apparent problem is addressed in Chapter 5, Section 5.5.

1.5. Significance of the model.

Emmorey and Fromkin (1988) comment that “Psychological models have generally been concerned with how lexical information is accessed or processed and have been less explicit about the representation and structure of the information. The nature of stored representations is important, however, because it may in part determine the nature of the access mechanisms” (124). The present analysis is concerned with this latter issue, that is, the “nature of stored representations” and the “representation and structure” of the

lexicon as it concerns verbal inflectional morphology. Any research that might shed light on how the mind organizes and processes language would benefit language instructors and second language acquisition researchers alike.

If the proposed model is a descriptively viable morpholexical model of Spanish inflectional verbal morphology, if it is capable of generating all forms of all Spanish verbs, and if it is consistent with an analogical model of lexical organization, then it may help to determine the nature of the access mechanisms. Future research possibilities might include empirical research on the psychological validity and pedagogical implications of the model. The primary aim of this analysis, however, is theoretical in nature and is intended to shed light on the organizing principles of the Spanish verbal system.

CHAPTER 2

MORPHOLOGY AND THE LINGUISTIC SIGN

2.0. The model proposed in the present study is a model of word formation based in form on the link between the two parts of the Saussurean ‘linguistic sign’, and it for this reason that this chapter discusses the nature of the linguistic sign, particularly from a morphological viewpoint. The concept of ‘word’ is considered, along with a brief discussion of various views of how and where words are composed, organized, stored, and accessed in the mental grammar. In other words, this chapter deals with how and where the morphological component works and whether it is a separate component of the grammar of a language.

Problems with traditional views of morphemes and morphology are addressed, such as the problem of incongruities in the function-to-form relationship. Since the proposed model is based on the correspondence between the descriptive word formation rules that represent a given form (the *signifiant*) and the respective semantic feature sets that represent the function of that form (the *signifié*), this chapter provides an explanation of the notion of ‘semantic feature sets’ and discusses why grammatical features might be considered as part of the overall semantic feature set of a given word. Finally, a brief explanation of Anderson’s (1995) view of words and morphology leads to the more detailed discussion of Anderson’s and others’ theories presented in the next chapter. This discussion also provides some theoretical background for the proposed verb model presented in subsequent chapters.

2.1 The Saussurean sign.

S.I. Hayakawa (1949) defines language as a system of symbols, pointing out that “human beings have agreed, in the course of centuries of mutual dependency, to let the various noises that they can produce with their lungs, throats, tongues, teeth, and lips systematically stand for special happenings in their nervous systems. We call that system of agreements *language*” (27).

Language, then, is a semiotic system of human communication. Saussure (trans. 1959) discussed in his lectures the symbols that are the building blocks of what we call language, maintaining that language is more than just “a naming-process only – a list of words, each corresponding to a thing that it names” (65). Saussure makes a distinction between the ‘sign’, the ‘signified’, and the ‘signifier’. The basic building block of this system of symbols is the *signe*, the ‘linguistic sign’, consisting of two elements: the *signifié* (the ‘signified’) and the *signifiant* (the ‘signifier’). Furthermore, Saussure maintains that “the linguistic sign unites, not a thing and a name, but a concept and a sound-image” (66). The *signifié* is essentially the mental representation of a thing or concept, or the meaning, while the *signifiant* is the “psychological imprint” (66) of the physical form, or sound sequence, that represents that concept.

In addition to this distinction between *signe*, *signifié*, and *signifiant*, Saussure considers language itself to have two dimensions, which he refers to as *langue* and *parole*. To Saussure, language is distinct from speaking: “Language is speech less speaking. It is the whole set of linguistic habits which allow an individual to understand and to be understood” (77). Language, however, is not just a private matter, since “for the realization of language, a community of speakers [*masse parlante*] is necessary” and

“language never exists apart from the social fact” (77). The present study deals primarily with *langue*, treating language as a system, but it also deals with *parole*, since the analyses are based on what people actually say as well as what the system is capable of generating.

Saussure’s notion of the *signifié* as a mental concept rather than as a physical thing roughly parallels Plato’s notion of the ‘ideal’ or the ‘archetype’. To Plato (*The Republic*, Book VI, trans. Rouse 1956), all human concepts and ideas are just imperfect copies of pure understanding, which is a phenomenon of a higher plane of reality. The word *bird*, for instance, causes a speaker of English to conjure up an image of a bird, although the bird in that image does not really exist. Although this bird image and all real birds exhibit common characteristics, there is no bird anywhere that is exactly identical to this mental image of a bird. Plato refers to an ideal, or a hypothetical archetype, of which all real birds and bird images are imperfect copies (308-311). Similarly, Saussure sees the word as a linguistic sign or symbol, but this sign is not just a physical symbol. It is, rather, is a dual entity made up of the mental imprint of the sound symbol (the *signifiant*) and the mental image of the thing or concept represented (the *signifié*), somewhat akin to Plato’s concept of images and ideals.

Language is a supremely complex organism, and the concept of *word* is not as simple a construct as the average person assumes it to be. If the symbols of language – its ‘signs’ – are such a complex dual entity, the question, then, is how we get from the signified to the signifier, from the meaning to the form (or the ‘psychological imprint’ of that form). How do we get from the mental concept to the spoken and perceived utterance, or to the signed or written equivalent thereof, or from the spoken and perceived

utterance to the mental concept? And if the form changes in different circumstances, or different syntactic environments, as verbs do in most languages, and nouns and adjectives do in some languages, then how are these changes implemented? How and where, and in what component of the mental grammar, are these changes, or inflections, added to the word? And what if these changes are not strictly agglutinative or linear, such as changing *take* to *took* or *taken*, or *poder* to *puedo*, or *decir* to *dije*? What if there is multiple exponence, meaning that morphological functions overlap and there are several categories represented by a single formative? In such cases we might not be able to distinguish exactly what each segment represents, as in the Spanish Preterite *dije*, *dijiste*, *dijo*, *dijimos*, *dijisteis*, *dijeron*, for example. What if the so-called inflections are completely suppletive, such as changing Spanish *ir* to *fue* or English *go* to *went*?

2.2. Where is morphology?

Most linguists agree that language comprises several different components in the brain, but there has been ongoing debate about what occurs in each component (Anderson 1992, Newmeyer 1986, Spencer 1991). The basic components of the grammar are: 1) the lexicon, which stores words, 2) the semantic component, which stores meanings, 3) the syntactic component, which governs word order and the relationships among words in a given sentence, and 4) the phonological component, which ‘spells out’ how the message is to be physically articulated. There is an ongoing debate about where morphology belongs and whether it comprises a separate component of the grammar (Matthews 1974: 3-8; Hall 1992: 6-8). Bloomfield (1933) and the structuralists of the nineteenth and early twentieth centuries considered the lexicon to contain only completely idiosyncratic

information, maintaining that “Any property of a word which can be predicted from, say, the phonology or the syntax, will therefore be excluded from the lexicon” (Spencer 1997: 47).

Although structuralists considered word formation to be the central focus of linguistics, the early generativists did not consider morphology to be an autonomous component of the grammar (Newmeyer 1986; Spencer 1997). In a generativist, rule-based model of grammar, certain morphophonemic rules apply in a specific order to generate the ‘grammatically correct’ form to add onto the base form, or lexeme, which is supplied by the lexicon. Skousen (1989) points out that Chomsky himself “is a structuralist par excellence” and that “For Chomsky (and virtually all other linguists today) there is no doubt that language is rule governed and that language behavior must be accounted for in terms of specific rules” (4). A generativist model would account for the generation of forms that the speaker has never heard before, and for regularized forms of child speech, such as an English-speaking child’s **bringed* (for irregular *brought*), a Spanish-speaking child’s **sabo* (for irregular *sé*) or **morido* (for irregular *muerto*), or even an adult Spanish speaker’s ‘incorrect’ **andó* (for the standard Spanish *anduvo*). Such a rule governed generativist model does not, however, account for irregular forms or for some attested but prescriptively ‘ungrammatical’ irregular forms such as **brang*.

A strictly lexeme-based model, on the other hand, claiming that all words are stored in the lexicon, would account for irregular forms, but would require entirely separate lexical entries for all inflected forms of all verbs. If all forms of all words are stored in the lexicon as whole words, then we might ask whether that means that all forms would have to have actually been heard by the speaker before he could produce

them. The so-called Lexicalist Hypothesis (Jackendoff 1972: 12-13), held by most generative grammarians, claims that no syntactic operations can be performed on words at the lexical level (Spencer 1997: 72-73). Although the Lexicalist Hypothesis originally referred to derivational morphology, the Strong Lexicalist Hypothesis later claimed this principle for inflectional morphology as well. One theory, known as Lexical Functional Grammar, was a theory of syntax based in the lexicon (Spencer 1997: 72-73). Halle (1973) proposed that Word-Formation Rules (WFR's) act on a list of morphemes stored in the lexicon to form fully inflected, concatenated words. He then proposed a filter that would prevent the overgeneration of unattested words and would thus account for idiosyncratic words that depart from the regular word formation rules.

Connectionist models of cognition view form-meaning connections as occurring throughout the whole neural network of the brain, not necessarily in a linear fashion, but rather, through parallel distributed processing (PDP) and through associative processes rather than the construction of rules. The strength of the connections is dependent upon activation of neural networks through associative processes rather than by abstract rules. Rumelhart (1989) explains that

Units interact by transmitting signals to their neighbors. The strength of their signals and therefore the degree to which they affect their neighbors are determined by their degree of activation.... It is this pattern of connectivity that constitutes what the system knows and determines how it will respond to any arbitrary input (139).

In general, non-rule-based approaches are referred to as ‘analogical’ or ‘exemplar-based’ approaches, with Rumelhart’s connectionist model a subset of this broad “class of possible models” (Skousen 1989: 56). This class of models “can all reasonably be referred to as ‘analogical’ and ... can be sharply differentiated from the kind of ‘symbolic’ or rule-based approaches that are exemplified by generative grammars in linguistics” (56). Skousen’s analogical model of language applies to all of language, from the completely regular to the completely suppletive. Skousen explains that “an analogical description predicts behavior by means of a collection of examples called the *analogical set*” (4), rather than predicting behavior based on generative, productive rules. For example, an analogical set in Spanish might include *habló, cantó, bailó, escribió, bebió, vivió* (all 3 Singular Preterite Indicative forms), or *hablamos, cantamos, bailamos, tenemos, escribimos, vivimos* (all 1 Plural Present Indicative forms).

Analogical models of language usage emphasize binding form to meaning through analogy, by searching for the exemplar that is closest to the form in question, both semantically and phonologically. Exemplar-based learning systems “involve hunting for the most similar instances (‘nearest neighbors’) to predict language behavior” (Skousen 1999: 1). Eddington (1999) explains that Analogical Modeling of Language (AML) “is a model of how memory tokens may be used to predict linguistic behavior” (2). AML accepts a whole-word view of the lexicon, with all known words stored as wholes in the mental lexicon.

When the need arises to determine the behavior of an unknown word, the lexicon itself is accessed. A search is conducted for the words most similar to the unknown word. The behavior of the word(s) most similar to the word in question

generally predicts the behavior of the word in question, although the behavior of less similar words has a small chance of applying as well (2).

For example, if the word *charlaron* is unknown to the speaker, he or she can predict the behavior of this word by accessing similar words in the lexicon and applying the patterns of those familiar words to the new word. The speaker accesses the already known words *hablaron*, *cantaron*, and *bailaron* and predicts the behavior of the new word based on the fact that all of these known forms occur with 3 Plural subjects in a context of completed past actions.

The process of making form-meaning connections, in either a connectionist or an analogical approach to language, depends upon the building of feature sets in the lexicon. The lexical entry of any given word consists of a set of features. These features include: 1) semantic features, which include all of the concepts associated with the word, that is, all of the features which make up the basic ‘meaning’ of the word (features such as: ‘animal, mammal, domesticated, four-legged, furry, says woof-woof’, etc.); 2) syntactic features, which specify word class (such as noun, verb, adjective, etc.); 3) subcategorization features, which specify what kinds of complements the word can have (such as whether and what kind of object a verb requires); 4) phonological features, which specify how the word is physically realized, meaning how it is pronounced or signed. In literate people, the orthographic representation, or spelling, is also included in the lexical entry for a given word.

Anderson (1995) views morphology as a process occurring in various components of the mental grammar, rather than in a bounded, separate module, although he says that whether morphology comprises a separate component of the grammar is not an especially

important aspect of his overall theory of morphology. Anderson maintains that while morphology is indeed a legitimate object of study in linguistics, the question of whether morphology is a separate ‘component’ of a grammar is a distinct theoretical issue in itself, and he doesn’t dwell on arguing one way or the other (1995: 8). Part of the morphological process may occur in the lexical component, part in the syntactic component, and part in the phonological component. In many instances there may be a morphological process involved rather than simply the affixation of a discrete, separable phonological unit, as in the traditional definition of morpheme as the minimal linguistic sign. Anderson views morphemes as ‘processes’ rather than as ‘items’; hence his term ‘amorphous morphology’ (1995). Anderson, for the most part, accepts a split-morphology hypothesis, however, maintaining that derivational processes take place in the lexicon, while inflectional morphology is handled by the syntactic and phonological components.

Anderson rejects an Item-and-Arrangement view of morphology, which sees morphemes as separate and separable ‘minimal units of meaning’, with word formation viewed as “the disposition of morphemes in a word” (Spencer 1997: 49). Anderson regards morphology as a process rather than a concatenation of morphemes, aligning himself more with Word-and-Paradigm theorists, who maintain that “there are certain generalizations which can only really be stated at the level of the whole word” and that “the notion of ‘inflectional paradigm’ seems to play some role in grammatical organization” (Spencer 1997: 52). For Anderson and other Word-and-Paradigm theorists, “the most important feature of inflectional systems is their tendency towards non-agglutination, manifested as overlapping exponence and fusion” (Spencer 1997: 221). The present analysis, like Anderson, also recognizes the problems inherent in a traditional

Item-and-Arrangement approach to morphology, especially the problems in the function-to-form ratio.

2.3. Function-to-form correspondence.

When considering the semantic properties of a given word, it is clear that there is never a one-to-one correspondence between meaning and form. If we were to separate all of the semantic features of a word into discrete parts and then represent each with a discrete phonological form, the resulting word would be impossibly long. For instance, if we had one morpheme to represent the concept of ‘noun’, another to represent ‘concrete’ as opposed to abstract, another to represent ‘animate’, another to represent the class of ‘mammals’, another for ‘nonhuman’, another for ‘canine’, and another for ‘domesticated’, then we would have a word composed of at least seven distinct phonological strings simply to signify the concept of ‘dog’. Then if we want to add the feature ‘plural’ to either the hypothetically long string or the actual English word *dog*, then we would add another phonological string, which in the latter case is the single consonant /s/ realized as [z]. If we wanted to add a diminutive feature to the word, we could add /i/ as a distinct phonemic segment to the original word, and we would know, as part of our native linguistic competence, that this segment would come between the main string and the plural string if we wanted to make the word plural as well as diminutive, such as *doggies*.

In traditional morphology, the content-bearing part of the phonological string we refer to as a ‘word’ is regarded as either a single morpheme or as a concatenation of separate derivational morphemes. For example, we regard the English words *dog* and

train as single morphemes, while the words *dogtrainer* and *untrainable* have three morphemes each. The first is a concatenation of the verb *train* with the derivational morpheme *er*, meaning ‘agent’, plus the free morpheme *dog* added at the beginning to indicate the object of this act of training. The second word, *untrainable*, is composed of three morphemes: the free morpheme *train*, as a verb, plus the bound morpheme *able*, meaning ‘capable of’, plus the bound morpheme *un* added at the beginning of the string to signify ‘not’. For the word *dogtrainers*, we would have one additional phonological string /s/ (realized as [z]) and one additional semantic feature: ‘plural’. For the nominalized word *untrainability*, we would add one phonological string to signify one semantic feature, the attribute of being *X*, in this case *untrainable*. As in the above example of *doggies*, we know, as part of our linguistic competence, the order in which these strings are concatenated.

In traditional morphology, the word *dogtrainers* contains three derivational morphemes (*dog/train/er*) and one grammatical morpheme, the noun plural marker /s/. However, rather than regarding these distinct phonological strings as linguistic signs in themselves, each with a separate *signifiant*, Saussure viewed the word itself as a linguistic sign. Although the structure of the above words can be consciously broken down as we have just done, the complete set of semantic features associated with a single word is viewed as being a single *signifié* in Saussure’s classic definition of the linguistic sign, even if it is a morphologically complex word.

If the simple one-syllable word *dog* is accepted as the sign for a whole set of semantic features, then it should not be difficult to view a seemingly complex polymorphemic word as likewise signifying a complex set of features as well. Just as

speakers do not normally separate all the semantic features associated with the word *dog* (such as those listed above), it is conceivable that the mind does not separate the semantic/grammatical features of a morphologically complex word like *dogtrainers* either. The mental image associated with the word *dogtrainers* is that of two or more humans with whatever trappings we might associate in our minds with people who train dogs, rather than a mental concatenation of four separate morphemes.

Linguists can consciously analyze and separate some morphologically complex words into discrete, identifiable parts, and speakers might use various parts of complex words in other complex words. However, the ability of linguists to analyze and separate words into parts does not necessarily mean that such analyses are available to the average speaker of a language. It does not necessarily mean that words are inherently separable in terms of their semantic and grammatical functions. Bybee (1988) sees word structure in a different way, regarding “the internal structure of a word as a set of relations with other words, rather than as a string of discrete meaningful sequences, that is, morphemes” (139).

Furthermore, the fact that linguists may be able to consciously analyze and separate morphologically complex words into discrete, identifiable parts at times does not necessarily imply that grammatical functions in a word are psychologically different from ‘true’ semantic functions. Chialant and Caramazza (1995) explain that for the ‘whole word representation hypothesis’, the part of a lexical form considered ‘morphological’ information (for example, the /ed/ in *walked*) is really just part of the semantic information associated with the given word. The meaning of the word *walked* is “represented by the same features that represent the meaning of *walk* plus the feature of

past tense” (59). Similarly, Bybee refers to ‘the semantics of past tense’ (1988: 137). The notion of past tense is just part of the overall meaning of the word *walked* or the Spanish word *anduvieron* (*walk*, 3 Plural Preterite Indicative).

This view of linguistic signs as whole words is consistent with Saussure’s (1916) classic definition of the linguistic sign, as opposed to the later, more widely held view of linguistic signs as morphemes, or parts of words (as in Bloomfield 1933: 160-163). In a whole-word view of the lexicon such as Bybee’s model (1985, 1988) and the present proposed model, the *signifiés* corresponding to the word parts traditionally called ‘grammatical morphemes’ are viewed as a subset of the overall semantic features of a word, rather than as something entirely separate. This is important for a model of morphology that emphasizes connections among forms based on overall semantic and phonological similarities, rather than on discrete, separable parts of whole words.

2.4. Form, function, and feature sets.

All of the above brings us to the question of verbs, which in most languages have rather complex structures that have been analyzed for centuries by linguists and grammarians. Although English has lost most of its inflectional morphology, Spanish, like its predecessor Latin, maintains a rich verbal morphology, allowing it to be a ‘null-subject’, or ‘pro-drop’ language, meaning that the verb can stand alone with an unstated but anaphorically understood subject. The semantic features of person and number are part of the set of semantic features of the inflected verb itself if we view the whole inflected verb as a ‘word’ in the Saussurean sense of the ‘linguistic sign’. Likewise, the features of tense, aspect, and mood are part of the entire set of semantic features

associated with the verb as a whole word. For example, the verb form *hablaban* is a sign in the classic Saussurean sense, that is, a phonological string signifying the whole set of semantic features including past tense, indicative mood, imperfect aspect, third person, and plural, as well as all of the semantic features associated with the notion of *speak*. We could represent the former of these features, meaning the features other than the ones associated with the lexeme ‘speak’, with the following shorthand notation: [+ 3 Pl, + Past, + Indic, + Imperf].

If we wished to separate all of the semantic features of the lexeme *speak*, we could also include them in the above feature set. For example, the feature set might include [+ verb, + action, + human agent, + noise, + emitted by vocal tract, + abstract or concrete referent], among other possible *signifiés*. There are at least two problems with this, however. In the first place, speakers are not accustomed to breaking down all of the semantic elements of a word in this way. Furthermore, when linguists or lexicographers describe the many semantic elements of a word, they are forced to use other words to describe the various semantic features of a given word. The semantic elements that make up the whole *signifié* of a sign, however, are really mental concepts, or, in Plato’s terms, they are mere images of the ‘ideal’ or ‘archetype’ of the given referent. So the *signifié*, or the set of features associated with a given word, is actually a set of abstract images or concepts in the mind, which may or may not actually be broken down as such in the mind of the speaker or language learner.

If the *signifié* in the mind is nothing more than an image or concept, or a complex conglomeration of images and concepts, then why should tense, mood, aspect, person, and number be singled out as separate, identifiable *signifiés* any more than all of the other

signifiés of the word? Should we separate these last five as discrete elements simply because we can separate them? But just because they can be separated in the *signifié* part of the linguistic sign does not mean that they can necessarily be separated as easily in their physical representation, nor does it mean that there will necessarily be a one-to-one correspondence between the signified and the signifier. In a hypothetical, strictly agglutinating language, we would have separate, discrete phonological strings to signify each of the many *signifiés* of the linguistic sign, but in the real world of real human languages this is not the case. Even in languages traditionally considered agglutinating, such as Hungarian and Turkish, the word is not, in fact, broken down into discrete phonological strings with each representing a single concept. The so-called grammatical functions may be represented by discrete phonological elements in such languages, but the entire feature set is not broken down in this way.

2.5. Problems in the form-to-function ratio.

Moreover, in most languages even these grammatical functions – tense, mood, aspect, person, and number – are not so simply broken down into easily separable units with a one-to-one correspondence between the signified and the signifier. As Bybee (1988) points out,

In the ideal case, all the phonetic material in an utterance can be assigned to one meaningful unit or another. In such a case the semantic and phonological connections run parallel. But not all linguistic material is organized in this fashion (128).

Spanish, for example, is a fusional synthetic language (Spencer 1997: 38), rather than being strictly agglutinative, since there is not always a one-to-one correspondence between form and function as there would be in a strictly agglutinating language. In other words, there is not always an entirely separate phonological string for each grammatical feature to be represented. There are instances of what Anderson (1995) calls ‘cumulative morphs’, “where more than one apparently independent dimension of a paradigm [i.e. different morphological categories] is expressed by a single formative” (54). For example, in the form *hablaras* (‘speak’, 2 Singular Imperfect Subjunctive), the /ara/ segment represents verb class and mood as well as tense. Similarly, in the Preterite form *habló*, the segment /ó/ represents tense, mood, person, and number all at once. A strictly agglutinating language would have entirely separate (and separable) phonological strings for each grammatical function associated with a given concept, such as person, number, tense, aspect, and mood. Spanish also exhibits syncretism, which means using a single form to represent two or more different functions, but not at the same time. The Spanish form *habla*, for example, signifies both 3 Singular Present Indicative, and 2 Singular Imperative. Similarly, /s/ in English is the 3 Singular verbal marker, as well as the noun plural marker and the noun possessive marker. In Spanish verbs there are also cases of what Spencer (1997) refers to as ‘multiple exponence’. An ‘exponent’ is the linguistic material that expresses grammatical categories, and multiple exponence refers to when “a single category is realized in more than one way within a word, in other words, when there is a many-to-one correspondence between form and function” (51). Spanish also exhibits ‘empty morphs’, such as Spanish verb thematic vowels, for which

there is no link to meaning. The /a/ in the ending of *hablan*, for example, is part of the verbal paradigm but has no semantic value.

Furthermore, there are in Spanish, as in most languages, many so-called irregular forms, forms which cannot be produced by simply adding segments to a fixed lexeme. Some of these may resemble the lexeme and may possibly be generated through phonological processes (see Chapter 5 of the present study), such as *pensar* → *piensa* (3 Singular Present Indicative), or *morir* → *muere* (3 Singular Present Indicative). Some may be irregular but have somewhat regular patterns in their irregularity, such as *traer* → *trajo* (3 Singular Preterite Indicative), or *decir* → *dijo* (3 Singular Preterite Indicative]. In both of these two types of irregularities, the irregularity is allomorphy, or alternation of form in the stem, rather than irregularities in the inflectional affixes. Still other verb forms are completely suppletive and do not resemble the base lexeme at all, such as the above *ir* → *fue* (3 Singular Preterite Indicative), or *ser* → *es* (3 Singular Present Indicative). Thus, Spanish verbal inflection is not just a simple matter of adding certain affixes in linear order to a fixed stem. There is also the problem of lexical gaps, or unattested forms, such as **abolo/abuelo* (1 Singular Present Indicative) from the verb *abolir*. And finally, there is the problem of ill-formed tokens, that is, forms deemed by most speakers as ungrammatical, yet well-attested in the language and seemingly generated by some pattern. These ill-formed tokens may appear in dialectal variations of the language, in child language, or in second-language-learner speech, for example: **andaron* for Spanish *anduvieron* (3 Plural Preterite Indicative), **sabo* for Spanish *sé* (1 Singular Present Indicative), or **brang* for English *brought* (Past Indicative).

2.6. Lexical stem sets.

In the case of completely suppletive forms (as with *ir* and *ser*), or even in the ‘irregular’ patterns above (as with *traer* and *decir*), there may be two or more stems stored separately in the lexicon, and these separate stems might then be defined as lexemes themselves, at least in Anderson’s model of morphology. Anderson (1992) suggests a “lexical stem set,” with various forms of a given lexical item already specified for certain morphosyntactic features in the lexicon. When choosing the particular stem to undergo word-formation rules, the one that matches the most features of the morphosyntactic representation is the one selected from the lexical stem set (133). The model proposed in the present analysis deals with stem allomorphy by employing Anderson’s notion of lexical stem sets and by selecting stem alternants based on matching features, as does Anderson’s model. The present analysis regards these lexical stem sets as parts of whole words in the lexicon, rather than as separate lexical entries in themselves.

Anderson’s lexical stem set hypothesis may or may not be consistent with the above discussion of whole words as complete linguistic signs linked in the lexicon to their associated feature sets, which include tense, mood, aspect, person, number features as well as the many semantic features associated with the sign. Whether or not these two notions are consistent with each other depends upon whether the word formation rules take place in the lexicon or in some other component of the grammar. If these word formation rules are realized in the lexicon, then they produce whole, inflected words that are then inserted into the syntax. However, if only the base lexeme, along with its allomorphic variants, is supplied by the lexicon, then the inflectional affixes would be

added in the phonological component, the syntactic component, or some other component of the grammar. This would be true whether the inflectional strings are seen as items, as in traditional morphology, or as processes, as in Anderson's model.

The model proposed in the present analysis accepts the former of these two possibilities, with the Word-Formation Rules seen as a component of a dynamic, multidimensional lexicon. Bybee comments that "Previously the lexicon has been conceived of as the mental counterpart of a dictionary, a list of forms set down once and for all," but she proposes "a more dynamic representation" of the lexicon. This dynamic view of the lexicon is important and relevant in that it can "account for the psychological, historical, and cross-linguistic effect of frequency on morphology" (1988: 131).

In such a dynamic, multidimensional lexicon, the contextual space is not strictly compartmentalized as it is in a rule-based system, with a single form or a single rule applicable to a given context (Skousen 1989: 22). Rather, for any given context there may be a number of possible rules or forms, although not all rules or forms are equally probable. An analogical approach "does not actually use rules to predict behavior. Instead, the analogical approach (as its name suggests) makes predictions in terms of individual occurrences" (22). Frequency of occurrence determines the strength of lexical connections, making one form more probable than another in certain contexts. When a given form (or part of a form) occurs with great frequency in similar semantic and/or phonological contexts, that form is more likely to be chosen (and used) than another form that may seem to fit a particular 'rule'. For example, although an apparent 'rule' such as $/X/ \rightarrow /X + o/$ for the feature set [+ 1 Sg Pres Indic] may seem to dictate a form such as

**so* for [*ser/be*, 1 Sg Pres Indic], the frequency of occurrence of the irregular *soy* (*ser/be*, 1 Sg Pres Indic), in the input preempts the more ‘regular’ – but incorrect – form.

The base alternants for verb forms are connected by semantic connections, and they may or may not be connected phonologically as well. For example, *ten* and *tuv* (two of the stem alternants for *tener* ‘have, possess’) are connected semantically and also slightly connected phonologically, since they share the initial phoneme /t/. On the other hand, *soy* (‘be’, 1 Singular Present Indicative) and *fue* (‘be’, 3 Singular Preterite Indicative) are only connected semantically, not phonologically. They are connected, nonetheless, based on the semantic notion of ‘be’. Derwing and Skousen (1989) contend that

a picture of the mental lexicon has emerged in which (a) words are represented in their full, 'undecomposed' forms and (b) the representations for morphologically related words are 'connected in some way'. Clearly, all this is highly compatible with a general 'analogical network' type of framework, particularly one where morphological variants are connected, with varying degrees of strength, with a central basic or 'root' word (59).

2.7. Lexemes and stems.

Since the initial variables that enter the blocks of Word-Formation Rules in the present proposed model are lexical stems, it is pertinent to consider the difference between a ‘lexeme’ and a ‘stem’. If different stems are possible for a given verb, meaning different allomorphic variants, then are these stem variants stored separately in the lexicon, all with equal status as *signifiants* in their own right, or are they just different

physical realizations of the same abstract lexeme? This view of ‘lexeme’ and ‘stem’ is somewhat parallel to the concepts of ‘phoneme’ and ‘allophone’ in phonology. In this sense, ‘lexeme’ and ‘phoneme’ are both abstract concepts, while ‘allophone’ and ‘stem’ are the mental imprints of the physical representations, or sounds, of those concepts. Spencer (1997) observes that in one sense a fully inflected form is a word itself, but he prefers to call a whole inflected form a ‘word form’. All inflected forms that are related by the same semantic content, such as all the conjugated forms of the verb *hablar*, for instance, are ‘lexemes’ according to Spencer (45), although they are realized in slightly different phonological forms.

Anderson (1995) comments that whether we consider morphology to be based on words or on morphemes depends upon our definition of the word itself. Shall we consider a word to be a full surface form, or is it a base form with no inflectional material? Some languages, such as English, add inflectional material to free forms that can stand on their own, while other languages like Spanish and Latin add inflectional material to base forms that never show up as a word on its own. For example, the base form or stem for *hablaban* (‘speak’, 3 Plural Imperfect Indicative) is /aβl/ (*habl*), which conveys the general meaning of ‘speak’ but never stands alone as a word. Anderson, therefore, comes to the logical conclusion that “it is not words but stems that function as the base of Word Formation Rules” (1995: 71). He then defines the notion of ‘stem’ as “word minus (productive) inflectional affixation” (71). Anderson then goes on to say that these forms (after derivational processes but before inflectional ones) have “a special status in the grammar: essentially, these are the ‘lexical items’ that are entered in a language’s dictionary” (71).

The model proposed in the present analysis regards fully inflected forms as words occurring in the lexicon “in their full, 'undecomposed' forms” with “the representations for morphologically related words... 'connected in some way” (Derwing and Skousen 1989: 59). In the proposed model, however, as in Anderson’s model, it is ‘stems’ – words “minus (productive) inflectional affixation” – that initially enter the blocks of Word-Formation Rules. But, contrary to Anderson’s view, it is whole words – ‘undecomposed forms’ – that are entered into the lexicon.

2.8. Words, semantic features, and grammatical features.

The model of Spanish inflectional verb morphology proposed in the present study views all inflected forms as words in their own right, but also as forms that are related to each other through analogy and through Word-Formation Rules. The proposed model does not in every case specify a stem as a distinct lexical item as defined above by Anderson. Some phonological strings in the proposed model are not lexical items per se, such as /fu/ in the suppletive Preterite model in Figure 6.7 of Chapter 6. Furthermore, some of the phonological material that is added to the X variables in the model are not morphemes in the classical sense either, such as the theme vowels (or class vowels) in Figure 6.1 and 6.2 in Chapter 6, since these theme vowels have no real semantic content.

In some languages it may be descriptively accurate, even obvious, to state simply that a given phonological string X signifies grammatical/morphological function μ and is affixed to the given lexeme in a prescribed manner. However, cross-linguistic evidence shows that real world human languages are not so neatly packaged, as discussed in Section 2.5 above. As Bybee points out, “In the ideal case, all the phonetic material in an

utterance can be assigned to one meaningful unit or another.... But not all linguistic material is organized in this fashion (1988: 128). And if we consider grammatical features as essentially no different from the many other semantic features of a whole, inflected word, as discussed in Sections 2.3 and 2.4 above, then it does not seem so obvious that they should be treated differently in the grammar.

Anderson (1995) views morphology as a process rather than simply the affixation of discrete, separable phonological units, as in the traditional definition of morpheme as the minimal linguistic sign. Anderson views morphemes as ‘processes’ rather than as ‘items’; hence his term ‘a-morphous morphology’, meaning not morphology without form, but morphology without morphemes. Although Anderson maintains that the question of whether morphology is a separate component of the grammar is not particularly relevant to his theory of morphology, he does view morphology as a process occurring in various components of the mental grammar, rather than in a bounded, separate module. For Anderson, part of the process may occur in the lexical component, part in the syntactic component, and part in the phonological component.

The present analysis presents an alternate view, one that follows Anderson’s model in form, but is a morpholexical approach, meaning that word formation, both derivational and inflectional, occurs in the lexical component of the grammar rather than in other components. The model proposed in the present study also treats the grammatical features of a word as essentially a subset of the semantic features in the total feature set of a word, since both grammatical features and semantic features may be difficult to separate into discrete units. As Chialant and Caramazza (1995) explain, for the ‘whole word representation hypothesis’, the part of a lexical form considered

‘morphological’ information is really just part of the semantic information associated with the given word (59). The present analysis also differs from Anderson’s model in that it sees rules as descriptions or patterns, rather than as true rules in a generative sense.

2.9. Conclusion.

This chapter has discussed the nature of the linguistic sign from a morphological viewpoint as a precedent to the model proposed in the present study, since the proposed model is a model of word formation based in form on the link between the two parts of the linguistic sign. The place of morphology in the mental grammar of a language has been considered from various viewpoints, pointing out the various possible ways of treating morphology as a component of the grammar or as a part of other components. We have discussed the form-function relationship, the problem of incongruities in the function-to-form relationship, and the problems in traditional views of morphemes as separate, separable, discrete units of meaning.

Since the proposed model is based on the correspondence between the descriptive Word-Formation Rules that represent a given form (the *signifiant*) and the respective semantic feature sets that represent the function of that form (the *signifié*), this chapter has provided an explanation of the notion of ‘semantic feature sets’ and has discussed the reasons for considering grammatical features as essentially just a subset of the overall semantic features of a whole word.

The notion of ‘lexical stem’ as related to morphology and the lexicon has been examined, since the proposed model depends upon lexical stems as the initial variables entering the model to undergo various Word-Formation Rules. We have established that

the present analysis regards the lexicon as a dynamic, multidimensional component of the grammar composed of whole words rather than just stems or parts of words, and that there are numerous multidimensional semantic and phonological connections among the various forms in the lexicon.

Finally, a brief explanation of Anderson's view of words and morphology ties this chapter's discussion in with the more detailed discussion of Anderson's and Bybee's morphology models presented in the next chapter.

CHAPTER 3

RULES, REPRESENTATIONS, AND MORPHOLOGY MODELS:

REVIEW OF LITERATURE

3.0. This chapter first presents a general background synopsis of morphology and morphemes, including structuralist and generativist views, before discussing Anderson's a-morphous morphology (1995) and his arguments against traditional treatments of morphology. The differences between inflectional and derivational morphology are then considered, along with the rationale for treating only inflectional morphology in the present analysis. An explanation of Aronoff's Word-Formation Rules and Anderson's Extended-Word-and-Paradigm model of morphology is then presented, leading to the subsequent explanation of the present model's Word-Formation Rules. This chapter also explains the ordering of Word-Formation Rules, including conjunctive and disjunctive ordering and the 'Elsewhere principle', as well as discussing lexical stems, stem alternants, lexical stem sets, and paradigms.

Skousen's (1989) analogical model of language is then discussed as it relates to Bybee's (1985, 1988) model of morphology. Bybee's "morphology as lexical organization" and her view of morphology as "a study of the relation between morphology and form" are then discussed, relating Bybee's model to the present morpholexical analysis of Spanish verbs, which also regards morphology as 'lexical organization'. Finally, Marslen-Wilson's (1987, 1989, 1999) 'cohort' theory of lexical processing is discussed as it relates to the present analysis. The theories and morphology models of Anderson, Bybee, and Marslen-Wilson all provide a formal and theoretical

basis for the model of Spanish inflectional verbal morphology proposed by the present study.

3.1. General background: morphology and morphemes.

Beard defines morphology as “the sum of all the phonological means for expressing the relations of the constituents of words, of words in phrases, and of the phrasal constituents of sentences,” (1995: 1). For Anderson, “Morphology is the study of the structure of words, and of the ways in which their structure reflects their relation to other words” (1988b: 146). Anderson contends that “the task of a theory of morphology is to bring order and coherence to our understanding of the way words are composed and related to one another” (1995: 2). Morphology, then, is about the structure of words and about how their structure reflects their relationship to other words, both within phrasal or sentential context and across the entire vocabulary of a language.

According to the above definitions, morphology is about words. But what is a ‘word’? Is a word based on phonological criteria or semantic criteria? Is inflection just phonology? Is inflection just phonological manifestations or changes, with the word maintaining the same underlying representation, with its semantic content basically unchanged? Saussure observed that a spoken word consists two components: 1) the sound-image, or mental imprint, of the word’s phonological realization (the *significant*), and 2) the concept that the sound represents (the *signifié*) (Saussure, trans. 1959: 65-70). The linguistic sign as the basic unit of language is commonly referred to as the ‘Saussurean sign’, although Greek Stoic philosophers centuries earlier had similarly defined the word as “a bilateral association of ‘the signifier’ and ‘the signified’ (Beard,

1995: 2). Beard, following Saussure and the Greeks, defines the word as “a symbol comprising mutually implied sound and meaning” (1995: 1). Although to a linguist the concept of ‘word’ may be difficult to define precisely, the word is “something central in the mechanism of language” (Saussure, trans. 1959: 111), and speakers of human languages seem to ‘know’ instinctively what a word is. This somewhat obvious observation was reinforced for the present author by a student in an introductory linguistics course several years ago. After a lengthy class discussion of the concept of ‘word’ and the difficulty for linguists in defining the ‘word’, the aforementioned student impatiently commented at the end of class, “Get a grip. You linguists are off in space somewhere. Don’t you know that even a two-year-old knows what a ‘word’ is?”

Saussure comments appropriately that “in the matter of language, people have always been satisfied with ill-defined units” (111). But even if we settle for an “ill-defined” notion of ‘word’, there remains the question of whether the word is the minimal unit of linguistic meaning. Is the word itself the basic linguistic sign, or is the minimal unit of meaning something smaller than the word, such as parts of words? Word makeup is the basic problem of morphology, and the basic problem of linguistics in general is how we get from the meaning to the symbol, especially if either the meaning or the sound is very complex or if there is not a one-to-one correspondence between the meaning and the sound. Individual words are made up of individual sounds, or phonemes, but how do speakers of a language group these phonemes together to signify an idea? At what point is it determined that a given combination of phonemes has a meaning, and is thus a minimal sign?

The word itself as a minimal sign is the basic premise of Saussure's view of language as a semiotic system. Although Saussure had considered the word to be the minimal sign, structural linguists of the 1940's and 1950's accepted parts of words as the minimal units of meaning; hence the morpheme came to be regarded as the minimal linguistic sign. To the structuralists, morphemes are "the smallest meaningful units" (Bloomfield 1933: 166), and a morpheme, more specifically, is phonological material bearing linguistic meaning. Bloomfield's "notion of 'morpheme' assumes a determinate phonological content" (Anderson 1988b: 152), although later structuralists viewed morphemes as abstract, regarding morphemes as related to their phonological realizations in the same way that phonemes are related to their respective phonological realizations. In the latter view, the variation or alternation of phonological realizations of the abstract morpheme is determined by phonological rules that act upon the morpheme.

Words are seen by structuralists as composed of abstract minimal units of meaning with concrete representations composed of strings of phonemes, which themselves are abstract representations of the actual sounds as they are realized in spoken language. The [im] in the English word *impossible*, for instance, represents negation, and although it is realized with a bilabial nasal in this word, it may be realized as [iŋ], with a velar nasal, in a word such as *incorruptible*. Likewise, the *-able* of *impermeable* is the phonological realization of the same *signifié* as the *-ible* in *incorruptible*, although, at least in a synchronic sense, these two latter words differ only orthographically, not phonetically. (The /a/ and the /i/ are both realized phonetically as an unstressed mid-central lax vowel.) Should we view different realizations such as [im] and [iŋ], or *-able* and *-ible*, as different morphemes? Are they different *signifiants* with the same referent,

or are they just phonetic (or orthographic) variations of the same *signifiant*, with both alternants having the same referent?

These phonological variants are traditionally referred to as ‘allomorphs’, just as phonological variants are referred to as ‘allophones’. So, then, shall the morpheme be considered an abstract concept? Or is a morpheme a string of phonological material rather than the concept that string represents? In general Saussurean terms, a morpheme would be both, since the Saussurean minimal sign consists of two elements as described above: its *signifié* and its *signifiant*. One aspect of traditional morphology is the study of allomorphy, which is concerned with the relationship between morphemes as abstract signs and the phonological realizations of those abstract signs. The other aspect of traditional morphology is morphotactics, which is concerned with how morphemes can be combined to form larger, more complex words. In other words, morphotactics concerns the arrangement of morphological material within a word and “the ways in which morphemes can be arranged into larger, hierarchically organized structures to construct complex words” (Anderson 1995: 14).

The structuralist view of morphology was adopted by generative linguists with the inception of generative linguistics in the 1950’s. The structuralist view of morphology dominated the early generativist view of word structure, although generativists like Chomsky defiantly departed from structuralist views in their overall view of language. In fact, according to Skousen (1989), “Chomsky himself is a structuralist par excellence,” since “For Chomsky (and virtually all other linguists today) there is no doubt that language is rule governed and that language behavior must be accounted for in terms of explicit rules” (4). Structuralism, as Skousen explains it, is “the traditional method for

describing behavior... in many diverse fields – from biological taxonomy to literary criticism” (3). Structuralism is “a system of classification,” determining “how a general contextual space should be partitioned” (3). Saussure, an early proponent of structuralism in the field of linguistics, defines language as “a self-contained whole and principle of classification” (1959: 9). Anderson describes the structuralist approach to language as one in which the investigator “was supposed to collect a large corpus of utterances, and then identify the similarities among them, as well as determine which utterances ‘contrast’ with one another” (1988b: 151). Although generative linguistics is more concerned with explaining how language works as an internal system rather than with the empirical description of language as it is actually spoken, generative linguistics, like structuralism, is also based on systems and rules.

However, although generative linguistics adopted the overall structuralist view of morphology, early generative linguists, unlike the structuralists, did not see morphology as basic to the study of language. Generativists came to regard morphology not as a separate field of study, but rather as a part of other fields. They treated allomorphy in the realm of phonology, and morphotactics in the realm of syntax. Anderson says that in early generativist theories “morphotactics was taken to be just a (not particularly well-defined) subpart of syntax within the grammar of a natural language” (1995: 16). Generative phonology views a given physical morpheme as having a basic underlying form, just as speech sounds have phonemes as underlying representations. In the generativist view, the phonological rules of a language operate on the underlying representation to produce morphemic variants, or allomorphs. Anderson, however, continues in the generativist tradition but comes to the conclusion that “allomorphy

cannot be reduced to the operation of rules by the phonology” (42). He points out that “When the operation of a given rule is conditioned by morphological factors, and refers to the particular morphemes (or classes of morphemes),” that operation does not “conform exactly to the same principles as one which is conditioned entirely by sound structure” (42). There is a need, then, according to Anderson, to consider morphology as a realm of study in its own right.

Syntacticians have had “sharp regard for the difference between domains internal and external to the word” (Anderson 1995: 47). Chomsky’s Lexicalist Hypothesis maintains that words are indivisible from the viewpoint of the syntax, that the syntax cannot operate at the word level (Jackendoff 1972, Chomsky 1970), and that words come fully formed from the lexicon. According to the Lexicalist Hypothesis, the properties of individual lexical items are not available to syntactic operations, since, according to Chomsky, lexical insertion takes place at S-structure (surface level) rather than at D-structure (deep structure level).

Syntacticians thus came to accept that rules distinct from those of the syntax operate in the lexicon to describe words. Thus, just as there is a need for morphology as a distinct realm of inquiry in linguistics, there is a need for a theory of lexical rules that form or describe words. Anderson argues that the rules of a given language “may divide into two distinct subsets, one of which is responsible for the composition of words and the other for phrases” (37). It is the former of these two subsets that Anderson is concerned with, and it is the inflectional subset of these rules that the present analysis deals with. If rules that are distinct from those of the syntax operate in the lexicon to describe words, then we need a theory of lexical rules, or, in Aronoff’s and Anderson’s

terms, we need a theory of Word-Formation Rules. Thus Anderson believes that the development of a theory of morphological structure is an important, distinct object of inquiry within linguistics.

Beard comments that “in the past few decades morphology has all but vanished from the agenda of linguistic inquiry,” (1995: 1). Although early generativists essentially ignored morphology as a relevant field of study, morphology is the foundation of traditional linguistics. The modern disciplines of syntax and phonology arose from the study of word structure, which traditional grammarians viewed as the central focus of the study of how language works. Word structure is “the product of interacting principles from many parts of the grammar: at least phonology, syntax, and semantics, in addition to the ‘lexicon’” (2). Morphology, then, at least in Anderson’s view, may not be localized, located in its own separate component of the grammar, but may actually be represented in various components of the grammar, and may be different for inflectional vs. derivational morphology.

Although Anderson maintains that morphology is indeed a legitimate object of study in linguistics, he regards as a distinct theoretical issue in itself the question of whether morphology is a separate ‘component’ of a grammar (8). Like Anderson, the present analysis of Spanish inflectional verbal morphology accepts the position that morphology is indeed a valid and important aspect of language, but does not make the claim that morphology constitutes a separate component of the mental grammar. The present investigation and proposed model also concur with Anderson that “morphological theory ought to have a rather different character than is commonly assumed” (9). The

proposed model, like Anderson, views “morphology as the study of the relations between words, rather than as the study of discrete minimal signs” (1).

The present morphological analysis of Spanish verbs, like Anderson’s model, is concerned with “words as grammatical, rather than as phonological, entities” (9), and also with words as semantic entities. The present study examines words as Saussurean minimal signs and tries to determine whether words are composed of several linguistic signs, each as a minimal unit of meaning as seen in the traditional view of morphology, or whether words can or should be seen as composed of a single linguistic sign, as in Saussure (trans. 1959) and Aronoff (1976, 1992). Like Anderson’s Extended Word-and-Paradigm (EWP) model of morphology, the present model is a structured system of morphological rules that eliminates word-internal boundaries. According to Anderson’s ‘a-morphous’ morphology (1995), the notion of words as comprising constituent minimal signs is not valid and is flawed for a number of reasons, which are discussed in the next section below.

3.2 Anderson’s a-morphous morphology.

Word structure according to early generativist grammar encompassed the premise that: 1) the morpheme is the minimal sign (as in the structuralist view); 2) allomorphy belongs to the phonological domain; 3) morphotactics belongs to the syntactic domain. Hence, there was no theory of morphology in early generativist grammar, and no need for a distinctly morphological study of allomorphy or of morphotactics. Early generativist grammar, therefore, seemed to recognize the ‘morpheme’, but not ‘morphology’.

Anderson (1995), however, takes the opposite stance: he recognizes morphology but not morphemes, hence ‘amorphous’ morphology, meaning not ‘morphology without form’, but rather ‘morphology without morphemes’. Anderson contends that “...the morpheme itself (at least as classically construed) is not an appropriate element to serve as the foundation for a general theory of word structure in natural language” (17). According to Anderson, words as wholes, not morphemes, ought to be the objects of inquiry in morphology. And so, along with Anderson, admitting that although we “cannot claim to have a full theory of what ‘words’ are... we can proceed as if we knew how to delimit words, as indeed we generally do” (21), and as a two-year-old generally does, as the above mentioned student pointed out.

In structuralist terms, the morpheme is the elemental unit of morphological analysis – ‘a linguistic form which bears no partial phonetic-semantic resemblance to any other form’ (Bloomfield 1933: 161; Anderson 1995: 49). In other words, according to the structuralist definition of the morpheme, a word cannot be divided further without destroying the correlation between the phonetic symbol and the semantic content, the two parts of the Saussurean linguistic sign. Even upon casual observation of a language, it is obvious that many words across the language, especially morphologically complex words, whether inflectionally complex or derivationally complex, show partial phonetic and semantic resemblances to each other. Thus, a theory of word structure must account for this observation. Bybee’s (1985, 1988) analogical model of morphology, as well as Anderson’s a-morphous morphology, also takes into account this relatedness among the words of a language, as will be seen more clearly below.

In the classical sense, a morpheme was to be “ ‘a minimal same of form and meaning’ –an indivisible stretch of phonetic (or phonological) material with a unitary meaning” (Anderson 1995: 49). This may seem consistent with the Saussurean sign, but Saussure himself considered the notion of the linguistic sign to apply to the word or complex form, not to simple forms, parts of words, or morphemes. Bloomfield considered a ‘morpheme’ to have “determinate phonological content” (Anderson 1995: 49), as Bloomfield’s statement that “Any morpheme can be fully described (apart from its meaning) as a set of one or more phonemes in a certain arrangement” (Bloomfield 1933: 167). In practice, however, Bloomfield seemed to consider allomorphs to represent the same morpheme, and this view is closer to the later definition of the morpheme. He states, for example, that “Strictly speaking, we should say that the morpheme... has two (or, sometimes, more) different phonetic forms, ... and that each of these alternants appears under certain conditions” (164).

According to Anderson, the post-Bloomfield definition of allomorph is “an abstract name for a class of ‘morpheme alternants,’” (Anderson 1995: 50). In other words, allomorphs are different phonological forms for the same semantic meaning, and these forms are in non-contrastive, or complementary, distribution. Thus, a ‘morpheme’ is an abstract concept or a class of alternants, just as a phoneme is an abstract concept realized as a set of phonetic alternants. A ‘morph’, on the other hand, is the actual phonological form; and an ‘allomorph’ is one of the alternant morphs corresponding to the same morpheme. These three terms – morpheme, morph, and allomorph – are roughly parallel to the concepts of phoneme, phone, and allophone, respectively.

As Anderson (1995: 50) points out, the basic properties of the classical morpheme were the following:

- 1) Morphemes are homogenous and indivisible atomic units of linguistic form.
- 2) Each morpheme in a given word is phonologically represented by exactly one morph, and each morph represents exactly one morpheme; i.e., there is a one-to-one correspondence between form and meaning.
- 3) The morphs themselves are consistently and uniquely (though not necessarily biuniquely) related to surface phonetic form.
- 4) The morphemes are arranged into a hierarchy that accounts for the word's internal structure.
- 5) Words are exhaustively composed of morphemes.

Morphology, then, is “a set of statements about how these abstract elements are distributed with respect to one another and organized into Immediate-Constituent structures (the morphotactics); and about how each is realized, in terms of its morphological and/or phonological environment (the statements about allomorphy)” (Anderson 1995: 50-51). Agglutinating languages – or agglutinative cases within a given language – are the neatest and easiest to fit into the above definitions of morpheme and morphology, but not all languages and words fit so easily. In Spanish, for example, although derivational morphology may be reasonably neat and predictable, inflectional morphology is not so easily analyzed and does not exactly fit the second of the above listed basic properties of the classical morpheme.

If we accept the above classical view, we are presented with several problems with the relation between meaning and form. In reality, there is not a one-to-one correspondence between meaning and form in most cases. There is not simply a series of separate and separable forms, with each representing a separate meaning. Spanish, for instance, has cases of empty morphs, such as Spanish verb thematic vowels, for which there is no link to meaning. Verbs may be *a*-class, *e*-class, or *i*-class, but these thematic vowels serve only as a sort of “morphological glue” (Anderson 1995: 54). Spanish also exhibits superfluous morphs, which are similar to empty morphs, except that they have a meaning. This meaning, however, is irrelevant or incompatible, as in Spanish adverbs with */mente/*. The suffix */mente/* is attached to the feminine form of the given adjective, although this feminine element – the */a/* in *rápidamente*, for instance – has no relevance in the synchronic semantic value of the word. In the historical development of the word, however, the feminine adjective ending had to agree with */mente/*, meaning ‘mind’ (Lathrop 1980: 153-154).

Cumulative morphs are another problem for a language such as Spanish. In the case of cumulative morphs, there may be a single formative that represents more than one semantic function, with the ratio of form to function ‘one to many’ instead of ‘one to one’. Multiple *signifiés* correspond to the same *signifiant*, as in Spanish Preterite forms, in which the form */ó/* corresponds to the notion of 3 Singular as well as Past tense. The morphs are not easily separated to correspond on a one-to-one basis with their corresponding semantic values representing person, number, tense, mood, and aspect. Thus we cannot say that these morphs are ‘minimal signs’, at least from the viewpoint of the *signifiant* half of the linguistic sign. Another case of a one-to-many ratio between

form and function is syncretism, for which a single formative may represent different semantic values in different contexts, that is, in different parts of the overall inflectional paradigm. In Spanish, the form *canta*, for example, represents both 3 Singular Present Indicative and Singular Familiar Imperative; likewise, a form such as *cantamos* can be either 1 Plural Present Indicative or 1 Plural Preterite.

There is also the problem of overlapping ('portmanteau') morphs (Anderson 1995: 54), or 'multiple exponence' (Spencer 1997: 51), for which the ratio of form to meaning is 'many to one' instead of one to one. In this case, more than one form corresponds to one meaning, as in English *written*, where the stem vowel change – from /aI/ to /I/ – as well as the /en/ ending signifies Past Participle. Similarly, in a Spanish verb form such as *supieron*, the /sup/ segment as well as the /yero/ segment signifies preterite function. Reciprocal conditioning, similar to portmanteau morphs above, is a problem as well. In the case of reciprocal conditioning, two components of a given form simultaneously depend upon each other to represent two or more *signifiants*, hence a form-to-function ratio of 'many to many', and therefore also inconsistent with the strict traditional concept of morpheme. The above example of *supieron* might be considered a case of reciprocal conditioning, as well as overlapping morphs or multiple exponence, since the /fu/ and the /eron/ parts of the word depend on each other to signify Past tense (Preterite Indicative, more specifically), but the /eron/ part of the word also signifies number and person. This case of a many-to-many ratio between form and function is inconsistent with the strict traditional concept of the morpheme, as are the cases of one-to-many and many-to-one ratios as discussed above.

Anderson contends that all of the above “suggests that the principles underlying the structuralist notion of the morpheme must be at least revised, if not abandoned” (1995: 56). McCarthy (1981) also proposed revising the notion of what constitutes the morphological analysis of a form, since the classical view sees morphemes as units of phonological material that are concatenated to create full word forms. McCarthy proposed that morphological and phonological analyses be viewed as coordinate but independent. McCarthy’s model is known as ‘non-concatenative morphology’, as opposed to purely affixational (concatenative) morphology. McCarthy’s non-concatenative morphology considers the relationship between morphological and phonological content as “a matter of associations within and among the planes of a complex representation (rather than as the concatenation of phonologically instantiated morphological units)” (Anderson 1995: 58-59).

Spencer (1997: 49-53) outlines the three structuralist models of morphology: Item-and-Arrangement (IA), Item-and-Process (IP), and Word-and-Paradigm (WP). As evidenced so far in this discussion, Anderson’s theory of morphology is not an item-and-arrangement (IA) approach to morphology. IA approaches emphasize word analysis rather than word formation and “sought to provide techniques for decomposing words into their component morphemes” (Spencer 1997: 49). Anderson’s theory of morphology essentially falls into the latter of the above three categories, Word-and-Paradigm, since this is the one of the three that recognizes, like Anderson, that there is often a many-to-many correspondence between morphological form and morphosyntactic function. A Word-and-Paradigm theory of morphology contends that there are certain generalizations about words and their relationships among each other that can only be explained at the

level of the whole word rather than at the level of its alleged constituent morphemes. A WP model recognizes the role of inflectional paradigms in the organization of a grammar. According to Spencer, the “key to the WP approach is our notion of the morphosyntactic word. Each inflected form has (at least) one morphosyntactic description (for example ‘past tense form’ or ‘dative singular of the masculine/neuter adjectival form’) and the grammar then makes available paradigms that specify the formatives which correspond to these categories” (Spencer 1997: 52). In a primarily agglutinating language the rules of correspondence between form and function may be relatively simple, with one morphosyntactic category for each formative and one formative for each morphosyntactic category. However, there is no inherent need for such simplicity or for such one-to-one correspondence, and, indeed, many languages are far more complex. A WP approach easily handles such problems as those described above, such as syncretism and overlapping exponence. The rules of a grammar can be written in such a way as to account for such phenomena, since “the morphosyntactic description is separated from the morphological formatives as such” (Spencer 1997: 52), given that the rules and representations are based on whole words rather than on morphemes.

Although Anderson’s theory of morphology is different from Item-and-Arrangement and Item-and-Process approaches in that it doesn’t recognize the morpheme per se and because it fits reasonably well into the Word-and-Paradigm category above, it is still essentially an Item-and-Process theory of morphology, because it is about morphology as a process rather than about morphemes as things. A process view “adds the relevant phonological material to a more basic stem” (Anderson 1995: 63). Anderson (1995) argues the following:

If we accept the evidence that the range of morphological possibilities in natural languages includes some processes that cannot properly be represented as the addition of an affix, we must conclude that a general morphological theory should admit both affixational and non-affixational rules. Since a process-based approach naturally accommodates affixation, but not vice-versa, the alternative we should prefer is to explore a theory of morphological processes (68).

Thus Anderson's model, although it is a Word-and-Paradigm model of morphology, especially as regards inflection, is also essentially an Item-and-Process model in that it stresses morphological processes, as opposed to lists of morphemes in the traditional sense. What is most significant here is that it is not an Item-and-Arrangement model, as is the structuralist view of word structure.

“If morphology is indeed more adequately represented in the general case by relations or processes than by discrete lexical-item affixes” (Anderson 1995: 69), as in the traditional view of the morpheme, then there are broad implications for the grammars of natural languages. “Instead of a lexicon of affixes, the morphology of a language would then consist of a set of rules, each describing some modification of existing forms that would relate them to other forms”(Anderson 1995: 69). For Anderson, morphological changes and phonological changes are nothing more than rules that modify the shape of a form. In this view, morphology is a set of rules rather than an inventory of lexical items.

The model proposed in the present study assumes such an a-morphous, or morpheme-less, approach to morphology, as will be seen more clearly when the model is

fully presented. In the proposed model, for instance, segments of phonological material are added onto the variable or the stem, but these segments do not by themselves necessarily correspond neatly to any semantic content, as we will see. The present model also concurs with Aronoff's (1976) and Bybee's (1985, 1988) word-based morphology, in which the relation between the two parts of the Saussurean linguistic sign is based on whole words rather than on morphemes, since as Anderson (1995) states, "... it is only at the level of whole words that form is associated with meaning" (69). Actually, both the forms and the meanings of words could be divided into smaller segments, such as dividing the semantic content of the word *boy* into the components 'young male human being'. However, although the classical morpheme assumes that there is a one-to-one correspondence between categories of meaning and aspects of form, this assumption is rarely the case in natural human languages.

3.3. What is inflection?

Aronoff (1976, 1992) concludes that if the basic linguistic sign is the word and not the morpheme, then morphological rules are relations between words, and these morphological rules (meaning the Word-Formation Rules) specify changes in properties that relate one word to another. A change in the form corresponds to some change in the syntax and/or some change in the semantics of the word, such as a different subject or a change from singular to plural.

But what is a word? Is it an existing surface word or something else entirely? What about cases where the base form never shows up as a form itself, as with Spanish verbs? If /aβl/ is the base form for *hablar*, then this notion is inconsistent with Aronoff's

claim that the basic sign is the whole word rather than the morpheme, since the stem /aβl/ is the base form of this verb and carries the semantic value of ‘speak’. According to Anderson (1995), Aronoff’s contention that the lexicon is ‘word-based’ rather than ‘morpheme-based’ only makes sense if the ‘words’ that feed into the Word-Formation Rules are “surface words minus certain material” (75), or what we have earlier referred to as lexical items in Anderson’s terms. Derived words, on the other hand, are made up of stems that do not contain such surface inflectional material. Inflectional morphology is physically peripheral to derivational morphology, occurring in the last level of word formation. Inflectional processes are applied to words only after all derivational affixes are added. Inflectional material or processes occur outside of derivational material or processes; in other words, inflectional prefixes precede derivational ones, and inflectional suffixes follow the derivational ones.

Anderson’s hypothesis regarding inflection and derivational processes is that “*Inflection is outside of derivation*. Material introduced into a form *F* as a consequence of the morphosyntactic properties *M* of the position that *F* interprets presupposes the prior application of all morphological processes involved in the derivation of the stem *S* on which *F* is based, and not vice versa” (Anderson 1995: 126). For example, Spanish speakers don’t say **hablamosante*, with the inflected *hablamos* as the stem, or **hablanante*, with the inflected *hablan* as the stem, but rather *hablante*, from the uninflected stem /abl/ (*speak*) plus the derivational affix /ante/ (agentive, meaning ‘one who does something’). In order to make this new derived word plural, we would add /s/ after the derivational processes have taken place, yielding not something like **hablsante*, but rather *hablantes*.

In Anderson's view, the 'lexical items' that are entered into the mental dictionary of a language are "the class of (possibly complex) forms representing the output of rules of derivation but prior to the operation of productive inflection" (Anderson 1995: 71). In other words, Anderson's lexicon is made up of words for which derivational processes (if any) have already taken place, but before inflectional processes occur. Inflectional rules are distinct from derivational rules in that derivational rules form lexical items, while inflectional rules convert these lexical items into inflected surface words. Thus Anderson modifies Aronoff's word-as-morpheme proposal by arguing that "it is not words but stems that function as the base of Word-Formation Rules" and that a stem is a "word minus (productive) inflectional affixation" (Anderson 1995: 71). Of course, as noted above, a given stem may already be morphologically complex in that rules of derivation may have already applied to form a new stem from some other stem.

However, although Anderson distinguishes between derivational and inflectional morphology as above, he himself admits that the two are not easily defined. Definitions are often circular, as in stating that inflection is grammatical material (as opposed to meaning-changing or word-class-changing material) that is added to a lexical stem, while, conversely, a stem is a surface word minus inflectional material. Anderson points out that "the traditional literature usually relies on a few examples to establish the difference between inflection and other sorts of morphology rather than an explicit definition" (1995: 77). In general, inflection doesn't change word class, nor does it change the basic meaning of the stem. Derivational processes change either the word class, as in *hablar* 'speak' vs. *hablante* 'speaker', or they change the basic meaning, as in *ventaja* 'advantage' vs. *desventaja* 'disadvantage'. Thus, 'semantic neutrality' could be

seen as a distinguishing factor of inflection. Although Anderson points out that even plural inflections change the meaning to a certain degree, since *dog* and *dogs* are not the same (79). Productivity is an “often cited basis for distinguishing inflection” (77), with inflectional processes noted as more productive than derivational ones. Anderson points out that the productivity criterion does not always hold up either, since some inflections in some languages are not very productive, such as the /en/ plural ending of English *oxen*. In addition, some derivational processes are indeed quite productive, as in the nominalizing /ing/ in English, as in ‘UFO sightings’. A further observation is that “words differing only in their inflection group together into paradigms” (79). But Anderson adds that “this is of little help in identifying ‘inflection’ unless we know when words should be said to belong to the same paradigm,” and the answer to the latter question is “when they differ only in inflection” (1995: 79).

Bybee (1985) mentions ‘predictability of meaning’ as a criterion for inflection (verbal inflection, at least) in that “The meaning of the category must be predictable with every verb,” adding that “If the grammar lists the meanings of the combinations with different stems, then the category is not inflectional” (27). This presumably might refer to derivational affixes such as /in/, which may have a negating meaning as in *incapacitar* ‘incapacitate’ or a directional (‘into’) meaning as in *ingresar* ‘to deposit; to enter/come in’ or *inculcar* ‘inculcate’. Verbal inflectional affixes, on the other hand, are essentially predictable in meaning, as in the 1 Singular Present Indicative ending /o/. Although verbal inflections may have a broad range of semantic possibilities and differing shades of meaning, as discussed further in Chapter 4 of the present analysis, the different

meanings for a given verb tense are all related to the same general sense of time, unlike the difference between the two meanings of */in/* in the examples given above.

Inflection seems to be the kind of morphology that is manipulated by the rules of syntax, but this goes against the notion in modern syntactic theory that the syntax cannot manipulate the internal structure of words. The Lexicalist Hypothesis from the generativist framework, as stated previously, maintains that the syntax neither manipulates nor has access to the internal form of words (Anderson 1988: 10; Spencer 1997: 72-73), although early generativists did regard syntax as capable of acting at the word level (Anderson 1995: 16). But Anderson contends that inflectional morphology does seem to be manipulated by the syntax. To Anderson, his theory of inflection is a theory about the apparent exceptions to the Lexicalist Hypothesis, since “inflection is precisely that morphology with respect to which principles of syntax and of word formation interact with particular grammars” (85). Anderson’s definition of inflectional morphology is “the area in which principles of syntactic structure and of word formation interact with one another; ... a theory of inflection is in some sense a theory of the ways in which the maximally strong form of the Lexicalist Hypothesis must be relaxed” (100-101).

For Anderson, then, inflection is where morphology and syntax overlap. But what information is exchanged between morphology and syntax? Anderson (1995) concludes that “there must be an analysis of at least some properties of words which is accessible to the syntax” (85). Anderson suggests, for example, that “we can allow the subcategorization requirements of individual lexical items to play the role of expressing

dependencies between ‘case assigners’ and the specific cases they require on some of their arguments” (120).

We could use a similar subcategorization designation in the lexicon to assign certain stem alternants to be used for certain tenses and even to specify verb class based on theme vowel. Anderson cites “devices invoked by other theories, such as allowing ‘morphemes’ to subcategorize for the particular affix (or class of affixes) which they can follow within a word” (124). For example, *traer* ‘bring’ would be specified (or subcategorized) as *e/i*-class, with the following stem alternants specified (subcategorized) for the given contexts:

traig / [Pres Subj] [1 Sg Pres Indic]

traj / [Pret Indic] [Imperf Subj]

tra / elsewhere

Bybee (1985, 1988), unlike Anderson, regards both derivational and inflectional morphology as occurring in the lexicon, since she accepts a whole-word view of the lexicon. The ‘whole word representation hypothesis’ describes the lexicon as “a full list of the lexical forms previously encountered – whether they are morphologically simple or complex” (Chialant and Caramazza 1995: 59). If both processes are part of the lexical component, then the above discussion about the ambiguity of inflection vs. derivation is not as important as it would be if derivation were in one component and inflection in another. Bybee regards inflection and derivation not as separate categories, but as part of a continuum, with full lexical expression on one end, syntactic expression at the other end, and derivation and inflection in between. Bybee explains lexical expression as “the

most highly fused means of expression,” and syntactic or periphrastic expression as “the most loosely joined” means (1985: 12), as shown in the following continuum:

lexical --- derivational --- inflectional --- free grammatical --- syntactic
< -----
greater degree of fusion

For example, while one language might use a completely different word to express ‘found out’ as opposed to ‘knew’, for example, another language might just make an aspectual or inchoative distinction of the same verb, as in Spanish *supo*, which essentially means ‘found out’ but is just a Preterite form of the verb *saber* ‘know’. Bybee (1985) maintains that

there are two factors that determine the likelihood that a semantic notion will be encoded as an inflectional category. First, the semantic notion must be relevant to the meaning of the stem to which it attaches. Second, it must be a very generally applicable semantic notion, or it simply will not apply to enough different items to be inflectional (19).

In addition, Bybee observes from cross-linguistic evidence that the point at which certain semantic/morphological distinctions happen to be on this continuum is specific to a given language. She identifies nine morphological categories that may occur at varying language-specific points on this continuum: valence, voice, aspect, tense, mood, number, person, person (object), gender. Furthermore, she found in the cross-linguistic evidence that the “degree of relevance in general” predicts the “order of occurrence of morphemes with respect to a root or stem” (33). This observation is consistent with the general notion that derivational morphemes occur closer to the stem than do inflectional morphemes.

All of the above categories except valence, voice, person (object), and gender are encoded in what is traditionally regarded as inflectional morphology in Spanish. The present analysis chooses somewhat arbitrarily to deal with Spanish inflectional morphology to the exclusion of derivational morphology, based on the traditional and somewhat intuitive notion of what is indeed ‘inflectional’. However, the decision to deal only with inflection in the present analysis does not necessarily preclude the possibility of a similar set of Word-Formation Rules for derivational morphological processes.

It is interesting and relevant that Alvar and Pottier (1983) point out that an Infinitive “no es estrictamente verbo” (‘is not a verb in the strict sense’) because “no tiene tiempo ni persona” (‘it has neither tense nor person’) (215). In fact, since the Infinitive lacks in verbal specification, it functions as a noun. Past and Present Participles also lack tense and person, but they function as an adjectives and adverbs, respectively. In this sense, then, Infinitives and Participles might be better regarded in the realm of derivational morphology, although the present analysis does include them in its Word-Formation Rules. Alvar and Pottier’s observation about Infinitives and Participles is evidence of the lack of a clear distinction between inflectional and derivational morphology, and their observation points out the somewhat arbitrary nature of the decision regarding what is included in the present analysis.

3.4. Word-Formation Rules.

Anderson proposes that “the morphology of a language consists of a set of Word Formation Rules [WFR’s] which operate on lexical stems to produce other lexical stems (which, if fully inflected, will be surface words)” (Anderson 1995: 71). In other words, in

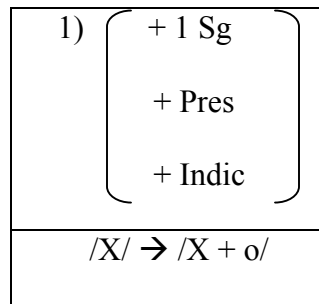
Anderson's theory of inflection, there is an organized system of rules that determine the phonological realization of morphological properties. Lexical stems + WFR's = other lexical stems + inflection = surface words. Morphology is a set of rules by which a word is related to other words and to a basic lexical entry. Although Word-Formation Rules may create semantically new lexical items, or new stems, by derivational processes, the Word-Formation Rules we are concerned with in the present study are not those that form lexical items. The present study is concerned with Word-Formation Rules that create fully inflected surface words from a given lexical item or stem. Anderson regards the operation of inflectional Word-Formation Rules as part of the process of lexical interpretation, but he maintains that these inflectional Word-Formation Rules must take place after the lexical interpretation of the stem, and after any derivational processes are applied to the stem.

According to Anderson's model, the grammar of a given language includes a set of inflectional Word-Formation Rules that "operate to map lexical words (actually, lexical stems) onto fully inflected surface words" (1995: 122). A given set of rules takes as its input from the lexicon a stem (the *signifiant* half of the Saussurean linguistic sign), along with its associated set of semantic features (the *signifié* half of the sign), including morphosyntactic information that is to be realized by an inflected form of the given stem. Each WFR operates on the stem to form a new stem reflecting the phonological result of application of the given rule. The phonological result of application of a given rule or rules may be the simple addition of affixal material, as in the addition of /s/ for the formation of an English 3 Singular Present Indicative form, or the more complicated addition of several phonological strings, which may or may not represent a one-to-one

ratio between form and function. In addition to suffixation, which is typical of a language such as Spanish, the phonological result of application of a rule may alternatively be prefixation or even infixation in some languages, or it may even be some other phonological change such as metathesis, substitution, deletion, or reduplication.

The structural description of a given rule includes conditions on the stem as well as conditions on the affixes or other inflectional changes involving the part of the fully inflected word that we traditionally call grammatical morphemes. For example, a given rule might apply to verbs interpreting positions where a [+ Preterite] form is required. “Each rule may be regarded as a sort of generalization of the notion of a ‘morpheme,’ whose form (or *signifiant*) corresponds to the rule’s Structural Change, and whose content (or *signifié*) corresponds to its Structural Description” (Anderson 1995: 123). These two aspects of the rule correspond to the bottom and top halves, respectively, of the verb model in Anderson’s model, as well as in the model in the present analysis as well. For example, in Anderson’s terms, for a rule such as the one shown in Figure 3.1, the bottom half represents the form (the *signifiant*), or the rule’s structural change, while the top half, on the other hand, represents the content (the *signifié*), or the rule’s structural description.

FIGURE 3.1. Sample Word-Formation Rule, representing a 1 Singular Present Indicative Spanish verb form such as *hablo, canto, bebo, vivo*.



Anderson argues that the traditional view of the morpheme is not adequate when there are aspects of a word's formal structure that cannot be represented as concrete morphemes in the traditional sense. Spanish Preterite forms, for instance, are difficult to divide into concrete morphemes, as Bull (1965: 118) shows in his three-part verb analysis. Bull (112-120) shows that in most cases Spanish verbs can be neatly divided into three parts in the following order: 1) the stem, which carries the semantic content of the verb, 2) the tense/mood/aspect markers, and 3) the person/number markers. The verbs *hablar, cantar, beber, vivir*, for example, can be divided as shown in Figure 3.2.

FIGURE 3.2. Bull's (1965) three-part verb analysis.

Stem (lexical symbol)	Tense/Mode/Aspect (TMA)	Person/Number (P/N)	Morphosyntactic information
Habl	a	s	Pres Indic 2 Sg
Cant	a	mos	Pres Indic 1 Pl/ Pret Indic 1 Pl
Habl	aba	n	Imperf Indic 3Pl
Habl	ara	s	Imperf Subj 2Sg
Beb	ía	is	Imperf Indic 2Pl
Viv	i	mos	Pres Indic 1 Pl/ Pret Indic 1 Pl

Spanish Preterite forms, however, cannot be so neatly divided as the above cases can. Bull attributes this irregularity in the system to this observation: “During the evolutionary phases through which Latin became Spanish, the Preterite did not undergo the regularization process exhibited by the other tense forms” and thus “the tense-mode-aspect slot is a mixture of debris left over from Latin” (118). The relevant point here is that for most Preterite forms the stem can be isolated, but we cannot easily separate the PN markers from the TMA markers. Consider the three Preterite paradigms shown in Figure 3.3, for example.

FIGURE 3.3. Preterite forms for three Spanish verbs.

Hablar		Beber		Vivir	
Singular	Plural	Singular	Plural	Singular	Plural
Hablé	Hablamos	Bebí	Bebimos	Viví	Vivimos
Hablaste	Hablasteis	Bebiste	Bebisteis	Viviste	Vivisteis
Habló	Hablaron	Bebió	Bebieron	Vivió	Vivieron

The dilemma here is that the /é/ and the /ó/, for instance, seem to be a TMA marker and a PN marker at the same time, as do the 2 Singular endings with /ste/. We still see the familiar 1 Plural /mos/ and the 3 Plural /n/, but not the familiar 2 Singular /s/, which occurs in virtually every other 2 Singular verb form in Spanish other than those of the Preterite. (There seems to be a powerful force of regularization at work in some varieties of Spanish, however, since the typical 2 Singular /s/ prevails in these varieties, yielding forms such as **hablastes*, **hicistes*, and **bebistes*.) Thus, since most of the Preterite forms are cases of multiple exponence, in which one formative represents several functions, there are indeed aspects of the formal structure of Preterite forms that cannot be represented as concrete morphemes in the traditional sense.

In exploring new alternatives to the traditional view of morphemes and morphology, Anderson cautions that we must at the same time “account (at least) for all of the things that the concept of the morpheme and a morpheme-based morphology were designed to cover” (1995: 71). There are two basic notions underlying the morphemic account of word structure: 1) Words that are related in meaning are often related in form as well. In traditional terms, this statement means that words related in meaning contain

the same morphemes. (For example, the words *unidentified*, *identified*, *identify*, and *identification* are all related in meaning and are also related phonologically in that they all contain the sequence *identify*. Likewise, the words *unidentified*, *uninhibited*, and *unrealistic* are morphemically related in that they contain the sequence /un/ and they all have negative meanings related to their respective base word.) 2) The constituents of a word's form can be organized into a hierarchical structure that also represents the internal organization of its meaning and relation to other words (71-72).

In Anderson's model, as well as in the present analysis, we can see the formal correlation among words that are related in form, such as in the rule shown in Figure 3.1, which is an actual rule from a block in the model proposed in the present analysis. The top half of this rule is the *signifié* half of the linguistic sign, while the bottom half is the *signifiant*. Thus, this rule means that when the semantic features [1 Sg Pres Indic] are present for any given lexical item (here meaning the stem, which is the X variable in this rule), then the fully inflected form is the given stem plus /o/. This rule shows that words whose semantic feature sets include the notions of first person Present Indicative, such as *hablo*, *canto*, *bebo*, *vivo*, are related to each other in form. And, likewise, if we produce an entire paradigm for any given verb, then we can also see the correlation among words with the same lexical stem, meaning words whose base lexical items share the same semantic value, such as *hablo*, *hablas*, *habla*, *hablamos*, *habláis*, *hablan*.

Anderson sees the constituents of words not as lists of morphemes, but rather as operations on the form of words, such as the rule presented above. Thus, to Anderson, "the formal constituents of complex words" are not "listed morphemes, but rather operations on the form of words," in other words, Word-Formation Rules. The formal

constituents of a word are “a set of changes made by phonological rules” rather than a “lexicon of listed word-like elements” (1995: 72). For Spanish *hablas* ‘speak’ [2 Sg Pres Indic], for example, rather than seeing this word as a concatenation of /aβl/ and /as/, or as /aβl/ + /a/ + /s/, we might say that the lexical stem /aβl/ undergoes a rule whereby /X/ → /X + a + s/, whenever the rule represents a verb with the property of being [+ *a*-class] [+ 2 Sg] [+ Pres Indic]. There is a relationship between /aβlas/ and /aβlan/ in that they share a base. But there is also a relationship between /aβlas/ and /kantas/ ‘sing’ [2 Sg Pres Indic], for example, in that they have both undergone the same morphological process, and they share the features [+ *a*-class] [+ 2 Sg] [+ Pres Indic]. Rules such as the above example “thus constitute the basis of relations among words within the total vocabulary of a language” (Anderson 1995: 72). The present analysis, however, unlike Anderson, regards all morphology as part of the organization of the lexicon, as in Bybee (1985, 1988), rather than regarding derivation in the lexicon and inflection in either the syntactic or phonological component.

3.5. Ordering of Word-Formation Rules.

It is essential to ensure that the final inflected form contains all the right material and that this phonological material occurs in the right order, in the right position, or, we could say, with the right morphotactic organization. We must also make sure that regular and irregular modes of inflectional marking are complementary, thus ensuring that there are no overlaps or double marking, such as there would be if we said **tooked* for English Past tense of *take* or **soyo* for Spanish *soy* ‘be’ [1 Sg Pres Indic]. Likewise, we must ensure that the rules do not produce inaccurate or unattested forms.

Anderson (1988b, 1995), maintains that the ordering of Word-Formation Rules is language-specific rather than being uniformly predictable from general principles, but his Extended Word-and-Paradigm (EWP) model does present some general ordering principles. Anderson comments that “The descriptive role of ordering in the current context is the same as that of devices invoked by other theories, such as allowing ‘morphemes’ to subcategorize for the particular affix (or class of affixes) which they can follow within a word” (1995: 124).

Anderson proposes that morphological rules of inflection be organized into blocks. The blocks of individual rule-and-feature sets in Anderson’s model, and in the model presented in the present study, are conjunctively ordered. ‘Conjunctive’ ordering means that the blocks apply sequentially. Conjunctive/sequential ordering means that the first block must apply before the second, the second before the third, and so forth, and if no rule applies in any given block, then the stem proceeds to the next sequentially ordered block unchanged. It should be noted here that although these ordering stipulations imply a temporal sequence, the production of words in actual speech is, in effect, practically simultaneous, since the time involved in producing a single word form in normal speech is a mere fraction of a second.

Although the blocks themselves are conjunctively ordered as specified above, the relation among rules within the same block is ‘disjunctive’. Disjunctive ordering means that the rules within the blocks are mutually exclusive; the first applicable rule must apply, and then no other rule in that block may apply after the application of that first applicable one. This stipulation prevents any other rule in the same block from applying, even though its respective structural description may fit. Disjunctive ordering means that

“some pairs (or perhaps larger sets) of rules are subject to a condition of mutual exclusivity, such that the applicability of one of them precludes the application of the other(s)” (Anderson 1995: 45). There is no implication of temporal sequence in the mind of the speaker, however. Disjunctive ordering in this case just means that “where rule R_i precedes R_j in descriptive order this means that the change(s) performed by R_j presuppose any changes performed by R_i to the form” (124).

Consider one of the blocks of rule-and-feature sets from the model proposed in the present analysis, shown in Figure 3.4 below. This block is the first of several conjunctively ordered blocks that produce the entire set of forms for the Present Indicative, Present Subjunctive, Imperfect Indicative, and Imperfect Subjunctive for all Spanish verbs with regular stems. If the lexical stem entering the block is */kant/* (*sing*), for instance, then the first rule produces */kanto/* ‘sing’ [1 Sg Pres Indic], and this form takes precedence over the more general Rule # 7 in the block, even though *cantar* is an *a*-class verb. It is completely arbitrary and irrelevant that this Rule 1 happens to appear as the first rule in the block in this example, however. The rules within the block could, in fact, be placed in any order, since the rules within the blocks are disjunctively – rather than conjunctively – ordered. The more specific rule applies before a more general rule; thus we say that the rules in the block are disjunctively ordered. Since the blocks themselves, on the other hand, are conjunctively ordered, this sample block, which is the first of several blocks, must apply before the second and subsequent blocks can apply.

FIGURE 3.4. A sample block of rule-and-feature sets from the proposed model., showing disjunctive ordering of rules.

1) $\left[\begin{array}{l} + 1 \text{ Sg} \\ + \text{ Pres} \\ + \text{ Indic} \end{array} \right]$	2) $\left[\begin{array}{l} + \text{ Pres} \\ + \text{ Subj} \\ + a\text{-class} \end{array} \right]$	3) $\left[\begin{array}{l} + \text{ Pres} \\ + \text{ Subj} \\ + e/i\text{-class} \end{array} \right]$	4) $\left[\begin{array}{l} + \text{ Imperf} \\ + \text{ Indic} \\ + e/i\text{-class} \end{array} \right]$
$/X/ \rightarrow /X + o/$	$/X/ \rightarrow /X + e/$	$/X/ \rightarrow /X + a/$	$/X/ \rightarrow /X + \acute{a}/$
5) $\left[\begin{array}{l} + \text{ Imperf} \\ + \text{ Subj} \\ + e/i\text{-class} \end{array} \right]$	6) $\left[\begin{array}{l} + \{ 1 \text{ Pl} / 2 \text{ Pl} \} \\ + i\text{-class} \end{array} \right]$	7) $\left[+ a\text{-class} \right]$	8) $\left[+ e/i\text{-class} \right]$
$/X/ \rightarrow /X + ye/$	$/X/ \rightarrow /X + i/$	$/X/ \rightarrow /X + a/$	$/X/ \rightarrow /X + e/$

As a result of this disjunctive ordering, the application of a rule that seems to fit the structural description may be blocked by a more specific rule within that same block of rules. The more specific rule must apply first within a block, before the more general ones may be considered. The principle involved here is the so-called ‘Elsewhere’ Principle, commonly known and invoked in phonological and morphological theory (as in Anderson 1969, Kiparsky 1973, Aronoff 1976, Anderson 1986, Zwicky 1986, and others). The ‘Elsewhere’ Principle states that “Application of a more specific rule blocks that of a later more general one” (Anderson 1995: 132). Or in Spencer’s words, we “present all the exceptional cases first and then say that ‘elsewhere’, such-and-such [i.e., the general case] happens” (Spencer 1997: 215). Anderson notes, however, that

To sum up the various types of disjunctive ordering discussed above, we shall say that disjunctive ordering applies under the following conditions (from Anderson 1995: 143):

- 1) Between rules, by stipulation (reflecting the organization of rules into blocks);
- 2) Between a specific rule and a later more general rule;
- 3) Between stems within a lexical stem set, where one stem is more specific than another consistent with the requirements of the position under interpretation;
and
- 4) Between a specifically characterized lexical stem and a general rule.

It should be noted that all of the above types of disjunctive ordering are really just alternate ways of stating the general stipulation of ‘more specific before more general’, except for #1. The first ‘type’ of disjunctive ordering (# 1 above) means that the organization of rules into blocks stipulates a certain order from the outset, meaning ordering of the blocks themselves, although within the blocks the rules apply from most specific to most general.

The sample rule-and-feature sets shown in the Figures 3.1. and 3.4 above, also known as Word-Formation Rules, are part of the model proposed in the present study. These Word-Formation Rules are formatted in essentially the same way as Anderson’s Word-Formation Rules, except that Anderson uses a binary system of feature sets including the values [\pm me] and [\pm you]. This is presumably an effort to simplify the system as much as possible and reduce the number of possible combinations, by reducing

the number of features that need to be explicitly specified. For example, a semantic value of [– me, – you] implies third person (singular or plural), and a semantic value of [+ me, + you] implies first person plural. In the present model, the semantic values are generally stated as positive values, although there is thus implied a possible negative value that may be used as the need may arise. Also, the present model uses the notations [+1], [+2], and [+3] for person instead of the [± me] and [± you] notations. This is simply because these notations seem more standard and seem to work better with the proposed model for Spanish verbs. Both ways of presenting the rule-and-feature sets – the notation and organization system used in the present model as well as Anderson’s – would probably serve equally well to generate the paradigms of a given language’s verbal system.

3.6. What is a ‘lexical stem’?

Aronoff (1992) defines stems “purely as units of form” (5) and says that the traditional definition of ‘stem’ is “that part of a complete wordform which remains when an affix is removed (5). These definitions seem circular, however, since ‘affix’ is often defined as a form that is attached to a stem. Aronoff later defines ‘stem’ as “the form on which a particular form is built.” (18) and then modifies the definition further: “We can now define a stem as the form of a lexeme on which a realization rule operates” (p 32). Anderson (1988a, 1988b, 1995) regards the operation of inflectional Word-Formation Rules as part of the process of lexical interpretation, but these inflectional Word-Formation Rules must take place after the lexical interpretation of the stem, and after any derivational processes applied to the stem. According to Anderson’s model, the lexicon

provides a single stem form that underlies all of its phonological variants. But then how do the individual phonological forms arise?

The organization of the Word-Formation Rules into blocks that are disjunctively ordered according to the Elsewhere Principle accounts for much of the complementarity among inflectional markers, since, as mentioned above, the rules in any given block are mutually exclusive. But in order to account for the complementarity of regular and irregular inflection, we need to define the concepts of ‘lexical stem’ and ‘lexical item’. A lexical stem in this sense can be regarded as “an idiosyncratic association of phonological, syntactic, and semantic properties” (Anderson 1995: 122), although there may be more than one “lexicalized set of phonological properties” associated with the same set of semantic and syntactic properties. That is to say that there may be one lexical stem with several suppletive variants, although some may be only partially suppletive. For Anderson, a lexical item is “a contextually restricted principle for interpreting terminal nodes in a syntactic structure” (132). ‘Lexical insertion’ means finding a lexical item consistent with the position to be interpreted, that is, one whose lexical characteristics and subcategorization requirements are not contradictory to the semantic and morphosyntactic features specified by the *signifié* half of a given Word-Formation Rule. The selected lexical stem, that is, the selected phonological form, is then inflected through the operation of the inflectional Word-Formation Rules that correspond to the morphosyntactic properties required by the context. The association between the lexical stem and its syntactic and semantic properties corresponds to the link between the two parts of the classic Saussurean linguistic sign as discussed above.

However, there is sometimes more than one phonological stem for a given set of syntactic and semantic features. Anderson (1995) refers to this set of alternate forms of a lexical entry (classically viewed as ‘allomorphs’) as a ‘lexical stem set, defined by Anderson as “a group of phonologically distinct stems ... with the same syntactic requirements and semantic interpretation, each associated with its own (partial) set of morphosyntactic properties” (133). In this way, a given individual stem may be associated with morphosyntactic features “that restrict the range of Morphosyntactic Representations in syntactic structure that they can interpret” (133).

The choice among the possible stems in a lexical stem set of a given lexical item is governed by the following principle: When there are several possible stems for a given lexical item, the only stem that may serve as the basis of the inflected form is the one specified for the maximal subset of the features compatible with the Morphosyntactic Representation to be interpreted (Anderson 1995: 133). This stipulation is specified in the present model, in which a more specific rule must apply before a less specific (more general) rule may apply within the same block or in an earlier block. If the context (and therefore the semantic/structural description) requires [*be/ser* Pres Subj], for example, specifying a semantic value corresponding to *be/ser* along with a Present tense and Subjunctive mood value, and if there is a rule whose associated feature set specifies [*be/ser* Pres Subj], then that rule must apply before a rule whose associated feature set specifies the more general [Pres Subj *e/i/class*].

A lexical stem set “may contain several stems with associated features, as well (perhaps) as a stem with no features beyond its lexical category” (133). In choosing a particular lexical stem from the set of possible lexical stems, it is necessary to choose

“the most specifically characterized stem possible, consistent with the features of the position given” (133). The uncharacterized (or default) stem is only available if there is no more specific stem available. Thus, according to Anderson’s model, a ‘regular’ inflected word is one for which there is only one stem in the item’s stem set, and that one stem is thus the uncharacterized (default) stem. For example, for the Spanish verb *hablar*, there is only one stem in the lexical stem set: /aβl/. This stem is uncharacterized, meaning that it does not specify any morphosyntactic features, and is thus the default form, making *hablar* a regularly inflected verb.

We shall see that in the present Spanish verb model, as well as in Anderson’s model, this notion of ‘stem’ will be extended to include more than just the initial stem entering the first block of the rule-and-feature sets. As the blocks progress conjunctively, the stem entering each successive block is the phonological form exiting the previous block. This new ‘stem’ may be one that has added phonological material in accordance with a rule applied in the block, or it may be the ‘elsewhere’ default form, meaning the form that entered the block in the first place, if there are no more specific rules in the block that correspond to the features required by the semantic/morphosyntactic context.

Thus, although Anderson maintains that ordering of Word-Formation Rules may be language-specific, there are three general principles at work in Anderson’s theory of inflection (1995: 132-134), all of which are just further clarification and reinforcement of the precedence of specific cases over more general ones.

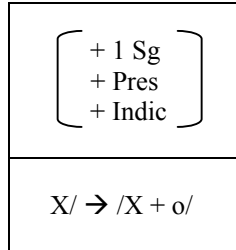
- 1) The ‘Elsewhere’ Principle, stating that “Application of a more specific rule blocks that of a later more general one.”

- 2) The only stem that may serve as the basis of an inflected form is one that is “characterized for the maximal subset of the features compatible with” the given Morphosyntactic Representation.
- 3) When a rule of the grammar would apply to a given stem on the basis of the features of a given position to be interpreted, application of the given rule is blocked if the feature set of the rule constitutes a subset of the lexical specifications of the stem. (134).

The example Anderson gives for this last principle is the English plural *oxen*, whose lexical stem set includes *ox* and *oxen*. What principle 3 above says for this particular lexical stem set is that application of the general rule for English noun plural formation, the rule that adds /s/ to the lexical stem, is blocked for the stem variant *oxen*, since the feature [+ Plural] is part of (that is, a subset of) the lexical specifications for the stem *oxen*. Thus, the above stated Principle 3 blocks the production of forms like **oxens*.

For Spanish, an example might be the stem set for the verb *ser*, a highly irregular verb with lots of suppletive forms. Depending on what inflection model we are working in, we could say that the stem set for this verb includes the following: *ser*, *se*, *fui*, *fue*, *era*, *soy*, *ere*, *es*, *so*, and possibly the single vowel phoneme /s/ (as in *siendo* and *sido*). The general rule for forming the 1 Singular Present Indicative form in the proposed model is shown in Figure 3.5.

FIGURE 3.5. General rule for the 1 Singular Present Indicative form for Spanish verbs.



This rule, however, is blocked for the stem /soy/, since the features [1 Sg Pres Indic] are a subset of the feature set for the stem /soy/. Actually, in this case, these four features [1 Sg Pres Indic] constitute the entire feature set for this particular stem, but the principle applies regardless. This is an example of how suppletive forms are dealt with in the present model. A completely suppletive form is unique, with a unique set of semantic/morphosyntactic features, and is thus more specific than the more general feature sets that apply to other ‘regular’ verbs; therefore, the application of the more general rules is blocked in accordance with the above three principles. A suppletive form that is a complete surface form with no apparent affixes, such as Spanish *soy*, may in itself be one of the stems of a lexical stem set. In such a case, the rule producing this single complete form is the most specific kind of rule in the model.

It is an interesting pedagogical anecdote, and perhaps worth future research, that a student of the present author recently produced the form **soyo* for the 1 Singular Present Indicative, which would be analogous to **oxens* in Anderson’s example above. The student apparently knew the ‘regular’ [1 Sg Pres Indic] /o/ affix and added this to the already suppletive form *soy*, thus producing a doubly marked form. The production of such a form could result from having sufficient input (or evidence) to establish the form *soy* as one of the members of the lexical stem set, but not enough input or evidence to

firmly establish the feature set associated with this stem alternant, that is, the feature set that limits application of the form just to the feature set [1 Sg Pres Indic] for the semantic value of *be/ser*. Similarly, many second-language learners, as well as young first-language learners, produce forms such as **sabo* (as opposed to *sé*) for [1 Sg Pres Indic] of *saber* ‘know’. In this latter case, the student or child presumably chooses the most common of the stems in the stem set and applies the ‘regular’ pattern by analogy to other 1 Singular Present Indicative forms.

‘Incorrect’ forms, then, may be produced when there is not sufficient input to establish a more specific rule, or a more specific feature set, to precede or override the more general rule. It is interesting to note that some varieties of English exhibit double plural forms such as *childrens* and *mens*. Similarly, in Mexican American Spanish, forms such as **vuelar* (for *volar* ‘fly’), **vuelamos* (for 1 Plural Present Indicative *volamos*), or **apreto* (for 1 Singular Present Indicative *aprieto* ‘grasp’) are common. What we would have to say about these latter regularly attested forms, however, as opposed to forms produced by second-language learners or young first-language learners, is that the feature sets of their lexical stem sets are different from those of Standard American English or Standard Spanish. The feature sets for such verbs for these speakers are consistent, however, with the norms for their speech community.

3.7. What is a paradigm?

According to Anderson (1995), “the paradigm of a lexical stem item in terms of its lexical stem set ... is the complete set of surface word forms that can be projected from the members of its stem set by means of the inflectional Word Formation Rules of

the language” (143). The paradigm “results from the interaction of the members of its stem set with the inflectional rules of the language” (134), and the only forms that can be part of the paradigm are forms that are consistent with the three principles listed in Section 3.6. Anderson (1995) says that it is “often argued that words differing only in their inflection group together into paradigms,” but he adds that “this is of little help in identifying ‘inflection’ unless we know when words should be said to belong to the same paradigm” (79). Ruiz (1986) defines paradigms as “groups of closely related forms, like all forms of a given verb in its various tenses” or as “a set of subsystems in which we find closely related forms” (104). Bybee (1985) defines “the lexical representation of a paradigm” as “a cluster of words bearing both a semantic and phonological relation to one another, in which one word (and sometimes more than one) is stronger than the others” (124).

It is arguable whether speakers of a natural human language normally have conscious awareness of the rules, patterns, and paradigms in their native language. According to Bloomfield (1933), psychologists maintain that normal speakers are not capable of the kind of reasoning implied by inflectional paradigms. If this is true, then “the normal speaker, who is not a linguist, does not describe his speech-habits, and, if we are foolish enough to ask him, fails utterly to make a correct formulation” (406). Bloomfield comments that linguists overestimate the ability of speakers to recognize their “speech-habits,” pointing out that linguists “forget that they owe this ability [to describe a language’s rules, patterns, and paradigms] to a sophisticated philosophical tradition” (406). Bloomfield says that such paradigms and other statements about the ‘rules’ of a

given language merely describe the action of native speakers of a language and do not “imply that the speaker himself could give a similar description” (406).

Although paradigmatic lexical organization does not imply a conscious awareness of paradigms as such, there is some evidence for the existence of some type of paradigm in the speaker’s consciousness (Aitchison 1991: 147-148; Bybee 1985: 51, 57, 64-65, 126; Wurzel 1984), because of processes such as analogical leveling and regularization in language change and variation. Pedagogically traditional verb paradigms based on tense/mood/aspect with six-member person/number-based sets for any given lexeme are actually consistent with Bybee and Pardo’s (1981) hypothesis of a hierarchical relationship among the surface forms of verb paradigms. Contrary to the generativist notion of an abstract underlying representation for a verb paradigm, in which the surface forms “do not have any official relation to one another” (947), Bybee and Pardo found evidence in empirical investigations with speakers that there is psychological validity to such paradigms and that there are indeed relations among surface forms. These researchers found strong evidence that “the person/number forms within one tense or aspect are more closely related to one another than they are to forms outside that tense or aspect” (947). Ruiz (1986) concurs with this observation when he comments: “The type of relationship among the forms of a given subset differs from the relationship among forms from different paradigms. Forms of the same paradigm are expected to be more closely related to each other” 104-105). In other words, alternations in the stem are more likely to occur within a tense or mood rather than across tense/mood/aspect boundaries. For the verb *poner* ‘put’, for example, the stem /pus/ occurs in all persons and numbers of the Preterite Indicative and Imperfect Subjunctive forms: *puse, pusiste, puso, pusimos,*

pusisteis, pusieron (Preterite Indicative); *pusiera, pusieras, pusiera, pusiéramos, pusierais, pusieran* (Imperfect Subjunctive). However, there is not a single stem variant that is found in all 1 Singular forms, as in: *pongo* (Present Indicative), *puse* (Preterite Indicative), *ponía* (Imperfect Indicative), *ponga* (Present Subjunctive), *pusiera* (Imperfect Subjunctive), *pondré* (Future Indicative).

Chapter 7 of the present study looks at the various ways these paradigms are organized for pedagogical purposes, and the proposed model presents one of the ways in which these paradigms might be organized, or at least connected, in the mental grammar. A Word-and-Paradigm theory of morphology contends that there are certain generalizations about words and their relationships among each other that can only be explained at the level of the whole word rather than at the level of its alleged constituent morphemes. A WP model recognizes the role of inflectional paradigms in the organization of a grammar. According to Spencer (1997), the “key to the WP approach is our notion of the morphosyntactic word. Each inflected form has (at least) one morphosyntactic description (for example ‘past tense form’ or ‘dative singular of the masculine/neuter adjectival form’) and the grammar then makes available paradigms that specify the formatives which correspond to these categories” (52).

3.8. Problems with Anderson’s model.

One problem with Anderson’s view is that of assigning semantic content to the phonological component. If the notion of plurality, for instance, is part of the set of semantic features associated with a given whole, inflected word, then we either have to have fuzzy boundaries between the semantic and phonological components, with some

semantic features associated with the phonological component, or we have to move the inflectional process back to the lexicon entirely. If, however, we view inflection as a syntactic process, then we are again blurring the boundaries between the various components of the grammar, and Chomsky contends that no syntactic operations can take place at the word level. If we abide by Chomsky's stipulation, then either inflectional operations are not syntactic in nature, or if they are, then they must take place in the syntactic component rather than in the lexicon. In this case, we would be attributing semantic features to the syntactic component, like the above suggestion of attributing semantic features to the phonological component.

According to Anderson, word structure is "a product of interacting principles from many parts of the grammar: at least phonology, syntax, and semantics, in addition to the 'lexicon'" (1995: 2). Morphology for Anderson, at least, may not be localized, as conventional views of grammar assume, located in its own separate component of the grammar, but may actually be represented in various components of the grammar, and may be different for inflectional vs. derivational morphology. However, given the gradient nature of derivational and inflectional morphology, as discussed above, the present analysis concurs with Bybee in regarding all morphology as part of the organization of a dynamic, multidimensional mental lexicon, discussed in more detail below.

3.9. Morphology as lexical organization.

As discussed throughout the preceding sections, linguistic analyses have traditionally been based on the notion that a language is a system of rules. This basic

premise of language as a rule-governed system applies to structuralists and generativists alike. However, others such as Skousen (1989), Bybee (1985, 1988), and Rumelhart (1989) have proposed different ways of accounting for language behavior. Rumelhart's connectionist model is a model of cognition that applies to other kinds of learning as well as language. Connectionist models of cognition view form-meaning connections as occurring throughout the whole neural network of the brain, not necessarily in a linear fashion, but rather, through parallel distributed processing (PDP) and through associative processes rather than the construction of rules. The strength of the connections is dependent upon activation of neural networks through associative processes rather than by abstract rules. Rumelhart (1989) explains that

Units interact by transmitting signals to their neighbors. The strength of their signals and therefore the degree to which they affect their neighbors are determined by their degree of activation.... It is this pattern of connectivity that constitutes what the system knows and determines how it will respond to any arbitrary input (139).

In general, non-rule-based approaches are referred to as 'analogical' or 'exemplar-based' approaches, with Rumelhart's connectionist model a subset of this broad "class of possible models" (Skousen 1989: 56). This class of models "can all reasonably be referred to as 'analogical' and ... can be sharply differentiated from the kind of 'symbolic' or rule-based approaches that are exemplified by generative grammars in linguistics" (56). Skousen points out that the term 'analogical' is not to be confused with the more general definitions and imprecise notions of analogy used in the past or in

other contexts, but rather involves an explicit, specialized definition. Skousen's analogical model of language applies to all of language, from the completely regular to the completely suppletive. Skousen explains that "an analogical description predicts behavior by means of a collection of examples called the *analogical set*" (4), rather than predicting behavior based on generative, productive rules.

An analogical set in Spanish, for example, might contain a number of 1 Singular Present Indicative verb forms, such as *hablo, canto, bailo, bebo, como, aprendo, escribo, vivo*, or perhaps all 1 Singular Present Indicative verb forms with velar insertion in the stem, such as *pongo, tengo, vengo, traigo, hago*. Traditional verb paradigms or conjugations, such as *hablo, hablas, habla, hablamos, habláis, hablan* (all Present Indicative forms for *hablar* 'speak') are an example of one kind of analogical set, but they are not the only kind. Any set of words that share semantic or phonological connections, or both, can be part of an analogical set, with varying degrees of relatedness among the members of the set based on proximity in form, that is, how alike the various forms are semantically and phonologically, and on the frequency of occurrence of the given forms.

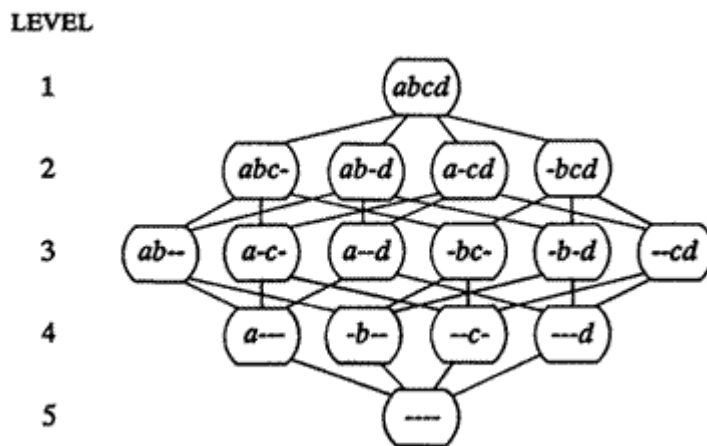
Unlike generativist models, Skousen's model is empirical and data-based. It is based on 'exemplars', or tokens of linguistic forms, and the analogical connections among those forms, along with the frequency with which each of those forms occurs. The analogical set for a given form *x* is determined by "looking through the data for 1) classes of examples that are the most similar to *x* and 2) more general classes of examples which behave like those examples most similar to *x*" (4). In addition to similarity with the given form *x*, the frequency of occurrence also contributes to determining an

analogical set. In contrast with a rule-based model, an analogical approach does not actually use rules to predict behavior, but rather “makes predictions in terms of individual occurrences” (22).

Eddington (1999) explains that “The behavior of the word(s) most similar to the word in question generally predicts the behavior of the word in question, although the behavior of less similar words has a small chance of applying as well” (2). He also clarifies the conditions under which a certain word will or will not be chosen from the analogical set, or, in other words, the restraints on the application of analogy to a given form. These conditions are: 1) proximity (how semantically and phonologically similar the word is to the word in question); 2) gang effect (if the example is surrounded by other examples having the same behavior); 3) heterogeneity (if there are intervening examples closer to the word in question but with different behavior). When the first two of these three conditions are present, they increase the likelihood that a certain form will apply, while the third, heterogeneity, decreases that likelihood. For example, for the verb *tener*, an analogical set might include *tenía, tenías, teníamos, teníais* (all Imperfect Indicative forms). A related analogical set might include *tienen, hablaban, escribían, cantaran* (all 3 Plural forms of different verbs and different tenses), so for a 3 Plural Imperfect Indicative form, the form *tenían* would be chosen by analogy to the words in proximity. However, for a 3 Plural Preterite form, there would be intervening heterogeneous forms such as *tuve, tuviste, tuvimos, tuvisteis* that are closer in meaning to the word in question (the 3 Plural Preterite form). Therefore, these latter forms would keep the /ten/ forms from being chosen, as in the incorrect **tenieron*, but rather, the correct word *tuvieron* would be chosen.

An analogical model, like connectionist models, regards language processing as parallel and as distributed throughout a network, rather than as sequential. There are networks of connections among linguistic forms, just as there are networks of neural connections among the data in other kinds of knowledge as well. In fact, Skousen (1989) points out that although his work specifically concentrates on language, “many of the results [of his analogical model and analyses] are applicable to more general kinds of behavior” (6). The diagram in Figure 3.6 represents part of such a connected network for four variables *abcd* (137).

FIGURE 3.6. Example of a parallel distributed network of connections for four variables *abcd*, showing the possible connections among these variables (Skousen 1989: 137).



An analogical model “appears to be rule governed” (5), as Skousen points out, but the notion of ‘rules’ for Skousen and Bybee is quite different from the generativist notion of rules. For Bybee (1988), “morphological rules and lexical representations are not separate from one another. Rather, morphological and morphophonemic rules are

patterns that emerge from the intrinsic organization of the lexicon” (125). Rules are not predictors or determiners of behavior, but rather are descriptions of behavior. Skousen (1989) contends that rules serve as “a kind of metalinguistic language that efficiently describes past behavior and allows us to talk about that behavior” (139), but that our viewpoint in doing so is structuralistic and is not necessarily predictive of new behavior. Comparing linguistic rules to scientific laws, Skousen maintains that “linguistic rules are metalinguistic devices that exist only in the minds of linguists,” just as the laws of physics “have no existence except in the minds of scientists” (140). Objects that fall to the ground when dropped may appear to be ‘obeying’ Newtonian laws of physics, and, similarly, speakers of a language may appear to be following rules in perceiving and producing language. Like Skousen and Bybee, the present model of Spanish inflectional verbal morphology regards rules as descriptions of lexical organization rather than predictors or determiners of linguistic forms.

Bybee’s model of ‘morphology as lexical organization’ (1988) is morpholexical, regarding both inflectional and derivational morphology not as a separate component of the grammar, but, rather, as a part of the lexicon. She proposes that lexical storage and organization are dependent upon the application of several principles: 1) “the ability to store strings of linguistic material in phonological and semantic representation,” 2) “the ability to form, among stretches of this material, connections of a semantic and phonological nature,” 3) “the accumulation of lexical strength due to token frequency,” and 4) the ability to organize sensory stimuli into categories” (139). She argues that this type of lexical storage and organization allow language users to “conceptualize the internal structure of a word as a set of relations with other words, rather than as a string

of discrete meaningful sequences” (139). She contends that this type of lexical organization solves a number of the problems with traditional views of morphemes as discrete units, the problem of ‘submorphemic’ units, for instance. By this she means units that are not necessarily directly associated with a semantic counterpart, such as the so-called ‘cranberry morphs’, dubbed this because the ‘cran’ part of the word has no identifiable meaning, although the ‘berry’ part does. This view also handles vowel class, as we see in Spanish *a*-class, *e*-class, and *i*-class verbs, in which these theme vowels have no semantic significance. Such a whole-word, connected-network conceptualization of the lexicon “allows the identification of a part of a word as a recurring meaningful unit without the necessity of assigning meaning to the remainder, and it allows the identification of phonological relations even in the absence of a clear semantic relation” (139).

Chialant and Caramazza (1995), in discussing the ‘whole word representation hypothesis’, describe the lexicon as “a full list of the lexical forms previously encountered – whether they are morphologically simple or complex” (59). For Bybee, the lexicon for any given speaker of a human language does indeed contain all previously encountered forms. Bybee’s (1985, 1988) model and the model proposed in the present analysis, however, do not regard the lexicon as a mere listing of forms, but rather as a multidimensional dynamic network of connections among forms. Whole words are connected to other similar whole words, with the connections based on semantic and phonological similarities. For example, the form *hablaban* ‘talk’ (3 Plural Imperfect Indicative) shares connections with *hablo* ‘talk’ (1 Singular Present Indicative), but *hablaban* also shares connections with *cantaban* ‘sing’ (3 Plural Imperfect Indicative)

and with *bailabas* ‘dance’ (2 Singular Imperfect Indicative). *Hablaban* and *hablo* share the semantic notions associated with ‘talk’ and they also share the initial phonemes /aβl/. *Hablaban* and *cantaban* share the semantic/ morphological features of [3 Pl] and [Imperf Indic], as well as the phonemes /aβa/ in the middle of each form. *Hablaban* and *bailabas* share the semantic/morphological feature imperfective past tense [Imperf Indic] and the phonemes /aβa/. but they do not share initial phonemes or final phonemes, nor do they share semantic features other than that of Imperfective Past tense. Thus these five words illustrate in a small way the myriad connections that exist among linguistic forms in the lexicon.

Chialant and Caramazza (1995) explain that for the whole word representation hypothesis,

Morphological information associated with a given lexical form (e.g., *walked*) would then be part of its semantic information (e.g., the meaning of *walked* is represented by the same features that represent the meaning of *walk* plus the feature of past tense) (59).

Similarly, the examples in the preceding paragraph show morphological information as essentially no different from the so-called ‘content-bearing’ semantic information represented by a given linguistic form. Anderson (1995) points out that when a noun is marked as plural, that marking does not leave the meaning of the word unaffected. For example, the meaning of the word *dog* is not the same as the meaning of *dogs*, even though ‘plural’ is an inflectional category in English. Anderson comments that “One could of course say that this is actually an instance of ‘inflectional meaning’ as opposed to some other, presumably more ‘genuine semantic’, meaning, but in the absence of a

serious theory of such a notion it simply begs the question” (79). The features traditionally called ‘grammatical’ or ‘morphological’ are regarded in this view as part of the overall semantics of a given linguistic form, especially given the gradient nature of the continuum between lexical expression and syntactic expression as shown above, and the lack of a consistent one-to-one relation between form and function. It is for these reasons that the model proposed in the present analysis views semantic and grammatical features as essentially the same thing, even though they can sometimes be separated and identified. They are both part of the overall meaning of a given word. The overall meaning of the English word *dogs*, for example, in addition to the semantic notion of ‘furry, four-legged domesticated canine’, includes the semantic notion of ‘more than one’, which linguists label ‘Plural’ for the sake of convenience and for notational purposes.

The semantic and phonological connections between and among forms in the lexicon can also be seen as an overlapping of forms or a matching of forms. Chialant and Caramazza describe the recognition of previously encountered words as “a matching procedure between the input stimulus and the lexical representation” (59). The semantic and/or phonological representations of one form can partially map onto, or match, the semantic and/or phonological representations of another form, as in the above examples of *hablaban*, *hablo*, *cantaban*, and *bailabas*. Parts of each of these forms could be mapped onto similar parts of each of the others, with the various combinations of these forms connected phonetically or semantically, or both, as in the examples discussed above. Some forms may be connected semantically but not phonologically, such as the initial phonemes of any inflected form of the verbs *hablar* ‘talk’, *charlar* ‘talk, chat,

prattle’, and *platicar* ‘chat, converse, talk’. Other forms may be connected phonologically but not semantically, such as the initial phonemes /kas/ of *casa* ‘house’, *casar* ‘marry’, *cascada* ‘waterfall’, and *cáscara* ‘rind, shell, peel’.

Bybee (1988) defines morphological relations as “semantic and phonological connections that run in parallel,” pointing out that “where semantic and phonological connections coincide exactly, morphological identity can be established” (127). The graphic illustration in Figure 3.7 shows examples of such overlapping semantic and phonological connections for the English words *cat*, *cats*, *caps*, *mats*, and *rats*. The bolder lines represent coinciding semantic and phonological connections, while the narrower lines represent phonological connections only. The word *cats* shares semantic and phonological connections with the singular *cat*, but it also shares semantic/phonological connections with *mats*, *laps*, *naps*, *tips*, *mits*, and all other plural nouns ending in /s/ as well, “on the basis of the shared semantic feature plural and the identity of the final fricative” (127).

FIGURE 3.7. Overlapping semantic and phonological connections for the English words *cat*, *cats*, *caps*, *mats*, and *rats* (Bybee 1988: 126-127).

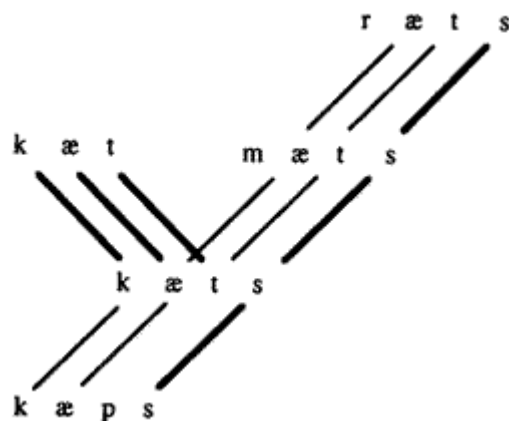
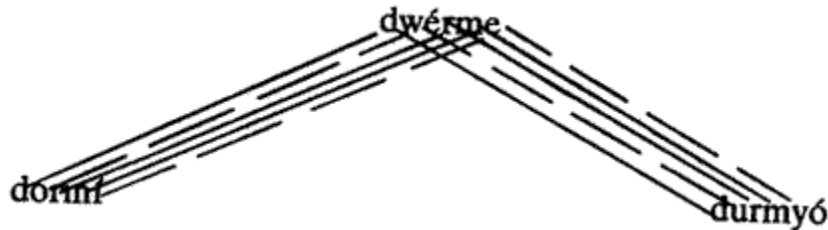


FIGURE 3.8. Overlapping semantic and phonological connections for forms containing the three allomorphs for the verb *dormir*, with solid lines representing relations of identity and broken lines indicating shared (but not identical) features (Bybee 1985: 126).



When semantic and phonological connections coincide exactly, that is, when they run parallel, those overlapping forms would be called ‘morphemes’ in traditional terminology. Bybee (1988) comments that the way speakers and linguists discover that a word consists of more than one morpheme is that “We find a relation of phonological and semantic identity or similarity between some subpart of the word and a subpart of another word” (127). The /aβa/ sequence in the above Spanish verb examples could be separated out and identified as a grammatical morpheme representing the past tense, imperfective aspect. In Bybee’s model, however, the speaker is not necessarily aware of these separated parts and does not dissect the words that way linguists are accustomed to do. The learning process in Bybee’s model occurs in this way: “When a new morphologically complex word is learned, it forms connections with existing lexical material on the basis of its meaning and phonological shape. The word is not physically dismembered, but its parts are nonetheless identified” (127), and those parts make analogical connections with similar parts of other words already existing in the lexicon.

However, most cases in Spanish are not as simple as the /aβa/ example given above or the English noun plural marker /s/ as shown in Figure 3.7. The traditional structuralist notion of the morpheme as ‘the smallest meaningful unit’ (Bloomfield 1933: 166) assumes ‘meaning’ to be the major criterion for identifying morphemes. Ideally, then, the phonological material and the semantic content should overlap completely and perfectly, with a precise one-to-one correspondence between meaning and form. In real languages, however, that is rarely the case, and it is not the case with Spanish verbal inflectional morphology. In Spanish, there are many cases of multiple exponence, with several morphological *signifiés* corresponding to one phonological form, as in *hablo*, *canto*, *bailo*, in which the /o/ signifies person, number, tense, and mood (1 Singular Present Indicative) all in one single vowel sound. There are also cases of syncretism, in which a single formative represents more than one meaning in different contexts, as in *habla*, which can be either a 3 Singular Present Indicative form or a Singular Familiar Imperative form. Suppletive forms pose another problem for the ideal one-to-one relationship between form and function, since verbs such as *ser*, for example, may have forms that share no phonological similarities with each other or with the infinitive form *ser* itself, such as *fue* (3 Singular Preterite Indicative) or *eres* (2 Singular Present Indicative).

Bybee (1988: 128-129) mentions two other problems presented by the traditional notion of morphology. One is the so-called ‘huckleberry morphemes’, ‘cranberry morphemes’, or ‘ceives’. Although these problems concern derivational morphology rather than inflectional morphology, which is the subject of the present analysis, they are mentioned here as further clarification of Bybee’s analogical model of morphology.

‘Huckleberry morphemes’ or ‘cranberry morphemes’ refer to words such as *huckleberry*, *cranberry*, or *strawberry*, in which the second part of the words are obvious and recognizable, but the first part of each word is meaningless as a morpheme in the traditional structuralist sense of ‘minimal unit of meaning’. Similarly, the English verbs *receive* and *deceive* seem to be related morphologically, but the *-ceive* part that they share in common has no meaning in itself.

Another quasi-morphological phenomenon that can be handled by Bybee’s analogical model but not by traditional approaches is the phenomenon of “phonaesthemes and other sub-morphemic units” (128-129). ‘Phonaesthemes’ are initial clusters that seem to have the same semantic feature in common, such as the /*sn*/ sequence in words pertaining to the nose, for example, *sneeze*, *snore*, *snort*, *sniff*, *sniffle*, *snivel*, *snoot*, *snout*, *snot*. Similarly, the /*fl*/ cluster denotes movement through air, as in *flit*, *fly*, *flutter*, *fling*. Although such sound sequences seem to represent semantic content, the remaining phonemic sequences in each of these words do not convey any meaning by themselves as do the clearly identifiable parts of polymorphemic words such as *unhappy*, *disable*, *rewrite*, *underestimate*. In terms of Bybee’s model, we could simply say that although the phonemic segments /*fl*/ in *flit*, *fly*, *flutter*, and *fling* share semantic connections, the phonemic segments following these initial sounds do not share semantic connections with any other words. Bybee’s model “allows the description of the full range of phenomena and also allows for differential status according to whether both semantic and phonological connections are made, and how strong these connections are” (129).

3.10. Strength of lexical connections.

Strength of lexical connections is an important aspect of Bybee's model, and strength of connections among forms is directly correlated to frequency of occurrence of the given forms. Thus frequency of forms is an important factor in an analogical model, unlike rule-based generativist models of morphology, which do not take into account the frequency of application of morphological rules. Bybee points out that "Each time a word is heard and produced it leaves a slight trace on the lexicon, it increases in lexical strength" (1985: 117). Conversely, infrequently used forms have lesser lexical strength, and, in fact, "An irregular form that falls below a critical frequency could disappear outright after a few generations" (Pinker 1999: 201). Forms that are not used often have a lower degree of lexical strength, which is an index of word frequency, while frequently used forms gain lexical strength, thus creating a "gradient representation of lexical strength" (Bybee 1988: 31). This notion of lexical strength accounts for the continued use of irregular and suppletive forms in a given language. In most human languages, it is typically the most common verbs of a language, rather than less commonly used verbs, that exhibit irregularity and suppletion. Irregular forms in low-frequency lexical items would not have the strength to continue in the language and would gradually fall out of use until they died out completely. Bybee says that lexical strength, or the lack thereof, also accounts for the regularization of infrequent irregular forms, such as **andó* for 3 Singular Preterite Indicative *anduvo*. "Patterns that range over large numbers of lexical items are highly reinforced or strengthened and apply more readily to new items, while patterns that are found in a smaller number of items are correspondingly weaker and less

apt to be productive” (125). Bybee comments that “A larger number of distinct verbs participating in the same pattern will serve to strengthen it” (138).

Similarly, Chialant and Caramazza maintain that the mapping or matching procedure is “sensitive to the relative frequency of occurrence of words” and therefore affects the speed of processing of lexical forms (1995: 59). Thus, as Bybee (1988) points out, “the difference between major productive rules, minor rules, and suppletion is just a matter of degree, not a matter of qualitative difference” (125). The degree of relatedness among forms can be seen as a continuum, so that rules and representations are essentially the same thing, not entirely separate entities. At one end of the continuum are completely suppletive forms, and at the other end are the highly regular forms that linguists group together according to rules that describe the regular patterns these regular forms exhibit. For example, the Spanish form *es* (*be/ser*, 3 Singular Present Indicative) would be at one end of the continuum as a completely suppletive form, since it has no semantic-phonological identity with other forms. A form such as *cantaban* would be at the other end of the continuum, since it can be described in terms of a ‘rule’, which could be written as: Verb *X* [3 Pl Imperf Indic] = Lexical stem + /aβa/ [Imperf Indic] + /n/ [3 Pl]. For the verb *cantar*, then, the rule would be realized as: ‘sing’ [3 Pl Imperf Indic] = /kant/ + /aβa/ + /n/, which is spelled conventionally as *cantaban*.

3.11. Degree of relatedness.

Bybee (1988) maintains that the degree of relatedness among forms and their consequent relative placement on the rules/representations continuum, depends upon how closely related the given forms are semantically: “The more closely related two forms are

semantically, the more likely they are to be similar morphophonemically” (130). Certain morphological categories are more tied to the overall semantic value of the verb than are other categories, such as aspect. Bybee points out that aspect is represented in some languages by an inflectional marker, while in other languages aspect for a given verb may be expressed lexically and thus expressed by an entirely different verb. For example, the English verbs *do* and *complete* represent imperfective and perfective aspects respectively, as do the verbs *know* and *realize*. In Spanish, however, the verb *saber* in most of its forms signifies imperfective aspect as in *know*, but the Preterite forms of *saber* show an aspectual difference and are thus closer to the English verbs *realized* or *found out*. Thus, the aspectual difference for these concepts is represented lexically in English, but morphologically in Spanish.

Bybee (1985: 11-37) also notes that the phonemic forms that signify morphological categories occur in approximately the same order in different languages, at least in cases for which it is possible to separate and morphologically describe the affixed segments. Morphological categories that are most relevant to the meaning of the stem of the verb are physically closer to the stem, while categories that are less relevant to the semantic value of the stem are physically more peripheral to the stem, that is to say, farther from the stem. Aspect, for instance, is typically closer to the stem, since aspect is closely related to the semantic value of the verb, as shown in the example discussed above. Person and number, however, are less semantically related to the stem and thus their signifiers are typically more peripheral. In the above mentioned forms *hablaban* and *cantabas*, for example, the /*aβa*/ segment signifying imperfective aspect (and past

tense as well in this case) is physically closer to the stem, while the PN markers /n/ [3 Pl] and /s/ [2 Sg] are affixed last and are thus more peripheral to the stem.

One effect of frequency of forms and strength of connections is what Skousen (1989) calls ‘gang effects’, explaining that

the gang effect measures the degree of regularity for any given context that actually occurs in the data. Typically, exceptional data occurrences have a gang effect of less than one, whereas regular data occurrences have a gang effect of greater than one (8).

Stemberger and MacWhinney (1988) explain ‘gang effects’ further: “Given storage [of items in the lexicon], items that are similar will tend to reinforce each other and lead to faster, more accurate performance” (103). Eddington (1999) points out that this ‘gang effect’ may account for ill-formed tokens such as **brang* (for *brought*, Past Participle of *bring*) in English. A gang effect would analogically link *bring* with the pattern *sing/sang/sung*, thus producing **brang* for some speakers, rather than linking with *brought*, since similar Past Participle forms such as *fought* (from *fight*) or *bought* (from *buy*) do not have bare infinitive forms that are analogous to *bring*, as does the rhyming form *sing*. According to Skousen, the common occurrence of ill-formed tokens such as these is a conceptual problem for rule-based generativist approaches. Another problem for rule-based models is the problem of redundancy, meaning that they have difficulty accounting for the apparent blocking of doubly marked forms such as **wroted* or **broughted*. (Skousen 1989: 7).

3.12. 'Rules' in the proposed Model.

In the Spanish inflectional verbal morphology model proposed in the present analysis, the constants in the rules are an example of the kind of semantic-phonological connections that serve as the basis for Bybee's (1985, 1988) model. For instance, the rule $/X/ \rightarrow /X + o/$, as in the first block of rule-and-feature sets in the Basic Model introduced in Chapter 1, Figure 1.3, shows an analogical link among all verb forms ending with $/o/$. In other words, this rule shows a connection among all verb forms whose *signifié* includes the semantic notion of 1 Singular Present Tense and whose *signifiant* ends with the phonological segment $/o/$. The rules in the proposed model are just a different way of illustrating the kinds of semantic-phonological connections shown in Bybee's graphics in Figures 3.7 and 3.8.

Also like Bybee's model, the proposed model accounts for suppletion and regular forms alike, and it views rules and representations as essentially the same thing. Unlike an Item-and-Process (IP) view of morphology, the present analysis concurs with Bybee in regarding items and processes (or representations and rules) as essentially the same thing. As Bybee says, "the best exemplar of a rule and the best exemplar of a representation are two poles of a continuum," and "some rules have properties we associate with representations while some representations bear a resemblance to rules" (1988: 121). Bybee's continuum of forms ranging from the completely suppletive to the completely regular is reflected in the ordering of the rules in the proposed model. Completely suppletive forms, like regular forms, are represented as rules in the proposed model, and are thus essentially no different from the representations for completely regular forms. The rules for suppletive forms, however, must apply earlier than the rules for the regular

forms, according to the model's rule-ordering stipulation of 'more specific before more general'. Furthermore, because of the nature of the rule-ordering in the model, the model cannot produce doubly marked forms such as **soyo* for *soy* {*be/ser*, 1 Singular Present Indicative). The proposed model also accounts for what Skousen (1989: 6) and the present analysis label 'exceptional/regular' forms, such as the Spanish Preterite Indicative forms *traje*, *trajiste*, *trajo*, *trajimos*, *trajisteis*, *trajeron*, which are irregular but have 'regularities in their irregularity'.

The rules in the proposed model, then, are not independent 'rules' in the generative sense as in Anderson's Word-Formation Rules, although the rules in the proposed model follow Anderson's model in form. The rules in the model proposed in the present analysis of Spanish verbs are, rather, descriptions of patterns and connections in the lexicon, as are rules and representations as described in Bybee's model. The rule-and-feature sets in the proposed model are really just descriptions or representations of the forms themselves, and these representations reflect the two-part nature of the linguistic sign, with the feature sets representing the *signifiés* and the 'rules' representing the *signifiants*.

Another way that the model proposed in the present analysis supports Bybee's (1985, 1988) model of morphology as lexical organization is in the notion of lexical strength and the notion of frequency of occurrence of lexical forms. The proposed model depends upon the building up of feature sets in the lexicon, that is, the establishment of connections between phonemic forms and feature sets, so that the correct form can be mapped onto its corresponding feature set. In Bybee's model, as mentioned above, every time a word is heard – or produced – it leaves a trace on the lexicon, thus increasing in

lexical strength. “So accumulated lexical strength depends on a certain degree of phonological and semantic correspondence” (1985: 117). Therefore, the semantic feature sets in the blocks of rule-and-feature sets of the proposed model depend upon sufficient input in order for the lexicon to build up the necessary connections among analogical forms and between individual feature sets and their corresponding phonemic forms.

An anecdotal example might serve to demonstrate the possible impact of insufficient input for the building up of semantic feature sets and for establishing strong connections between feature sets and their corresponding forms. A four-year-old Spanish-speaking child known to the present author was overheard playing with a ‘superhero’ action figure. After a vigorous rescue mission enacted by the child with the figure, the figure fell to the floor and the child pronounced: “Está morido.” The child’s mother then corrected: “Está *muerto*.” Then the child vehemently protested, insisting: “¡No! ¡Está morido! La palabra NO es *muerto*! ¡Es *morido*!” This anecdote illustrates the importance of Bybee’s (117) notion of lexical strength related to frequency of occurrence, as well as the notion of suppletive and irregular forms represented as specific rules that take precedence over more general rules. The young child who insisted that **morido* was the ‘correct’ form of *muerto* had obviously not heard enough occurrences of the irregular Past Participle form *muerto* to establish a specific suppletive rule. He apparently produced the form **morido* based on analogical connections with other Past Participle forms that ended in /iðo/, such as *comido* (from *comer* ‘eat’) or *bebido* (from *beber* ‘drink’), which he is likely to have heard much more often than *muerto*. In other words, there was not sufficient lexical strength for the suppletive form *muerto* to block application of the more general rule/representation that would yield **morido*.

3.13. Marslen-Wilson: a cohort model of lexical output.

Marslen-Wilson's (1989: 7-8) cohort model of lexical processing contends that words are often recognized upon the listener's hearing the initial phonemes of the word, and if the word is part of a longer utterance, then the recognition of the word is reinforced by the syntactic and semantic context of the utterance as a whole. As soon as a listener hears the initial phonemes of a word, acoustic-phonetic information is mapped onto the lexical level during the process of lexical access and selection. Upon hearing the initial phonemes of a word, the listener instantaneously accesses a group of possible words ('cohorts') that the whole word could be. As Emmorey and Fromkin (1988) explain, "the first one or two phonemes serve to 'activate' all words in the listener's lexicon which begin with that initial sequence, and these words form the 'word-initial cohort'" (126). It is not just phonetic cues that narrow the group of possible cohorts for an initial sound sequence, however. In actual communicative speech, the semantic and syntactic requirements of the given context, along with the collocations, or words surrounding the given sound sequence, also help to narrow the cohort group.

This initial group of cohorts could be very large if the word is an isolated word out of any meaningful situational context. In Marslen-Wilson's (1989) view, all words in the lexicon with the same initial sound or sound sequence would be accessed upon hearing a given initial sound or sounds. Upon hearing the initial syllable /ma/ of *madrugar* 'get up early in the morning', the hearer accesses all words that begin with this syllable, such as *magullar*, *manosear*, *mantener*, *masacrar*, *matar*, *manipular*, *mascar*, *manejar*, including words other than verbs, such as *mano*, *malo*, *material*, *mantel*, *manga*, *mamá*, *maíz*.

Marslen-Wilson maintains that “the speech signal is rich in what we can call partial information – that is, anticipatory cues to the identity of an upcoming segment” (9). His research shows that there are phonetic cues, such as vowel nasalization, that further restrict the possible cohorts for an upcoming segment. For instance, all of the above words that begin with /ma/ followed by a nasal consonant would have this nasalized vowel. So, upon hearing an initial /ma/ sound sequence, but with the /a/ nasalized, the listener would access only words in which the /m/ sequence is followed by a nasal consonant, such as *mantener*, *manipular*, *mamar*, *manosear*, *mamullar*. According to Marslen-Wilson’s theory, this cohort of possible words is accessed even before the nasal consonant itself is heard.

In some cases, a word may be recognized at the outset, solely on the basis of its initial phonemes, that is, if it is uniquely defined by context and sensory information prior to the end of the given word. Emmorey and Fromkin (1988: 126) point out, for example, that the English word *dwindle*, can be recognized as soon as the /dwI/ segment is heard, because the /I/ immediately distinguishes the word from all other English words that begin with /dw/. Marslen-Wilson (1989) maintains that semantic constraints imposed by the context limit the possible cohorts for the base form of an inflected word, while syntactic constraints imposed by the context limit the cohorts for the fully inflected form. The group of cohorts gets smaller as more phonemes are added to the initial phonemes, so that if the word is not correctly recognized at the onset, it will be fully recognized by the time the last phoneme is heard.

Marslen-Wilson’s cohort model of lexical input is reinforced by so-called ‘tip-of-the-tongue’ states, in which subjects are often able to say what the initial sound or

syllable of a word is, even though they may be unable to think of the entire word. Such ‘tip-of-the-tongue’ states are evidence that words are stored in the lexicon by their initial sound segments, that is, by some sort of phonemic or phonetic organization. (Emmorey and Fromkin 1988: 127; Garrett 1988: 72.) It should be pointed out, however, that this observation does not mean that such a phonemic or phonetic organization is the only kind of organization system in the lexicon, however. In fact, Emmorey and Fromkin (1988), as well as Bybee (1985, 1988) argue for a componential, multidimensional lexicon, as does the present analysis. Bybee’s inter-connections in the lexicon relate words based on semantic, phonemic, and morphological bases, and there are likely many semantic sub-groupings as well.

If Marslen-Wilson’s (1989, 1999) proposed cohort effect is an accurate description of the lexical processing of input, meaning the aural perception of speech sounds, then the reverse may also be true for the processing of output as well. The same might also be true for a speaker’s deciding what comes after the initial phonemes of a word in the syntactic and semantic context of a larger utterance. The semantic sense of the utterance would allow selection of a base form, and then the syntactic and semantic context would narrow the group of cohorts further, so that the correct inflected form is produced. This notion of a cohort model for output processing is consistent with the Spanish verb model proposed in the present analysis, and is also consistent with an analogical model of morphology, as in Skousen (1989) and Bybee (1985, 1988), as illustrated in the next paragraph below.

The first phoneme of a Spanish verb is the first phoneme of the stem of the given verb, as in /a/ of /aβlaβan/ (*hablaban* ‘speak’, 3 Plural Imperfect Indicative). According

to a hypothetical reverse view of Marslen-Wilson's cohort model, the speaker would anticipate a group of possible phonetic sequences, or the group of all possible words (the 'cohorts'), that could be formed from the initial phonemic sequence of the given word. Out of context, simply as an isolated word, the group of cohorts would be large, very large, in fact, in the case of the above /a/ as an initial phoneme. This group of cohorts would contain all of the words in the Spanish language, verbs or otherwise, that begin with /a/. Within the semantic and syntactic context of an entire utterance, however, the group of cohorts would be narrowed down considerably, just as in the cohort model for lexical input processing. The feature set associated with the given verb form, along with the feature sets of the collocated words in the utterance, meaning the words in close proximity to the word in question, would determine the remaining phonemic sequence of the given verb form. The semantic and syntactic requirements of the given context, along with the collocations, help the listener to narrow the cohort group. In the present analysis and proposed model, the semantic feature sets of the Word-Formation Rules would contribute to the narrowing of the initial cohort group proposed by Marslen-Wilson's model.

This application of Marslen-Wilson's (1989, 1999) cohort model to productive as well as receptive lexical processing is consistent with the way verbs are built up in the lexicon in the present model. The initial X variables entering the blocks of the proposed model are analogous to the initial phonemes in Marslen-Wilson's model. The semantic feature sets corresponding to each rule in the proposed model specify the semantic (and sometimes syntactic) constraints that narrow the selection of the final word from the list of possible cohorts, thereby aiding in the selection of the phonemes that follow the initial

X variable. In other words, the rules in the proposed model could be regarded as a description of the mapping process described by Marslen-Wilson. The rules in the proposed model describe incremental additions of sounds to an initial sound sequence, such as /ten/ → /tenía/ → /tenías/. The feature sets in the upper half of the rule-and-feature sets are semantic contributors to the determination of the phonemic strings represented by the rules in the lower half of the blocks. The collocations, that is, the ‘neighbor words’, and the feature sets of those collocated words, as well as the syntax of the entire utterance, determine the semantic features for the given verb form and thus contribute to the determination of the final phonemes of the given verb.

Since the goal of linguistics is to discover how the mind organizes and processes complex utterances in such an inconceivably short period of time, any evidence that could possibly shorten that processing time should be considered. In Marslen-Wilson’s cohort model, the group of possible cohorts is narrowed down considerably by the semantic and syntactic context of the entire utterance, as opposed to the larger group of cohorts that the initial phonemes of an isolated word would anticipate. In the verb model presented in the present analysis, the group of cohorts is narrowed down not only by the semantic and syntactic context of the utterance and by the collocations, but also by the feature set associated with the given verb itself. It would follow, then, that the group of cohorts would thus be narrowed down even further, thus shortening the processing time.

Let’s look at an example. In the above case of /a/ as in *hablaban*, the group of cohorts associated with that initial phoneme would be very large, as mentioned above, if we consider the word *hablaban* in isolation. If we consider the collocations alone, the use of *ellos* or any plural noun, for example, may narrow the group of cohorts down to

those ending with /n/. If the utterance is *Ellos hablaban con sus amigos por teléfono* ‘They were talking with their friends on the telephone’, however, then there is no way of narrowing down the cohort group to those containing /aβa as the middle string if we consider only the collocations, unless we extend the concept of collocations to the entire context rather than just the words in close proximity to the word in question. On the other hand, if we consider the feature set associated with the verb in question, which would be [‘speak’, 3 Pl Imperf Indic], the correct form /aβlaβan/ would be produced. However, if the goal is to speed up the processing, then a combination of analogy, collocation, and rule-and-feature processing would narrow down the group of possible phonemic combinations, or cohorts, more than would any one of these processes alone would do.

For the inflectional verb model proposed in the present analysis, what is relevant about this cohort model for either receptive or productive lexical processing is that this cohort theory could add one more dimension to the processing of input or output. Adding another dimension, that is adding another aid to quickly accessing the whole correct form, could help speed up the processing of the given form by helping to narrow down the choice of possible rules from the blocks in the proposed model, and consequently, narrow down the choice of possible sound sequences that can follow the initial sound sequence entering the blocks in the model.

3.14. Conclusion: Reconciling Anderson, Bybee, and Marslen-Wilson.

This chapter has presented some general background information on morphology and morphemes, followed by discussions of Anderson’s (1988a, 1988b, 1995), Bybee’s

(1985, 1988), and Marslen-Wilson's (1987, 1989, 1999) respective theories regarding the lexicon and word formation. Aronoff's (1992) and Anderson's (1988a, 1988b) Word-Formation Rules and his Extended-Word-and-Paradigm model of morphology were presented as the formal basis for the model proposed in the present analysis of Spanish verbal morphology. This chapter has also explained the ordering of Word-Formation Rules used in the proposed model, including conjunctive and disjunctive ordering and the 'Elsewhere principle', as well as discussing lexical stems, stem alternants, lexical stem sets, and paradigms. Anderson's (1995) notion of 'a-morphous' morphology was explained as a partial theoretical basis for the proposed model, but problems with Anderson's model then pointed towards an alternate theoretical basis for the proposed model.

Skousen's (1989) Analogical Model of Language was introduced and discussed as it relates to Bybee's (1985, 1988) model of morphology and her view of "morphology as lexical organization" and as "a study of the relation between morphology and form". Bybee's theory of connections among semantically and phonologically related forms, along with the significance of lexical strength and frequency of occurrence of lexical forms, has been discussed and related to the proposed model, which also regards morphology as 'lexical organization'. Finally, Marslen-Wilson's (1987, 1989, 1999) cohort theory of lexical processing was discussed and related to the present analysis, extending Marslen-Wilson's theory to lexical production as well as lexical processing.

The theories and morphology models of Anderson, Bybee, and Marslen- Wilson all provide a formal and theoretical basis for the proposed model of Spanish inflectional verbal morphology, although some aspects of these theories may seem contradictory with

each other. Anderson's model provides the formal basis for the model, and his notion of an a-morphous morphology is essentially consistent with the proposed model, but the proposed model, unlike Anderson's, is not a generativist model. The proposed model also parts with Anderson in that the present model considers inflectional morphology to be based in the lexicon, regarding the lexicon as a dynamic, multidimensional, interconnected component of the grammar, as does Bybee.

The proposed model appears to be rule-based, since it is indeed a collection of Word-Formation Rules like those described by Anderson, but we have seen that the proposed model views rules differently than does a generativist model. The rules in the model, as in Bybee's model, are descriptions of patterns of lexical organization and connections, and these lexical rules are seen as indistinct from lexical representations. Hence, representations and rules, or items and processes, are just opposite ends of a lexical continuum. On this continuum, lexical strength is at times inversely proportionate to regularity of form, since the highly lexicalized suppletive forms at one end of the continuum are dependent upon high frequency of use for their survival. The more regular representations/rules at the other end of the continuum are less dependent upon frequency, since their regularity itself makes their forms more predictable. The rules in the proposed model that are most specific, those at the 'suppletive' end of the continuum, are the rules that must apply first, before application of the more general rules at the other end of the continuum. The application of these 'most specific' rules, or even their very existence for a given speaker, is dependent upon the lexical strength of those rules, and that lexical strength is dependent upon sufficient input to establish strong connections between the form and its function.

Marslen-Wilson's (1987, 1989, 1999) cohort model of lexical processing is related to the proposed model by suggesting that a similar process might help to explain lexical production, as well as lexical processing. The phonological segments represented linearly in the rules of the proposed model are analogous to the linear sequence of phonemes uttered in a word, and the corresponding semantic values represent part of the contextual component of Marslen-Wilson's model. In addition, Marslen-Wilson's model is relevant because it demonstrates another dimension of the dynamic, multidimensional process of lexical processing and production. If several psychological processes can operate simultaneously in receptive or productive linguistic tasks, it is conceivable that this multidimensional processing would make the tasks faster and more efficient.

The Spanish verb model proposed in the present analysis is thus a product of, and hybrid of, several models and theories. The proposed model's Word-Formation Rules, although formally based on Anderson's generativist morphology model, are actually a different conceptual illustration of Bybee's analogical model of morphology. In addition, the proposed model supports Marslen-Wilson's cohort model of lexical processing as well.

CHAPTER 4

FORM VS. FUNCTION

"Every language contains terms that have come to attain cosmic scope of reference, that crystallize in themselves the basic postulates of an unformulated philosophy, in which is couched the thought of a people, a culture, a civilization, even of an era. Such are our words 'reality, substance, matter, cause', and . . . 'space, time, past, present, future'" (Whorf 1956).

4.0. Grammatical labels are an important and necessary part of the descriptive Spanish verb model proposed in the present study, since the model depends upon the link between the *signifié* and the *signifiant* for a given verb form. The semantic values of the *signifiés* are also important for the connections among forms in the lexicon, as in Bybee's (1985, 1988) model of morpholexical organization, since forms are connected via semantic and/or phonological similarities. It is for these reasons that this chapter focuses on form and function of the linguistic sign and considers the discrepancies between grammatical labels and the reality they represent. The grammatical labels used in the proposed model, and, in fact, wherever and whenever they are used, are nothing more than a convenient shorthand notation used by linguists for specific grammatical forms of a language, just as a word is a shorthand notation for the concept it represents. Such grammatical labels are necessarily overly simplistic, since there is no way to describe in a single word the many functions that a given form may represent. This chapter explores the reality of time and tense, and then explains the various semantic values that these grammatical labels may represent for Spanish verbs in the proposed model. This analysis is a necessary precedent to the currently proposed Spanish verb model, since the semantic/grammatical notations used in the *signifié* part of the rule-and-features sets are a sort of shorthand notation for

the many possible semantic notions linked to the corresponding *signifiant* of a given Word-Formation Rule.

4.1. The reality and necessity of verb tense markers.

This chapter considers the semantic values of the labels given to the various verb forms in Spanish and discusses whether these labels represent what they appear to represent. In other words, do the Present tenses¹, for example, necessarily represent the present time in a literal sense? Does the Past Subjunctive form always literally represent an event in the past?

In fact, one could argue, as does Weinrich (1964), that tense/mood/aspect (TMA) markers may be redundant and unnecessary, since the context generally provides the time frame in the form of adverbs. Weinrich contends that “frente a los muchos matices de los adverbios de Tiempo, aparecen los tiempos del verbo como un instrumento del lenguaje bastante romo” (‘Given the many shades of meaning of adverbs of time, verb tenses seem to be a rather pointless instrument of language’) (12). [The translation is the present author’s.] For example, if one says *Caminaron por el parque ayer* (‘They walked through the park yesterday’), the tense marker in the verb could be considered redundant and unnecessary, since the adverb *ayer* provides the time reference. Nevertheless, Weinrich goes on to say that linguists and philosophers have always made a distinction between real time and verb tense, noting that Aristotle defines the verb, as opposed to the noun, as “palabra determinada según el cronos” (‘a word determined/ defined according to time’) (15). [The translation is the present author’s.] Weinrich refers to Spanish in

¹ In the present study, the various words for PN/TMA distinctions are capitalized when they are used as grammatical labels for the actual forms, while the words that refer to PN/TMA in the *real* world, in real time, are not spelled with initial capital letters.

particular when he adds that a verb is “lo que en español solo puede expresarse, o traducirse, con temporal, es decir, palabra con determinación temporal” (‘that which in Spanish can only be expressed, or translated, with a time reference, that is, a word with temporal determination’) (5). [The translation is the present author’s.]

The relevant point here is that Spanish does have a complex verb system with various ways of representing tense, mood, and aspect and that these tense designations are somewhat language specific. Weinrich (1964) points out that a verb tense is part of a verb form and part of a verb system “que varía de un idioma a otro” (*that varies from one language to another*) (15). Verb tense, like real time, can be divided into three parts, recognized at least since the time of Homer, as Weinrich points out: “Calculus, el adivino, anunciaba ‘lo pasado, lo futuro, y lo presente’ (Iliada I: 70)” (cited in Weinrich 1964:16). However, further variations of meaning, or ‘matices de significado’, can be specified by various language-specific gradations of these three basic time reference points. This chapter looks at the philosophical implications of these time references and addresses the diverse meanings of the various gradations of past, present, and future designation in Spanish. Clarification of these tense designations is important for the proposed model, since the grammatical notations (such as [+ Pres Indic], [+ Imperf Subj], [+ Cond], etc.) in the *signifié* half of the sign are really just shorthand notation for all the various *signifié* that these grammatical terms may represent for any given verb form.

4.2. Time and tense in the real world.

In Spanish, the word for ‘verb tense’ and the word for ‘time’ are the same: *tiempo*. English, however, has two distinct words: *tense* and *time*, as does German:

Tempus ‘tense’ and *Zeit* ‘time’. Thus, in Latorre’s (1968) translation of Weinrich (1964), the first chapter is titled “Los tiempos, no el Tiempo,” from which we can infer that Latorre uses the former, with a lower case <*t*>, to refer to verb tense, and the latter, with upper case <*T*>, to refer to real time. The implication of the title, and thus the topic of discussion of Weinrich’s first chapter, is that ‘tense’ (defined by Webster as ‘a distinction of form in a verb to express distinctions of time’) does not necessarily correspond to real ‘time’ in the real world.

Manuel Seco (1999) of the Real Academia Española, which is the officially recognized authority on Spanish grammar and usage, comments that ‘time’ in a literal sense is only expressed in the Indicative and Imperative moods in Spanish: “Pero el sentido puramente «temporal» de los tiempos no existe más que en el modo *indicativo* – que dispone de todos los presentes, pasados y futuros... y en el *imperativo* – que tiene siempre una orientación al futuro más o menos inmediato” (‘But the purely temporal meaning of [verb] tenses only exists in the indicative mood, which encompasses all of the present, past and future tenses... and in the imperative – which always has a future, more or less immediate, orientation’) (272). [The translation is the present author’s.]

Subjunctive mood forms, however, can refer to events in the present, past, or future, but it is the context rather than the verb form itself that determines the real-world time denoted by the phrase: “El *subjuntivo* puede «hablar» de un hecho presente, pasado, o futuro; pero esto no lo «indica» el mismo verbo, sino el sentido general de la frase o la situación en que se habla” (‘The Subjunctive mood can ‘speak’ of an action in the present, past, or future; but this is not indicated by the verb itself, but rather by the overall meaning of the sentence or the situation in which it is uttered’) (272). [The translation is the present

author's.] However, as discussed further in Section 4.3 below, Seco's (1999) generalization about the Indicative mood's literal time values does not always hold true, as Seco himself admits when he discusses the other possible uses of the Present Indicative forms in Spanish (270-271).

This observation about descriptive grammars has an important impact on the model proposed in the present study. The blocks in the model consist of rule-and-feature sets, in which the top half represents the *signifié* (the mental concept, or the set of semantic features for a given verb), and the lower half represents the *signifiant* (the physical realization, meaning the phonological material that represents the given mental concept). Since the semantic features in the upper half of the sets are grammatical terms, such as those representing tense, we need to clarify what these semantic-grammatical designations mean in terms of real time. The grammatical terms used for tense, mood, and aspect are just shorthand abbreviations for the wide range of semantic information that these designations represent in the mind of the speaker, just as any given word is just an abbreviation for the wide range of mental images associated with that word. We shall see that the designations are not really so simple as 'past', 'present', and 'future', although we are accustomed to labeling verb forms as such. The discussion below attempts to clarify the range of meanings, or *signifiés*, that are represented by the shorthand abbreviations (such as Pres Indic, Imperf Subj, etc.) given in the top half of the rule-and-feature sets of the Spanish verb model proposed in the present analysis.

4.3. Present, past, and future reconsidered.

All verb tenses are based on a present-time reference point from the viewpoint of the speaker or writer. Thus, if a speaker or writer uses a Past tense form, the listener/reader assumes that the action took place at some point of time in the past from the speaker's or writer's viewpoint. Likewise, a Present tense form theoretically conveys that the action is taking place in a present time frame from the speaker/writer's viewpoint, and a Future tense form represents an event that has not yet occurred from the speaker/writer's viewpoint.

Weinrich (1964) comments that “Después de tantas autoridades de la Antigüedad quedó asegurada la supervivencia de la triple división del Tiempo hasta la época moderna” (‘After so many ancient authorities, the tri-part division of time has persisted into the modern age’) (17) [The translation is the present author's.] From a real world perspective, however, it could be argued that there is no literal present time, or at least that the literal present is infinitesimally small. Since there is a time lapse (admittedly a very small time lapse) between what we experience through our senses and what our brains perceive and process, then everything that we think we see, hear, taste, smell, or touch in the present is actually a memory of the past, not an actual present-time experience or present-time awareness. It is only because of the human faculty of memory that we have a linear concept of time, with time divided into three parts designated as past, present, and future. If we consider time as a linear construct (at least for illustrative or, perhaps, metaphorical purposes), and if we imagine the present as a vertical line intercepting time, then everything before that hypothetical line is in the past, and everything after that line is in the future. But the intercepting line, like any line in

mathematical terms, has no width dimension. The past and the future meet at this line; hence the literal present is purely imaginary, like the points on a circle or the width of a line segment. In fact, Weinrich says that, according to Aristotle, “el presente... solo es la frontera... entre el pasado y el futuro. Tiene tan poco cuerpo como un punto o como el filo de un cuchillo” (‘The present is only the border between the past and the future. It has as little substance as a point or a knife’s edge’) (17) [The translation is the present author’s.]

A recent article by Paul Davies in *Scientific American* (September 2002) maintains that “Our senses tell us that time flows: namely, that the past is fixed, the future undetermined, and reality lived in the present. Yet various physical and philosophical arguments suggest otherwise” (40). Davies discusses the tri-part time distinction that Weinrich says has been held by human beings (at least in Western civilizations) since the time of Aristotle or before. Davies observes the following about time:

In daily life we divide time into three parts: past, present, and future. The grammatical structure of language revolves around this fundamental distinction. Reality is associated with the present moment. The past we think of as having slipped out of existence, whereas the future is even more shadowy, its details still unformed. In this simple picture, the ‘now’ of our conscious awareness glides steadily onward, transforming events that were once in the unformed future into the concrete but fleeting reality of the present, and thence relegating them to the fixed past (40-41).

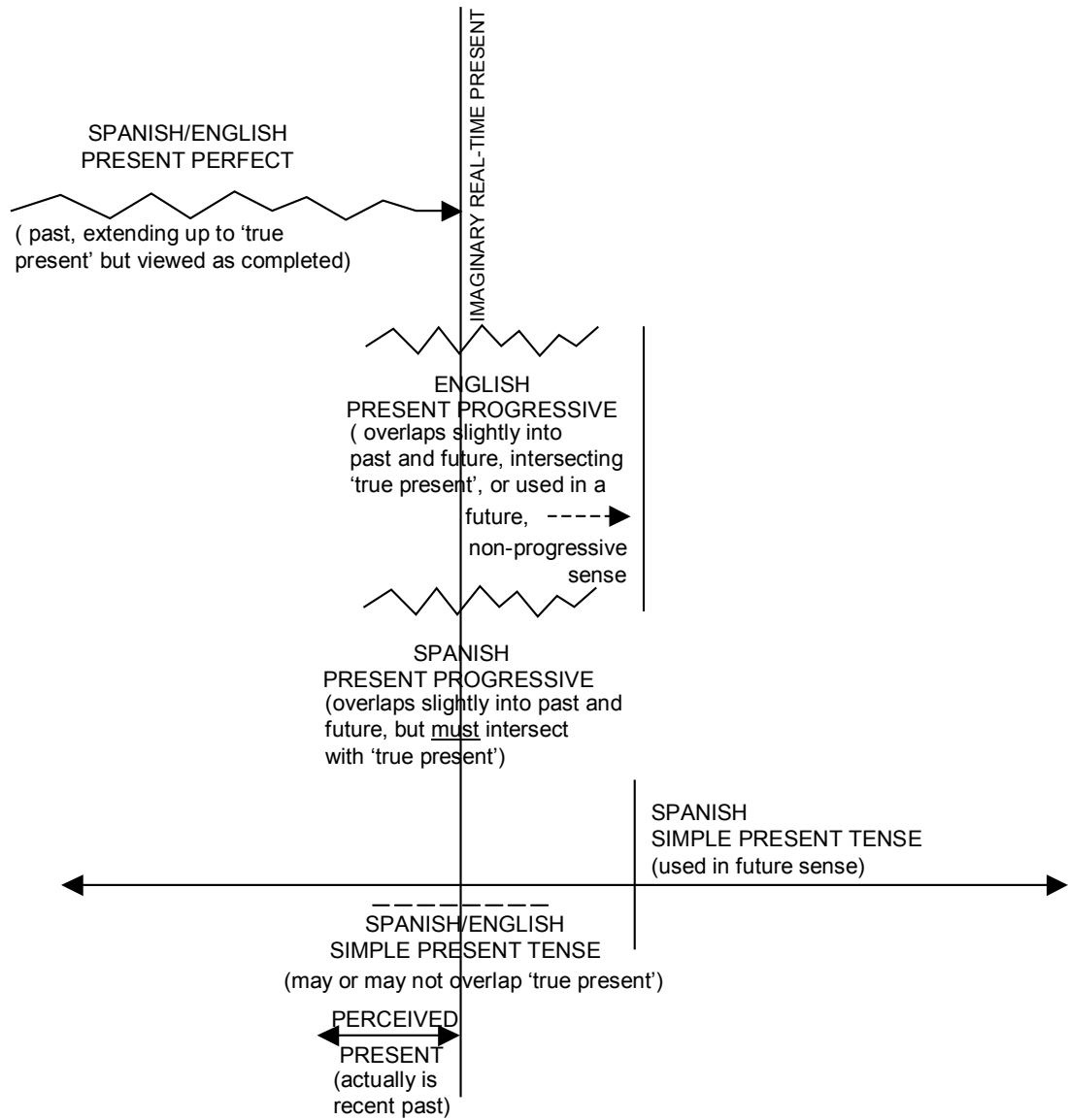
Davies also comments that the human concept of time as a flowing, linear construct may be due to the nature of memory and the human brain, maintaining that “the formation of memory is a unidirectional process – new memories add information.... We might perceive this unidirectionality as the flow of time” (47). According to Davies, Albert Einstein once wrote to a friend, “The past, present, and future are only illusions, even if stubborn ones” (cited in Davies: 41). Davies explains further that Einstein’s conclusion “stems directly from his special theory of relativity, which denies any absolute, universal significance to the present moment” (41). Perhaps it is only because of the human capacity for memory that we have a concept of time at all, since, as discussed above, what we experience as ‘present’ is really just the memory of what we have already experienced. Damasio (2002), in another article in *Scientific American*, says that “The ability to form memories is an indispensable part of the construction of a sense of our own chronology” (68).

Thus when humans speak of the ‘present’, we are being somewhat liberal in accepting part of the past and part of the future in our realm of what we consider the ‘present’. Furthermore, the degree to which we extend the notion of ‘present’ into the past and future is quite subjective and varies from one person to another, from one situation to another, and from one culture to another. Weinrich says that, of course, we have to come to some kind of understanding about what constitutes the ‘present’, since we have a common-sense notion of ‘present time’. He cites Arne Klum (1961) in this regard, explaining that one of the concepts of ‘present time’ divides “como punto matemático la línea del Tiempo en dos mitades; el otro corresponde a ‘una porción de Tiempo’ alrededor de este punto” (‘divides the time line into two halves like a

mathematical point; the other corresponds to a portion of time around this point') (cited in Weinrich: 18). [The translation is the present author's.]

In the literal sense, then, there can be no verb form to represent the 'true' present, fleeting or nonexistent as it is, and, in fact, there is no such form in Spanish. The so-called Present tense forms all have semantic values that extend beyond the literal (if existent at all) present time in the real world. They may overlap that hypothetical present time line and extend into the past or the future or both, or they may not even touch that hypothetical line at all. Seco comments that "Los tiempos verbales se aplican con frecuencia a momentos de la realidad que no les corresponden propiamente" ('Verb tenses frequently apply to moments of reality that do not properly correspond to those verb tenses') (1999: 270). [The translation is the present author's.] Spanish verb forms whose labels do not correspond to the reality of their context will be discussed further in Sections 4.4 and 4.5.

The graphic in Figure 4.1 is a representation of time as linear, flowing from left to right, with the real-time 'present' intersecting the time line in the middle. The graphic also shows the placement of the various semantic values of the Spanish Present tense forms in relation to the vertically intersecting real-time 'present'. English is also included in Figure 4.1 for comparison.



TIME & TENSE in SPANISH & ENGLISH
SPANISH/ENGLISH INDICATIVE MOOD TENSES
vs.
'REAL TIME'
WITH TIME SEEN AS A CONTINUUM

FIGURE 4.1. Real-time present vs. verb-form present for Spanish and English.

4.4. Present Indicative forms and their semantic values.

Given the above, why, then, do we have Present Tense forms, and why do we talk about the ‘present’? Indeed, philosophers and pop-psychology gurus often urge us to live in the present, to remember that the past is gone and cannot be changed, and that the future is never within our grasp. Faulkner, in his 1948 novel *Intruder in the Dust*, contends that "It's all NOW you see. Yesterday won't be over until tomorrow and tomorrow began ten thousand years ago" (190). It could be argued, however, that ‘living in the present’ is only meaningful in a metaphorical (and, perhaps, thus motivational) sense, because there is no real present. So what do we mean by ‘Present tense’ when we refer to verb forms?

The designation ‘Present tense Indicative mood’ in Spanish can have any of several possible meanings, and in the mind of the listener/reader, the decision as to which is applicable depends upon the context of the entire utterance. If one uses a Present tense Indicative mood verb form, he/she might mean a progressive action that extends a short distance into the past and a short distance into the future, thus crossing and including the imaginary intersecting line that we like to call the ‘present’. For example, if one says *Escribo en la computadora* (‘I write/am writing on the computer’), she might be describing an ongoing action spanning the immediate past and the immediate future and crossing that hypothetical real-time present line. Of course, one might alternatively use a Present Progressive form to represent this meaning, but we are currently discussing simple Present tense forms.

A second possible meaning for this shorthand designation ‘Present tense Indicative mood’ is to refer to a habitual action, one that has occurred more than once in

the past and is expected to continue to occur indefinitely in the future. Note here that it is not even necessary for the action itself to coincide with that hypothetical ‘present’ time line, even though we label this action as ‘Present tense’. We could say *Voy al supermercado todos los martes* (‘I go to the supermarket every Tuesday’), and we would use the Present tense form of the verb even though today might actually be Wednesday. What this statement means is that the speaker has habitually gone to the supermarket on Tuesdays in the past and expects to continue doing so in the indefinite future.

A third possible function of the designation ‘Present tense Indicative mood’ is to refer to an action that is expected to take place in the near future, such as *Mañana voy a la playa* (‘I’m going/ will go to the beach tomorrow’). This usage is clearly not a literal use of the so-called Present tense, nor does the English Present Progressive form *am going* in this case refer to an actual present event. In both cases, the Spanish *voy* and the English *am going*, the real time is the future, not the present. This future function of the Present tense is typically presented in textbooks as referring to the ‘immediate future’, but even this is a subjective, relative term. One can say *Voy a la playa el verano que viene* (‘I’m going to the beach next summer’), which may be several months away, or even *Voy a la playa en diez años* (‘I’m going to the beach in ten years’). Although it is clear in either of these latter two cases that we are stretching the meaning of ‘immediate future’, it is nevertheless perfectly grammatical to use the Present tense form in any of these three cases (*mañana, el verano que viene, or en diez años*) or in many other such cases as well, where the true meaning is future rather than present.

A fourth possible use of the Present tense Indicative mood forms is the so-called historical present, using the present “in place of the preterite to render the past more vivid” (Da Silva and Lovett 1965: 9), as in:

Se dio la orden de tomar el pueblo, y nuestras tropas se lanzaron al ataque.

Salvan los parapetos, se *apoderan* de la primera línea de trincheras, y a los veinte minutos, el enemigo *inicia* una retirada precipitada. [Italics are the present author’s.] (‘The order was given to take the town, and our troops rushed to the attack. They cleared the parapets, gained possession of the first line of trenches, and twenty minutes later, the enemy began a hasty retreat.’) (9). [The translation is Da Silva and Lovett’s.]

The italicized verbs in the passage above are Present tense forms, although the verbs in the first sentence (*dio* and *lanzaron*) are Preterite forms, and it is clear that the actual time frame is past, not present. This ‘historical present’ is also used in informal English, at least in some dialects, as in *So I walk into the store and this guy tells me I can’t come in barefoot*, where it is obvious that the speaker is speaking in a past time frame.

In Spanish, however, the historical present is used in standard varieties of the language and in formal written registers, as in the following passage from *España y los españoles* (‘Spain and the Spaniards’) (Cirre and Cirre 1981):

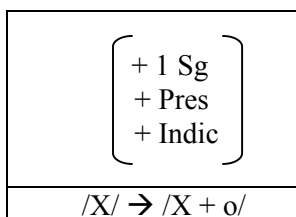
A pesar de todo, en este período sigue aumentando la riqueza nacional y la burguesía se hace más numerosa e influyente. Pero, al propio tiempo, aumentan las diferencias de nivel de vida entre las clases sociales privilegiadas y las clases proletarias, especialmente en el sur de España. (‘In spite of everything, in this period the national wealth continues to increase and the bourgeoisie becomes

more numerous and influential. But at the same time, the differences in standard of living between the privileged social classes and the proletariat classes increases, especially in the South of Spain.’) (93). [The translation is the present author’s.]

All of the verbs in the above passage are Present tense forms, even though the statements are semantically in a past time frame.

So, as illustrated by all of the above examples , the notation [+Pres Indic] used in the feature sets of the blocks of the proposed model is merely a shorthand abbreviation for any of the four concepts discussed above. If a given semantic feature set includes references (implicit or explicit) to any of the above four time concepts, represented by the simple notation [+Pres Indic] in the rule-and-feature sets in the proposed verb model, that concept will then be linked to the appropriate physical form corresponding to that feature set. In other words, the appropriate phonological form will be mapped onto the given feature set, as in the rule-and-feature set shown in Figure 4.2, which produces /aβlo/ (*hablo* ‘I speak’) when the X variable is /aβl/.

FIGURE 4.2. Rule-and-feature set yielding /aβlo/ (*hablo*, ‘speak’, 1 Sg Pres Indic).



It is only possible through the context of the entire utterance to determine which of the above four possible notions the given form semantically represents, even though the feature set designates [+ Pres Indic]. Of course, the speaker himself would not

necessarily be consciously aware of the linguistic label ‘Present Indicative’, but rather, he or she would just be aware of the semantic notion corresponding to one of the above described uses of the Present Indicative forms, and would thus mentally link this concept to a Present Indicative verb form.

4.5. Present Subjunctive forms and their semantic values.

So far this section has been discussing Present Indicative forms. However, it is also true for Present Subjunctive forms that the semantic notion linked to the form is not necessarily truly ‘present’ in real time. The Subjunctive mood can refer to an event in the present, past, or future, and it is the context rather than the verb form itself that determines the real-world time denoted by the phrase. In fact, the Present Subjunctive is more likely to carry the notion of a future time frame than a present one. If one says *Cenamos cuando lleguen los invitados* (‘We’ll have dinner when the guests arrive’), the verb *lleguen* (3 Plural Present Subjunctive) clearly has a future meaning, as does the Present Indicative *cenamos* (1 Plural Present Indicative) as well. Likewise, if a speaker says *Quiero que vengas a la fiesta* (‘I want you to come to the party’), *vengas* carries a future semantic value, although the verb *quiero* (1 Singular Present Indicative) conveys a present value in the first sense described in the last section. It is a progressive action – albeit a mental state rather than a physical action in this case – extending a short distance into the past and a short distance into the future, crossing and including the imaginary intersecting line that we like to call the ‘present’.

The author of the present analysis saw anecdotal evidence for the psychological reality of the future meaning of this so-called ‘Present Subjunctive’ form when tutoring a

native Spanish speaker in English some years ago. Upon hearing a Spanish form (such as *lleguen* or *vengas* in the above examples) called ‘Present Subjunctive’, this student insisted that that his teacher was wrong, that this was not a Present tense verb, but rather a Future tense verb. He was right, of course, in the semantic sense, although linguists and grammarians label such forms ‘Present Subjunctive’.

Some uses of the Present Subjunctive overlap that hypothetical present time line in the real world, however, just as the Present Indicative does at times. In the sentence *Dudo que ella llegue hoy* (‘I doubt that she’ll arrive today’), *llegue* is used in a future sense, but if one says *Dudo que ella hable japonés* (‘I doubt that she speaks Japanese’), then the verb *hable* is used in the same sense as the habitual usage of the Present Indicative as discussed above, meaning that I doubt that she knows Japanese, has spoken it in the past, or will be doing so in the indefinite future. On the other hand, the speaker may mean that he doubts that she is speaking Japanese at the moment (she may be speaking Korean instead), and in this case *llegue* is used in the progressive sense as discussed above. Or the speaker may even mean that he doubts that she will speak Japanese in a lecture she is to give tomorrow.

The Present Subjunctive might also appear in a historical present context if the verb in the main clause of a sentence in such a context happens to be a verb that triggers use of a Subjunctive form in the dependent clause. For example, the above Da Silva and Lovett (1965: 9) quote could have said *Se da la orden que los soldados tomen el pueblo* (‘The order is given that the soldiers take the town’) with the verb *da* in the main clause in the Present Indicative. In this case, the Present Subjunctive form *tomen* must be used in the dependent clause since it expresses indirect discourse, but the semantic value of the

utterance is past nevertheless. As Seco (1999: 273) notes, the choice of verb tense in a dependent clause is determined by the tense of the main verb of the sentence, meaning the verb in the main clause, but the form may not conform to the reality of the context.

In summary, then, the simple Present Indicative in Spanish has at least four different semantic values other than ‘true present’. Present Subjunctive forms can be used in any of the Present Indicative contexts discussed above, and, in addition, is quite commonly used in a very clear sense of real-time future meaning as well.

4.6. Past forms and their semantic values.

Preterite and Imperfect Indicative forms truly represent an action or state in the literal past, at least from the viewpoint of the speaker or writer. The forms *dio* (*give*, 3 Singular Preterite Indicative) and *lanzaron* (*‘throw’*, 3 Plural Preterite Indicative) in the Da Silva and Lovett (1965: 9) passage cited above, clearly have past semantic values. The verbs *hablaba* (*speak*, 1 Sg Imperf Indic) and *sonó* (*‘ring’*, 3 Sg Pret Indic) in the following sentence also have a past time reference: *Yo hablaba por teléfono cuando sonó el timbre* (*I was talking on the phone when the doorbell rang*). We could invent a future context for this sentence, such as a person trying to concoct an alibi for future use and saying *Les diré que yo hablaba por teléfono cuando sonó el timbre* (*‘I’ll tell them I was talking on the phone when the doorbell rang’*). But even in this case, although the action is future from the viewpoint of the speaker of the main clause, it is nevertheless truly past in semantic value from the viewpoint of the speaker in the indirect discourse within the larger utterance.

The Imperfect Subjunctive form, on the other hand, may or may not represent an action or state actually occurring in the real-world past. For example, one might say *Ella me pidió* (3 Singular Preterite Indicative) *que le trajera* (1 Singular Imperfect Subjunctive) *un vaso de agua* ('She asked me to bring her a glass of water'). In this case, whether or not the speaker actually has already brought her a glass of water is not implied by the verb. He may indeed have already done so (true past), or it may be something he has yet to do so in the future (not past in meaning), or he may not ever do so (likewise not past in meaning). In any of these three cases, the Imperfect Subjunctive form actually conveys a 'subsequent' semantic sense; in other words, it is future from the point of view of the subject (*ella*) and action (*pidió*) of the main clause. All we can infer from this context is that what is asked for (*que le trajera un vaso de agua* / 'that I bring her a glass of water') is future from the viewpoint of the main clause, but we cannot infer from the verb form whether the request has actually been carried out or not. Thus, the verb is not truly past in function, unless the action has indeed already taken place, but the verb does not indicate whether this is true or not. Bybee's (1985: 61) conjugation chart for the Spanish verb *cantar* ('sing') labels the forms 'Present', 'Past', and 'Subsequent', with the 'Subsequent' category including only Future and Conditional. This categorization seems somewhat lacking, however, since, as we have seen above, other tenses may also have 'subsequent' semantic values, as in the 'glass of water' example above.

Another common use of the Imperfect Subjunctive form is in hypothetical 'if-then' statements, with a Conditional form in the 'then' clause: *Si yo fuera* (1 Singular Imperfect Subjunctive) *rey del mundo, no habría* (3 Singular/Plural Conditional) *guerras* ('If I were king of the world, there would be no wars'). Note that the English form *were*

in this translation is also a Past form (but subjunctive mood in function), just as the Spanish *fuera* is a Past form (Imperfect Subjunctive, that is) as well. Although the context is hypothetical, and in that sense doesn't really have a real-world time frame at all, the sense of the hypothetical statement is that of the present (present, at least, in the liberal sense discussed above): *Si yo fuera rey del mundo* – 'If I were king of the world at the present time', that is to say.

The Conditional form used in hypothetical 'if-then' contexts, such as *habría* in the main clause of the example given above, also has a functionally present (in the liberal definition) time sense. For example: *Yo te ayudaría* (1 Singular Conditional) *si tuviera* (1 Singular Imperfect Subjunctive) *tiempo* (I would help you if I had time'); or *Si yo fuera* (1 Singular Imperfect Subjunctive) *rey del mundo, no habría* (3 Singular/Plural Conditional) *guerras*. Both the Conditional and the Imperfect Subjunctive forms in these examples have functionally present, although hypothetical, meanings.

The use of the Imperfect Subjunctive form in polite requests is also semantically present in real time (assuming the liberal definition of 'present' again): *Quisiera* (1 Singular Imperfect Subjunctive) *un vaso de agua, por favor* ('I would like a glass of water, please'), meaning that I would like a glass of water now, in the present or the near future. The Present Conditional form, like the Imperfect Subjunctive, is also used for making polite requests, and in this case is also in the present time frame: *Me gustaría* (3 Singular Present Conditional) *un vaso de agua, por favor* ('I would like a glass of water, please'). Just like *Quisiera* in the example given above, the verb form *gustaría* conveys the meaning that I would like a glass of water right now, in the present time, or in the near future.

In addition, the Present Conditional form may be used in a subsequent sense, somewhat like the first example of Imperfect Subjunctive usage above in *Ella me pidió que le trajera un vaso de agua* ('She asked me to bring her a glass of water'). The Present Conditional, however, is not used after a main clause verb that requires Subjunctive in the dependent clause, as in the above example, but rather when the verb in the main clause is factually stating in the past that something is going to occur in the future from that past viewpoint. For example: *Ella dijo que su hermano vendría a la fiesta* ('She said that her brother would come to the party'). This indicative mood statement of fact differs from the above example (*Ella me pidió que le trajera un vaso de agua*) that requires a Subjunctive form in the dependent clause, but both statements convey the meaning of 'subsequent' action in their respective dependent clauses. In the case of the Conditional *vendría*, just as with the Imperfect Subjunctive *trajera*, we cannot infer whether the action was actually carried out or not, only that the action is subsequent – or future – to the action of the verb in the main clause.

The various semantic notions represented by this so-called Conditional form perhaps account for the inconsistency in labeling this form in the various sources cited in the previous chapter. This form, referred to in this study as 'Present Conditional', is labeled several ways in various texts (see Chapter 7): 'Conditional' (with no tense designation); 'Presente, Modo Condicional'; 'Pospretérito, Modo Indicativo'; 'Potencial Simple o Imperfecto, Modo Indicativo'; 'Condicional, Modo Indicativo'. It is interesting to note, however, that none of these designations is sufficient to describe the actual functional use of the form. The term 'Condicional' (or English 'Conditional') works for the hypothetical usage of the form, as in *Lo haría* ('I would do it'), but it doesn't cover

the subsequent usage as in *Ella dijo que su hermano vendría a la fiesta* ('She said that her brother would come to the party'). The designation 'Pospretérito' (Seco 1999, Zatarain et al 1998) covers this subsequent usage, but does not apply to the hypothetical usage. It is also interesting to note that some texts refer to a Conditional 'Mood' (Dozier and Iguina 1999), as opposed to Indicative, Subjunctive, and Imperative Moods. Yet at the same time, some sources refer to 'Condicional, Modo Indicativo' (Real Academia Española 1973). When used in the subsequent sense, or 'Pospretérito', the Conditional form is indeed indicative in mood, as opposed to subjunctive, as seen in the examples above contrasting Subjunctive and Conditional usage in the dependent clauses of the examples. The designation of Conditional 'mood' seems to make sense only in the hypothetical usage; the same is true for the designation 'Potencial Simple' (Real Academia Española 1931).

It is interesting to note that, in general, the real-time semantic value of the Subjunctive forms is moved up one time frame into the future as compared with their Indicative mood counterparts and their grammatical tense designations. For example, in real time the Present Subjunctive is generally one time frame ahead of the Present Indicative, as well as one time frame ahead of the real-time present, at least in many contexts, as in: *Quiero que vengas* (2 Singular Present Subjunctive, but future in meaning) *a la fiesta* ('I want you to come to the party'). Likewise, the Imperfect Subjunctive in the subsequent sense (as in the *que le trajera un vaso de agua* example above), is one time frame past the Imperfect Indicative and the Preterite Indicative, even though the form *trajera* is labeled 'Imperfect Subjunctive' or 'Past Subjunctive'. Even in the hypothetical use of the Imperfect Subjunctive, it is one time frame ahead of its 'Past'

or ‘Imperfect’ designation, as in *Si yo fuera rey del mundo*, conveying a present (although hypothetical) meaning in spite of its being labeled as ‘Past’.

Likewise, the compound Pluperfect Subjunctive is really just past from the speaker’s viewpoint, as opposed to the Pluperfect Indicative, which we could call ‘pre-past’ from the speaker’s viewpoint. For example, in *Ella dijo* (3 Singular Preterite Indicative) *que había hecho* (3 Singular Pluperfect Indicative) *el trabajo* (‘She said that she had done the work’), the action of speaking (*dijo*) is ‘past’ from the viewpoint of the main clause speaker, while the action *había hecho* is ‘pre-past’ from the main clause speaker’s viewpoint, meaning that it is one time frame before the verb in the main (first) clause. The Pluperfect Subjunctive, however, is one time frame ahead of the Pluperfect Indicative, making it simply ‘past’ in real time. For example, if one says *Ojalá que hubiera llegado a tiempo* (‘I wish that he/she had arrived on time’), the real time sense is simply past, not pluperfect or pre-past. This may be hard for English-speakers to see, since English Subjunctive mood usage follows the same pattern of being one time frame ahead (in real time) of the time frame implied by the grammatical designation of the tense, as in: *I wish I were* (called ‘Past Subjunctive’, but actually present in real time), or *I wish I had been* (called Past Perfect/Pluperfect Subjunctive, but really just simple past in real time).

Figure 4.3 is a graphic representation of real time vs. grammatical time/tense, including the various Past Indicative forms as well as the Present Indicative forms shown in Figure 4.1 above. This graphic illustrates the ‘one-step-ahead’ nature of Past tense forms in Spanish.

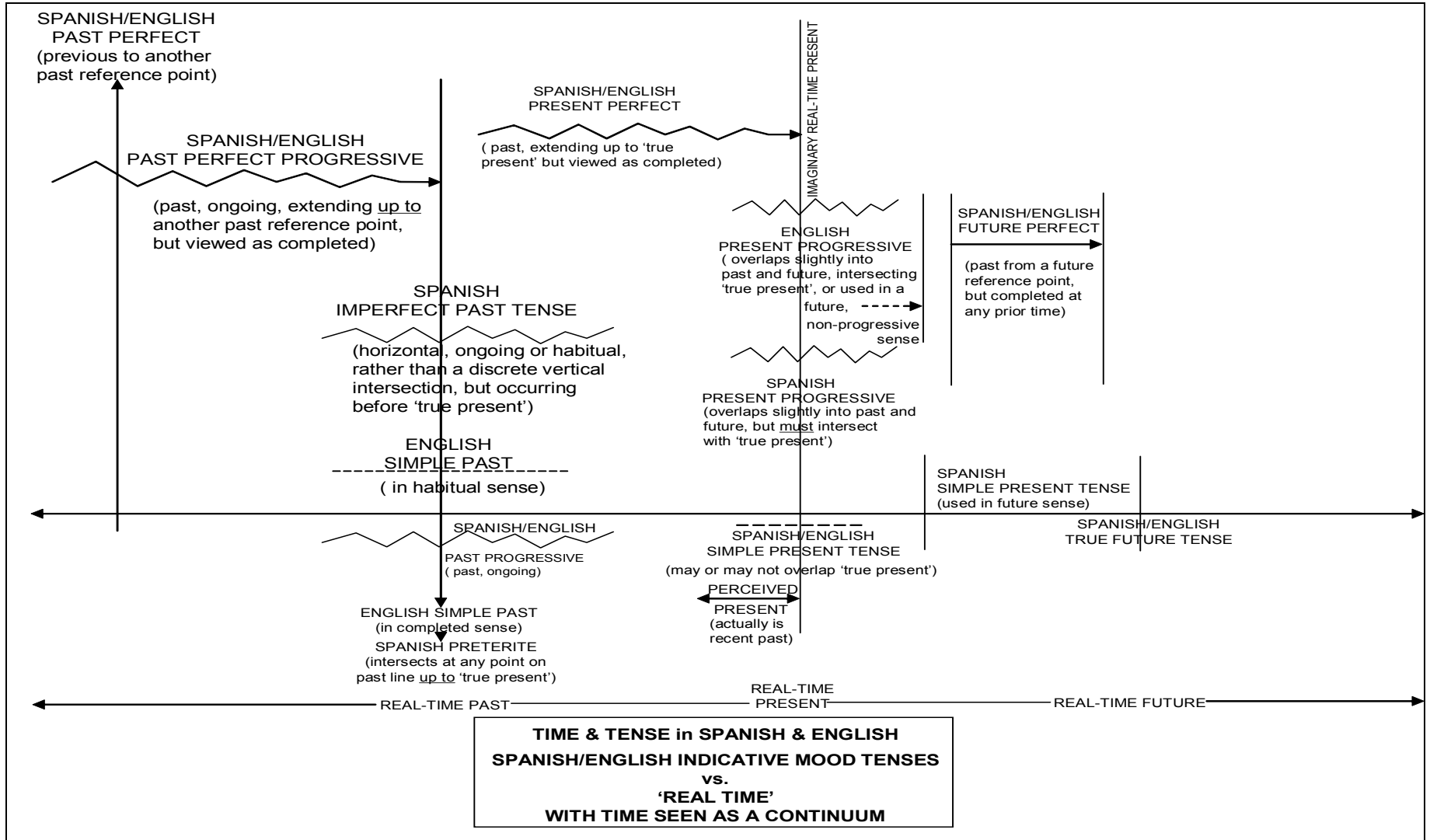


FIGURE 4.3. Real time vs. grammatical tense, showing Spanish and English Indicative mood forms on a time line.

4.7. Conclusion.

All of this discussion of real time (meaning how verbs actually function in Spanish) vs. tense time (meaning how verb forms are traditionally labeled in Spanish) demonstrates the wide range of semantic values represented by the many verb forms in the Spanish language. Therefore, when certain grammatical designations are used in the upper half of the rule-and-feature sets in the verb model proposed in the present analysis, these designations, such as [+ Pres Indic, + Imperf Subj] etc., are really just shorthand notation, or abbreviations, for the many possible semantic values linked to the *signifiant*, the other half of the Saussurean linguistic sign, represented by the bottom half of the rule-and-feature sets of the proposed verb model. In many if not most cases in Spanish, the context of a given utterance and the collocations, or other words surrounding the given verb form, are necessary to help determine the true semantic value of the verb form in question. In the model proposed in the present analysis, the semantic value of a given verb form is significant because the model depends upon the link between the *signifié* and the *signifiant* of the given form. The semantics of a given verb form are also significant to the model because of the importance of connections among semantically and phonologically similar forms in the lexicon. Thus, when grammatical notations/abbreviations are used in the feature sets of the proposed model, these notations represent a whole range of possible semantic notions, as described in this chapter, and these semantic notions are further defined by the context of the entire utterance.

PART II
THE PROPOSED MODEL

CHAPTER 5

INTRODUCTION TO THE MODEL

5.0. The model presented in full in this chapter and the next is essentially a set of rules that attempt to condense, generalize, and describe the patterns of Spanish inflectional verbal morphology as concisely and elegantly as possible. The model is intended to produce all attested forms without producing any unattested forms, and without any unnecessary redundancy or superfluous rules. The essence of the proposed model is the link between the two parts of the linguistic sign as defined by Ferdinand de Saussure: the *signifié*, which is the mental concept, the meaning, or the semantic content of a given word, and the *signifiant*, which is the mental imprint of the physical form, that is, the sound, that represents the given meaning of the word. In the present model of Spanish verbal morphology, as in Anderson's Extended Word-and-Paradigm model, the *signifié* is represented by semantic/grammatical feature sets, and the *signifiant* is represented by rules that transform a given stem into a fully inflected word.

The Spanish verb model presented in this chapter is based in form primarily on Stephen Anderson's (1982) Extended Word-and-Paradigm model of morphology. The present analysis of Spanish inflectional morphology corroborates Anderson's Extended Word-and-Paradigm model, as well as Anderson's (1995) theory of 'a-morphous morphology', since the proposed model, like Anderson's, is a 'morpheme-less' treatment of morphology. However, unlike the Word-Formation Rules in Anderson's model, the 'rules' in the present model are regarded as descriptive patterns of lexical organization

rather than as independent, generative rules occurring in other components of the grammar after lexical insertion.

Thus the proposed model is morpholexical, with the lexicon viewed as a dynamic component of the grammar rather than a static listing of lexical entries. The present analysis supports Bybee's (1985, 1988) analogical model of morphology, since the proposed model, like Bybee, regards rules and representations as essentially the same thing and emphasizes the role of semantic and phonological connections among words. The proposed Spanish verb model is also consistent with Marslen-Wilson's (Marslen-Wilson 1987, Warren and Marslen-Wilson 1987) cohort model of lexical processing, although Marslen-Wilson's model is not a morphology model itself. Marslen-Wilson's cohort model, along with Bybee's analogical model, provides a theoretical base for how words are built up in the lexicon in the proposed model, in either receptive or productive tasks.

The proposed model of Spanish inflectional verbal morphology has been introduced in small increments in earlier chapters of the present work, beginning with the brief synopsis presented in the Introduction. The current chapter now presents a broader introduction to the Basic Model for categorical/regular verbs in the Present Indicative, Present Subjunctive, Imperfect Indicative, and Imperfect Subjunctive. Then the proposed model's treatment of irregular verbs is discussed, including the categories designated respectively as phonologically variant, exceptional/regular, and idiosyncratic/suppletive.

5.1. Theoretical basis of the model.

The model of Spanish inflectional verbal morphology proposed here is based in form primarily on Anderson's (1982) Extended Word-and-Paradigm (EWP) model of morphology, explained in more detail in Chapter 3. In addition, it is theoretically based on Anderson's 'a-morphous morphology' (1995) and Bybee's analogical model of morphology (1985, 1988) and on Marslen-Wilson's cohort model of lexical processing (1987).

The present analysis corroborates Anderson's EWP model by demonstrating that his proposed blocks of rule-and-feature sets and his proposed ordering system can be successfully applied to the inflection of Spanish verbs, thereby producing all the necessary paradigms for all tenses and moods in the language. The present model also corroborates Anderson's theory of 'a-morphous' morphology, since the proposed model, like Anderson's, is a 'morpheme-less' treatment of morphology. However, the model proposed in the present study, unlike Anderson's model, is morpholexical. The Word-Formation Rules in the present model are regarded as taking place in the lexical component of the grammar, although there is multi-dimensional interaction with other components of the grammar. In the proposed model the lexicon is seen as a dynamic component of the mental grammar with many levels of analogical connections, rather than a static list of words or morphemes. Anderson (1995), on the other hand, accepts a split-morphology thesis, meaning that derivational morphology is regarded as taking place in the lexicon, while inflectional morphology takes place in another component (or other components).

The proposed model and analysis also support Bybee's (1985, 1988) analogical model of morphology, which emphasizes the role of phonological and semantic connections among words that are similar in form or meaning. The constants in the proposed model (which are, for the most part, the inflectional material of a given verb, or parts thereof) can be regarded as the parts of words that are 'similar in form or meaning' in Bybee's model. For instance, in Bybee's model there would be an analogical link among all words that end in /mos/ (as in *hablamos*, *comemos*, *escribimos*). That link is seen in the proposed model when /mos/ appears as one of the constants in the model, as in the rule /X/ → /X + mos/ for feature sets that contain [+1 Plural]. Similarly, the rule /X/ → /X + ba/ for feature sets that include [+ a-class; + Imperf Indic] shows the link among forms such as *hablaba*, *hablaban*, *amaba*, *amábamos*, *cantaba*, *cantabas*, *amabas*, and other Imperfect Indicative forms for *a*-class verbs. In other words, there are many verbs in Spanish, each with a different semantic and phonological value for the X variable, that can be mapped onto each of the phonological strings that make up the constants in the proposed model, just as there are many layers of connections among words in Bybee's model.

Marslen-Wilson's (Marslen-Wilson 1989, Warren and Marslen-Wilson 1987) cohort model, along with Bybee's analogical model, also provides a partial theoretical base for how words are built up in the lexicon in the proposed model, in either receptive or productive tasks. The present Spanish verb model is partially consistent with Marslen-Wilson's cohort model of lexical processing, although Marslen-Wilson's model is not a morphology model itself, and although Marslen-Wilson accepts a morphemic, decompositional view of the lexicon, rather than a whole-word view (1987: 102-103).

Marslen-Wilson's (1989: 7) cohort model of lexical processing contends that words are often recognized upon the listener's hearing the initial phonemes of the word. According to Marslen-Wilson's model, as soon as a listener hears the initial phonemes of a word, acoustic-phonetic information is mapped onto the lexical level during the process of lexical access and selection. Upon hearing the initial phonemes of a word, the listener immediately accesses a group of possible words ('cohorts') that the whole word could be. Marslen-Wilson maintains that "the speech signal is rich in what we can call partial information – that is, anticipatory cues to the identity of an upcoming segment" (9).

This initial group of cohorts could be very large if the word is an isolated word out of any meaningful situational context. In Marslen-Wilson's view, all words in the lexicon with the same initial sound or sound sequence would be accessed upon hearing a given initial sound or sounds. It is not just phonetic cues that narrow the group of possible cohorts for an initial sound sequence, however. The semantic and syntactic requirements of the given context, along with the collocations, or words surrounding the given sound sequence, also help to narrow the cohort group. In actual communicative speech, then, the context and collocations narrow the selection of possible cohorts from among this larger group.

This productive application of Marslen-Wilson's cohort model is consistent with the way verbs are built up in the lexicon in the present model. The initial X variables entering the blocks of the proposed model are analogous to the initial phonemes in Marslen-Wilson's model. The semantic feature sets corresponding to each rule in the proposed model specify the semantic (and sometimes syntactic) constraints that narrow the selection of the final word from the list of possible cohorts, thereby aiding in the

selection of the phonemes that follow the initial X variable. In other words, the rules in the proposed model could be regarded as a description of the mapping process described by Marslen-Wilson. The rules in the proposed model describe incremental additions of sounds to an initial sound sequence, such as /ten/ → /tenía/ → /tenías/, as in *Tú tenías diez dólares al entrar en la tienda* ('You had ten dollars upon entering the store').

For the inflectional verb model proposed in the present analysis, what is relevant about this cohort model for receptive or productive lexical processing is that this cohort theory could add one more dimension to the processing of input or output. Adding another dimension, that is adding another aid to quickly accessing the whole correct form, could help speed up the processing of the given form by helping to narrow down the choice of possible rules from the blocks in the proposed model, and consequently, narrow down the choice of possible sound sequences that can follow the initial sound sequence entering the blocks in the model.

5.2. Verb categories analyzed by the present study.

Skousen, in his *Analogical Modeling of Language* (1989), describes "three basic types of behavior" that a theory of language must account for: categorical, exceptional/regular, and idiosyncratic (11). 'Categorical' forms are those with completely regular patterns, such as English past forms with /ed/. 'Exceptional/regular' forms, on the other hand, are those with regularities in their irregularity, we might say. Forms in this latter group, although irregular compared to the first group, exhibit patterns and regularities of their own, such as so-called strong English verbs with past forms that

can be grouped together based on similarity of form, such as *sang/rang*, *blew/flew/grew*, *bought/fought/brought*.

Using Skousen's three-member classification system as a starting point, the present study discusses the following categories of Spanish verbs, categorized according to the patterns and similarities of their respective paradigms:

- 1) 'categorical/regular' verbs: verbs traditionally viewed as completely 'regular';
- 2) 'phonologically variant' verbs (a subset of Skousen's 'exceptional /regular' forms): verbs with regular inflectional affixes, but with vowel alternations or diphthong/mid-vowel alternation in the stem; traditionally called 'stem-changing verbs';
- 3) 'exceptional/regular' verbs: verbs with stem alternants that have some phonological similarities with each other, but are not of the above vowel-alternation type; commonly called 'irregular' verbs in teaching textbooks, although (like the phonologically variant group) the irregularity is primarily in the stem;
- 4) 'idiosyncratic or suppletive' verbs: verbs with forms that do not resemble the regular stems in the given verb's lexical stem set;
- 5) 'lexical gaps': verbs with nonexistent, unattested forms that cause 'gaps' in their respective inflectional paradigms;
- 6) 'ill-formed tokens': forms deemed by most speakers to be 'ungrammatical' (or are ungrammatical in the standard form of the language), although they are well-attested in the language as a whole and conform to some regular pattern.

5.3. How the proposed model works.

The essence of this model of Spanish verbal morphology is the link between the two parts of the Saussurean linguistic sign, that is, the link between the *signifié*, represented by semantic/grammatical/collocation feature sets, and the *signifiant*, represented by rules or patterns that transform the initial phoneme or phonemes into a fully inflected word. As in Anderson's model, the rules in the model presented in this analysis are based on constants and variables. The constants are the phonological material that many words have in common – the phonological strings commonly referred to as grammatical morphemes. The variable for any given word, at least at the outset of the model, is typically the phonological string that carries the semantic content of a given word, such as /aβl/ for 'speak', as in *hablo, hablas, habla..*

The term 'content-bearing' refers to the part of a word that conveys the word's semantic content, that is the variable in the proposed model, as opposed to the part or parts commonly called 'grammatical morphemes', although so-called grammatical morphemes could actually be considered as part of the overall semantic value of a given word, as discussed in Chapter 2.3. For example, there is a semantic difference between the mental image of one apple and the mental image of a dozen apples. Likewise, there is a difference in meaning between a verb that designates action occurring at the present moment and a verb that designates an action that took place a hundred years ago. Just because it is possible to separate some words into discrete, identifiable parts, this does not necessarily mean that words are inherently separable in terms of their semantic and grammatical functions, nor does it necessarily imply that grammatical functions in a word are psychologically different from 'true' semantic functions.

In the production or comprehension of a spoken word, the speaker or listener begins with an initial sound or sound cluster. For most verbs in Spanish, this initial segment is the so-called content-bearing part of the word, such as /kant/ in a form such as *cantábamos* ('sing', 1 Plural Imperfect Indicative). In the model proposed in the present study, the true content-bearing segment for any given verb is the variable that enters the first block of the model. However, as the model progresses through the blocks of rules and features sets, the variable part of the model collects more phonological material. Hence, the variable entering successive blocks is usually altered in some way from the original variable, and the resulting final form of the word is the entire string of phonological material that exits the last block after picking up one or more of the constants, which in most cases is the phonological material commonly referred to as grammatical morphemes. We might say that the stem 'grows' as it progresses through the blocks of Word-Formation Rules, but it is no longer a 'stem' in the traditional sense after it adds additional phonological material and proceeds as the new variable to the next block of rules. It is for this reason that it is advantageous to use the term 'variable' in the model instead of 'stem'.

Many verbs in Spanish have more than one stem in their respective lexical stem sets. Therefore, it must be specified at the outset which of the stem alternants is to apply to what rules in the model as the X variable. The Appendix of the present study specifies the stem alternants for a sample list of Spanish verbs, along with their respective feature sets. Each verb's group of stem alternates, with mutually exclusive feature sets for each alternant, specifies which stem will serve as the X variable in the rules of the model.

The purpose of the model is to recognize and organize as elegantly as possible the patterns that exist in the Spanish verbal system as a whole, in order to generalize as much as possible and thus simplify the description of the system, thereby suggesting how this complex system might be organized and processed in the mind of a speaker. Although the investigative process involves looking for generalizations and finding the lowest common denominators in the system and working from the most general to the most specific, the model itself is essentially organized from the most specific to the most general, as in Anderson's model (1992). The rule-and-feature sets in the model are disjunctively ordered within each block, also as in Anderson's model. This means that the rule-and-feature sets do not necessarily apply in the order in which they appear in a given block, but rather, from most specific to least specific. The rules are mutually exclusive within a given block; that is, only one rule within a block is possible. Once a rule applies in a given block, the resulting form exits the block, and this string of phonological material is the new /X/ variable that proceeds to the next block.

The distinction between square brackets [] and curly brackets { } is important for understanding the model. When a feature or a group of features is enclosed within regular brackets, the brackets mean that all of the features within the brackets must apply in the given rule. For example, if a rule includes the feature set [+ Pres, + Indic, + 1 Sg], then the resulting form must have all of these features; thus the resulting form will be a First Person Singular Present Indicative form such as *hablo* or *escribo*. When more than one feature is enclosed within curly brackets, however, the enclosed features are mutually exclusive, meaning that only one of them can apply for any given form. For example, if a rule includes the feature set {1 Pl; 2 Pl}, then only one of the enclosed features can apply

for any given form. There may be a combination of regular and curly brackets, as in the following rule from one of the blocks: [+ {1 Pl; 2 Pl}; + *i*-class]. This feature set means that the resulting form will be either 1 Plural or 2 Plural, and, in addition, it must also be classified as *i*-class.

Anderson uses a binary notation system, using [+] features and [-] features in the feature sets. The present model uses only the [+] notation whenever possible for the sake of simplicity, with the assumption that, for instance, a [+ Pret] notation in a given feature set implies a [-] (negative) value for tenses other than Preterite. However, in some cases it may be necessary to specify a [-] value, such as in one of the rules in the Preterite models in Figures 6.5 and 6.6. The negative notation is necessary to block the production of an incorrect form, as will be seen in the discussion in that section.

Let's look at the first block in the model, shown In Figure 5.1. Any stem, or, in a more general sense, any word-initial set of phonemes, may enter the block, but the given stem (or phonological string) may only proceed further through the blocks if its own semantic and grammatical specifications do not contradict one of the feature sets in the block. The initial phonemes of the verb *hablar*, for example, are /aβl/, and its associated feature set would specify a semantic value of 'speak', as well as being specified as *a*-class. Since *hablar* is a categorical/regular verb and has no alternate stems as some verbs do, there would be no other specifications in its feature set that would bar the application of the rules in Block I, except for the *a*-class specification of *hablar*, which would bar the application of any set that includes *e/i*-class in its feature set. On the other hand, for a verb such as *tener*, each of its alternate stems or initial phoneme groups (/ten/, /tyen/, /teng/, /tuβ/, /tendr/) would have grammatical feature specifications associated with the

given stem as part of its associated semantic/grammatical/collocation feature set, in addition to its semantic value of ‘have/possess’. For example, the stem /tuβ/ would include in its feature set the specification {[+ Pret Indic] [+ Imperf Subj]}, meaning that a form containing the /tuβ/ stem will be either Preterite Indicative or Imperfect Subjunctive, but not both, of course. (We know this because curly brackets are used in the notation.)

For any categorical/regular verb, there is one and only one stem available to enter the blocks for the given verb: the infinitive form minus the /ar/, /er/, or /ir/ ending. It should be noted, however, that the infinitive form is not necessarily considered a ‘base’ form from a psychological standpoint, and is, in fact, generated in the proposed model by rules presented in the next chapter. Bull comments that “Spanish verb forms are always made up of three parts, a stem plus two suffixes,” adding that “This is readily observable in the infinitives and both participles” (1965: 112). Thus, the infinitive itself begins with a stem. Furthermore, Alvar and Pottier (1983) maintain that the Infinitive is not really a verb form in the strict sense of the term ‘verb’, since the Infinitive is unspecified for tense and person and actually functions as a noun (215).

The Infinitive form is used here only as a reference for identifying the default stem, meaning the stem that applies if there is no alternant stem specified. For the purposes of the present analysis, it is convenient to refer to the Infinitive stem as a reference point, since the Infinitive form is the form typically listed in dictionaries and vocabulary lists. This default stem is the initial variable for any given verb in the proposed model unless there is a rule that specifies another stem, hence another variable. For categorical/regular verbs, such as *hablar*, *comer*, or *vivir*, however, we can proceed

through the first block with no problems. The first block of the proposed basic verb model is shown in Figure 5.1, where any ‘regular’ stem may enter the model.

FIGURE 5.1. First block of the proposed Basic Model for Spanish verbs.

1) $\left[\begin{array}{l} + 1 \text{ Sg} \\ + \text{Pres} \\ + \text{Indic} \end{array} \right]$	2) $\left\{ \left[\begin{array}{l} + \text{Pres} \\ + \text{Subj} \\ + a\text{-class} \end{array} \right] \right\}$ + <i>e/i</i> -class	3) $\left\{ \left[\begin{array}{l} + \text{Pres} \\ + \text{Subj} \\ + e/i\text{-class} \end{array} \right] \right\}$ + <i>a</i> -class	4) $\left[\begin{array}{l} + \text{Imperf} \\ + \text{Indic} \\ + e/i\text{-class} \end{array} \right]$
/X/ → /X + o/	/X/ → /X + e/	/X/ → /X + a/	/X/ → /X + ía/
5) $\left[\begin{array}{l} + \text{Imperf} \\ + \text{Subj} \\ + e/i\text{-class} \end{array} \right]$	6) $\left[\begin{array}{l} + \{ 1 \text{ PI} / 2 \text{ PI} \} \\ + i\text{-class} \end{array} \right]$		
/X/ → /X + ye/	/X/ → /X + i/		

In the model proposed here, the individual numbered squares with the ‘formulas’ in them are the rule-and-feature sets of the model. A careful examination of the block shown in Figure 5.1 shows that all of the rules in the block are mutually exclusive. Rule 1, and *only* Rule 1, yields the 1 Singular Present Indicative form, such as *hablo* or *vivo*. This form ending with /o/ is the most specific in the block, since, unlike the other forms in the block, its feature set contains person and number (PN) as well as tense and mood (TM). After Rule 1 applies, the resulting form exits the block and proceeds to the next block, and this exiting form is the new variable for the next block. No other rule can apply within this sample block; hence, we say that the rule-and-feature sets in the model

are disjunctively ordered within each block. In fact, for a 1 Singular Present Indicative form, no other rule in any successive block has a matching feature set. Therefore, the form produced by Block I is the final form, such as *hablo, canto, bebo, vivo*.

Another rule in the block, Rule 4 (only Rule 4), yields the basic Imperfect Indicative form for *e/i*-class verbs, such as *bebía* or *vivía*. Person and number are not specified at this point, since the model seeks the lowest common denominators and aims for as many generalizations as possible. The phonological material that signifies PN will be added later in another block, except in the case of the 1 Singular and 3 Singular forms. The 1 Singular and 3 Singular forms (for the Imperfect Indicative for *e/i*-class verbs) do not need to be specified as such in any of the rule-and-feature sets, since the form (such as *bebía* or *vivía*) produced by Block 1 is also the final form in the case of 1 Singular and 3 Singular Imperfect Indicative for *e/i*-class verbs. So if no further phonological material is added on, as specified by further rule-and-feature sets in subsequent blocks, then the form that exits from Block I (such as *bebía* or *vivía*) then exits the final block as is, as a sort of default form, with a ‘zero morph’, if we wish to use traditional terminology. We might formulate a rule for such 1 Singular and 3 Singular forms, specifying a zero morph, such as $/X/ \rightarrow /X + \emptyset/$ for the feature set [+ 1 Singular] or [3 Singular]. However, in the interest of simplifying the rules and not including any superfluous rules, it is more economical not to include such a rule, but to allow the 1 Singular and 3 Singular forms to exit as default forms, unspecified for PN. It is interesting to note that according to Bybee (1985: 50-51) and Spencer (1997), “In many inflectional systems, 3rd person singular is the default P/N specification” (Spencer 1997: 218).

While the rules within each block are ‘disjunctively’ ordered, the blocks themselves are ‘conjunctively’ ordered, meaning that if a rule can apply in a given block, then it must apply. For example, Rule 1 of Block 1 above must apply for a feature set that includes [1 Sg, + Pres, + Indic], producing regular 1 Singular Present Indicative forms such as *hablo, canto, bebo, vivo*). Once this rule has applied, no other rules in that block can apply, and the resulting form then must exit the block and proceed to the next sequentially ordered block, and so on through all of the blocks in the model.

The Elsewhere Principle applies within all of the blocks, meaning that application of a more specific rule blocks that of a later more general one (Anderson 1995. p. 132), thus assuring mutually exclusive affixation, as with the 1 Singular Present Indicative form of Rule 1 of Block 1 above. Since, for example, the specification of the feature set [+1 Sg, + Pres, + Indic] is more specific than just the specification of [+ *a*-class], the rule /X/ → /X + o/ takes precedence over the more general rule in Block I that specifies [+ *a*-class]. Therefore, only the more specific form – in this case, a 1 Singular Present Indicative form such as *hablo, canto, bebo, vivo* – can exit a given block. This more specific rule/form blocks application of the more general rule stating /X/ → /X + a/ for *a*-class verbs.

Anderson (1995: 132) points out that according to the Elsewhere Principle, a specific rule can only block the application of a more general rule in the same block or a later more general rule in a subsequent block. In other words, a more specific rule in a later block does not block the application of an earlier more general rule, that is, a rule in a previous block that has already applied. There does not at this time appear to be any instance in the present proposed model for which such a stipulation would be necessary,

but it is nevertheless worth clarifying in case of further developments in the proposed model.

The top sections in each block represent semantic/grammatical/collocation feature sets, with individual features in the feature sets enclosed in brackets and represented by the notation [+ Grammatical Function]. It is important to note that the [+] notation is not used anywhere in the present model to mean ‘plus’, in the sense of adding something onto a word, such as an affix. The [+] notation is used here in the same way it is used in syntactic and phonological theory, that is, to designate a positive value for a given feature. In the present model, the features are grammatical features, and the [+] notation means that a given form has a given feature, such as a particular tense, as opposed to some other tense for which the form would have a [-] value. Within these brackets, all of the features in that set must apply. However, as mentioned earlier, curly brackets, as in {+ Infin, + Fut, + Cond}, indicate mutually exclusive features within those brackets.

Each of the feature sets given in the blocks is part of the semantic feature set conceptually associated with the *signifié* of a whole inflected word. A feature set for an actual word would actually include all of the semantic features associated with the given word as well as the so-called ‘grammatical’ features. In this morphological analysis, however, only the grammatical features (including verb class) are listed in the feature sets, since the phonological material corresponding to these ‘grammatical’ features are the constants in the model, while the ‘content’ part of the word is the variable in this model. It should be noted, however, that in the model proposed here, these grammatical features are seen as part of the whole set of semantic features, not as something separate, as explained earlier in Chapter 2.3. This stipulation is primarily for the purpose of

emphasizing that there is not always a one-to-one correspondence between function and form, and because the proposed model accepts a whole-word view of the lexicon. As discussed earlier, just as we are not accustomed to breaking down the so-called ‘content’ part of a word into discrete morphemes, and in fact it is not possible in some words (such as *dog* or *casa* ‘house’), it is not always possible to isolate and break down inflectional material in this way either. For example, Spanish Preterite affixes are not easily separable in terms of person, number, tense, mood, and aspect, but rather, the indicators for PN and TMA are fused together.

It should also be noted that the designation of verbs as *a*-class, *e*-class, or *i*-class is considered here to be part of the semantic/grammatical/collocation feature set of a given verb, and would be considered part of a native speaker’s linguistic competence, although the average speaker himself probably wouldn’t consciously divide verbs into classes in this way. Spencer comments that “deciding which ... formatives are to be regarded as inflections proper” is “a theory-internal choice” (1997: 217). In discussing Italian verbal morphology, he says that

The theme elements [that is, vowels, as in Spanish verb class based on the theme vowels *a*, *e*, *i*] are often thus regarded as derivational, stem-forming suffixes, which would suggest that their presence should be accounted for by other rules (namely, derivational rules in the lexicon). However, in a certain sense these theme vowels are part of the paradigm and interact with other, genuinely inflectional, suffixes in complex ways. Therefore, we shall assume that it is our morphological rules which introduce the theme vowels (217).

In a sense, then, it is somewhat arbitrary and a ‘theory-internal’ choice to regard vowel class as part of the Word-Formation Rules for Spanish inflectional morphology.

If there is not some sort of mental grouping of verbs according to vowel class, then it is hard to explain how the ‘correct’ vowel comes to be used in the verb endings. In Bybee’s (1985, 1988) analogical model, verbs would be grouped together according to vowel class, or theme vowel, by virtue of their phonological similarities. In Bybee’s model, it doesn’t matter whether there is any semantic significance to the vowels, since her model is based on semantic and/or phonological connections among forms. However, connections are stronger where there is semantic and phonological identity. Furthermore, as explained in Chapter 3 of the present analysis, Bybee’s gradient view of inflection vs. derivation means that both of these morphological processes occur on a continuum between lexical expression on one end and syntactic expression on the other, with both inflection and derivation somewhere in between. In this view, then, it is not so important to distinguish between derivation and inflection regarding the theme vowels.

The reason the term ‘semantic/grammatical/collocation features’ is used here in reference to the feature sets of the Word-Formation Rules is that although the designation of verb class has no descriptive content and does not convey any real semantic information, verb class does implicate certain specified forms that must be collocated with the given form in question. The notion of ‘collocation’ used in the present analysis extends to neighboring phonological material that may be part of a given word, as well as to whole words that are neighbors of a given form or word, such as a particular preposition that typically follows a given verb. Collocations, at least in this broad sense of the word, are relevant to the present morphological analysis in that they represent

regular patterns of association among forms in the lexicon. For example, the segment /*a*/ in Imperfect Indicative forms (in Rule 4 of Figure 5.1) requires that it be collocated only with forms that fit in with the *e/i*-class verb patterns.

If the word ‘semantic’ is seen in a literal sense, where every semantic feature has descriptive content and represents a real conceptual referent, then one might question including verb class in the sets of semantic features. However, verb class based on theme vowel might be considered as a sort of subcategorization feature of a given verb. Subcategorization specifications are considered by Chomsky, Radford and other theoretical linguists (Radford 1997: 59-69) to be a part of the semantic (or grammatical) feature sets of words, although it might be argued that at least some of these subcategorization specifications are not literally part of the semantic feature sets any more than verb class based on theme vowel might be. If we say that a given verb is subcategorized as transitive, for example, it means that it must be followed by a complement, as opposed to intransitive, meaning that it cannot have a complement. Although the feature of transitivity is related to meaning in that a complement may (by definition) complete the meaning of the given verb, the basic semantic content, or meaning, of the verb generally remains constant no matter what the complement is. For example, if one says *Yo compré* (‘I bought’), the overall meaning of the utterance is incomplete unless the speaker specifies what it was that he bought. However, whatever the complement is, whether I bought a Mercedes or a pack of chewing gum, makes no difference to the meaning of the verb itself, but rather just to the complete utterance. Of course, if the verb is used in an idiomatic or metaphorical sense, such as *I bought the witness*, the meaning is different, but that is not a factor of the verb’s subcategorization

regarding transitivity. Likewise, the English verb *bathe* can be either transitive or intransitive, but the basic semantic content of the verb is the same in either case. The fact that *bathe* may be subcategorized as transitive or intransitive and that the Spanish verb *comprar* requires a complement may be no more truly semantic in nature than verb class would be for a Spanish verb. So we could perhaps say that the verb *comprar* is subcategorized as [+ transitive] and [+ *a*-class]. Furthermore, for the purposes of the present analysis, it could simply be stipulated at the outset that in the present model of Spanish verbal morphology, verb class is considered part of the semantic feature set, thereby perhaps slightly relaxing the definition of ‘feature set’ for lack of better terminology.

The bottom half of each block represents the *signifiant* part of the linguistic sign. It is the phonological realization of the semantic and grammatical features of the *signifié* represented in the block above. (Actually, to Saussure the *signifiant* is the ‘mental imprint’ of the physical realization.) The /X/ variable entering the first block, enclosed between slashes as phonemes are conventionally notated, typically represents the content-bearing stem or lexical variant, although it is not necessary to see the /X/ variable in this way. All parts of a word are just parts, just phonological strings, as in Anderson’s a-morphous, morpheme-less morphology. In the present verb model, semantic features and grammatical features are all seen as having semantic value, and the phonological strings are just seen as variables or constants that are collocated with certain specified other phonological strings. As more phonological material is added to the word, more semantic/grammatical/collocation features are also added, but these features may not be entirely separable either in the top (*signifié*) half of the set or in the bottom (*signifiant*)

half of the given rule-and-feature set, as they would be in a hypothetical strictly agglutinative language. The notation in the lower half of each of the sets in the blocks represents one of the morphological rules that transform the initial phonological string into the fully inflected phonological form of the word; thus the symbol \rightarrow means ‘becomes’ or ‘is realized as’.

Although all of the feature sets in a given block are mutually exclusive, meaning that only one can apply, it is not necessary that one of the sets apply in a given block. If none of the feature sets in a block applies, then the phonological form entering the block exits the block unchanged. In this model, therefore, there is no need to specify a zero-morph, as mentioned above, such as in the 1 and 3 Singular forms of several of the Spanish verb tenses: 1 Singular and 3 Singular Imperfect Indicative *a*-class *hablaba*, for example, or 1 Singular and 3 Singular Imperfect Indicative *e/i*-class *bebía* or *vivía* mentioned above. If there is never a specification in the feature sets for 1 Singular or 3 Singular after the block in which *hablaba*, *bebía* or *vivía* are produced, then these forms will be the final forms for the respective verbs. When one of the rules does apply in any given block, however, then the form that exits that block enters the next block as the new variable /X/ and then undergoes some changes, adding phonological material, so that a different form exits the final block. In the case of the above three forms, *hablaba*, *bebía* and *vivía*, later blocks have other PN specifications in their feature sets, such as the one that adds /s/ for 2 Singular and the one that adds /mos/ for 1 Plural, thus producing *hablabas*, *bebías* and *vivías* for 2 Singular Imperfect Indicative and *hablábamos*, *bebíamos* and *vivíamos* for 1 Plural Imperfect Indicative.

5.4. Basic model: Present Indicative, Present Subjunctive, Imperfect Indicative, and Imperfect Subjunctive.

For all three Spanish verb classes (*a*-class, *e*-class, *i*-class), the basic model presented in Figure 5.2 is capable of producing the following forms for all categorical/regular verbs: Present Indicative, Present Subjunctive, Imperfect Indicative, and Imperfect Subjunctive. The following abbreviations are used in the model: Pres, Indic, Subj, Imperf, as well as the numerals 1, 2, and 3 for first, second, and third person designation and Sg and Pl for singular and plural number designation. At the outset, upon entering the blocks, /X/ = the content-bearing stem, which for most Spanish verbs is the infinitive form minus the /ar/, /er/ or /ir/ ending. Stem allomorphy and suppletion will be discussed in more detail in later sections of this chapter and the next.

For the sake of easy recognition of words by the reader, orthographic representations may be used in discussing the model, but phonological representations are used in the model itself, including the use of /y/ for the high front glide instead of the orthographic <*i*>. This will be important later in the discussion of the possible influence of *yod* (the semi-vowel or glide [y]) in producing some phonological variants of certain verbs. To avoid confusion, conventional spelling (with no slashes, of course) is used when giving an orthographic representation in the text, as opposed to using slashes for the phonological representation of a word or string of phonemes.

FIGURE 5.2. Basic Model: Present Indicative, Present Subjunctive, Imperfect Indicative, and Imperfect Subjunctive.

Block I

1) $\left[\begin{array}{l} + 1 \text{ Sg} \\ + \text{Pres} \\ + \text{Indic} \end{array} \right]$	2) $\left\{ \left[\begin{array}{l} + \text{Pres} \\ + \text{Subj} \\ + a\text{-class} \end{array} \right] \right\}$ + <i>e/i</i> -class	3) $\left\{ \left[\begin{array}{l} + \text{Pres} \\ + \text{Subj} \\ + e/i\text{-class} \end{array} \right] \right\}$ + <i>a</i> -class	4) $\left[\begin{array}{l} + \text{Imperf} \\ + \text{Indic} \\ + e/i\text{-class} \end{array} \right]$
/X/ → /X + o/	/X/ → /X + e/	/X/ → /X + a/	/X/ → /X + ia/

5) $\left[\begin{array}{l} + \text{Imperf} \\ + \text{Subj} \\ + e/i\text{-class} \end{array} \right]$	6) $\left[\begin{array}{l} + \{ 1 \text{ Pl} / 2 \text{ Pl} \} \\ + i\text{-class} \end{array} \right]$
/X/ → /X + ye/	/X/ → /X + i/

Block II

1) $\left[\begin{array}{l} + \text{Imperf} \\ + \text{Indic} \\ + a\text{-class} \end{array} \right]$	2) $\left[\begin{array}{l} + \text{Imperf} \\ + \text{Subj} \end{array} \right]$
/X/ → /X + ba/	/X/ → /X + ra/

Block III

1) $\left[+ 2 \text{ Sg} \right]$	2) $\left[+ 1 \text{ Pl} \right]$	3) $\left[+ 2 \text{ Pl} \right]$	4) $\left[+ 3 \text{ Pl} \right]$
/X/ → /X + s/	/X/ → /X + mos/	/X/ → /X + is/	/X/ → /X + n/

Let's look at an example of a form produced by these blocks of rule-and-feature sets. For a feature set specifying the semantic notion of 'dance', for example, with a TMA value of Present Subjunctive and a PN value of 2 Singular, the stem entering the first block would be /bayl/, the only stem in the lexical stem set for the verb *bailar*. This stem would be specified as [+ a-class] and would enter Block I as the X variable, where the only set that matches is the set in Rule 2, stating that /X/ → /X + e/. Thus the form exiting Block I is /bayle/. This form now enters Block II as the new X variable, where we see that there are no matching feature sets. The form /bayle/ then exits the block with no changes and then enters Block III as the X variable for that block, where Rule 1 has the matching feature set of [2 Sg]. Therefore, the final form exiting the last block is /bayles/, conventionally spelled *bailes*.

Block III of Figure 5.2 shows the lack of need for specification of zero-morphs in the model. There is no feature set given for 1 Singular or 3 Singular, since the /X/ variable exiting Block II and entering Block III is itself the 1 Singular and 3 Singular form for all of the above tenses and moods, except for the 1 Singular Present Indicative, which dropped out in the first block as the most specific rule in the block. There was not a more general rule that could have applied in a later block to the 1 Singular and 3 Singular, because there were no later feature sets that included either of these two features. Hence, the form exiting the final block as a sort of default form is itself the 1 Singular and 3 Singular form for the other tenses and moods specified above. This fact is interesting, given the fact that a child learning his or her first language is likely to hear and use the first and third persons singular more than any other. In addition, according to Bybee (1985), the 3 Singular Present Indicative form seems to be the default verb form in

first language acquisition for speakers of Spanish. According to Bybee, it is this form that Spanish-speaking children learn first, and it is this 3 Singular Present Indicative form that children use in place of all other verb forms until they acquire the other forms (50, 59-60). As noted earlier, Spencer (1997) also comments that the 3 Singular form is the default PN specification “in many inflectional systems” (218).

5.5. Categorical/regular forms.

The above rules produce all of the categorical/regular forms, that is, all of the forms for verbs traditionally viewed as completely ‘regular’, for all verbs in the Present Indicative, Present Subjunctive, Imperfect Indicative, and Imperfect Subjunctive. Other verb categories, including phonologically variant forms, exceptional/regular forms, idiosyncratic/suppletive forms, lexical gaps, and ill-formed tokens, will be covered later in this chapter.

An astute Spanish-speaking observer might think at this point that the model has produced an unattested form in the 2 Plural Present Indicative of *i*-class verbs, since according to the rules in the above blocks, the 2 Plural Present Indicative form for the verb *vivir* (‘live’) would be *viviis*, with a double <*i*>. Although this is not the correct spelling of the form, it is in fact the correct pronunciation if we regard each <*i*> as a separate syllable. If so, then the model has not produced an ungrammatical form after all, because the penultimate syllable, meaning the next-to-the last /*i*/ in this case, would be stressed according to normal Spanish stress rules, yielding *viví-is*. (The hyphen is inserted here just to show hiatus, not to indicate spelling.) If this is true, then what happens is that the unstressed second /*i*/ in the affix is deleted in speech, following a

natural language tendency toward lenition or deletion of unstressed vowels in final syllable position (Lathrop 1980: 66), and is thus not represented in the traditional orthographic representation of the form. Therefore, the form must be spelled with a written accent over the <i> in the affix, since with only one <i>, instead of two treated as separate syllables, the normal stress rules are broken.

If we were to create an extra rule which would generate the form with just one /i/ instead of two, then the stress would be wrong. According to the natural stress rules of Spanish, the penultimate syllable is stressed if the final consonant is /s/, contrary to the pronunciation of the actual form *vivís*. An additional rule to create a form with only one /i/ in the affix would yield */βiβis/, with the first syllable (the stem syllable) stressed as per the normal Spanish stress rule, instead of the actual form /βiβis/ with the last syllable stressed. However, if we leave the two /i/'s there in the form generated by the model, as in /βiβiis/, and assume hiatus rather than diphthong (since each /i/ is affixed independently in the model), then, according to normal Spanish stress rules, we do indeed get stress on the penultimate syllable, which is the first affixed /i/. Then if the unstressed second /i/ in the affix is deleted due to lenition, then the model does yield the form orthographically represented as *vivís*, stressed on what is at this point the last syllable. According to conventional orthography, this written form must then be spelled with a written accent over the <i> in the last syllable, since it does not conform to normal Spanish stress patterns.

5.6. Phonologically variant forms with diphthong/mid-vowel alternation.

The verbs traditionally called ‘stem-changing’ verbs have stems that exhibit alternation between a mid vowel such as /e/ or /o/ and a diphthong such as /ye/ or /we/

(<ie> and <ue>, respectively, in conventional orthography). One way to treat this stem alternation would be to specify in the lexicon certain morphosyntactic contexts (the specific TMA and PN features) in which each stem alternant appears. For example, for the verb *pensar* ‘think’ we could list the stem alternants /pyens/ and /pens/ and specify certain feature sets for each as follows, based on morphological criteria:

/pyens/ / 1,2,3 Sg Pres Indic; 3 Pl Pres Indic; 1,2,3 Sg Pres Subj; 3 Pl Pres Subj
/pens/ / elsewhere,

meaning that /pyens/ occurs in the following grammatical contexts: 1, 2, 3 Singular Present Indicative; 3 Plural Present Indicative; 1,2,3 Singular Present Subjunctive; 3 Plural Present Subjunctive /pens/ occurs in all other contexts.

If the respective grammatical feature sets are specified for each of the above stems in the Basic Model, the model would produce the following forms (presented here in standard orthography, but with non-standard stress marks added to show stressed syllables, and with the diphthong/mid-vowel alternation shown in bold type):

Present Indicative: *piénso*, *piénsas*, *piénsa*, *pensámos*, *pensáis*, *piénsan*;

Present Subjunctive: *piénse*, *piénses*, *piénse*, *pensémos*, *penséis*, *piénse*,

The model would also produce all other forms, which use the /pens/ alternate, such as the following:

pensár (‘think’, Infinitive), *pensába*, *pensábas*, *pensábamos*, *pensábais*, *pensában*
(all Imperfect Indicative forms), *pensé*, *pensáste*, *pensó*, *pensámos*, *pensásteis*,
pensáron (all Preterite Indicative forms), among others.

However, although verbs such as these ‘stem-changing’ verbs are considered irregular, they have regular patterns in their irregularity. Although the particular morphosyntactic contexts for each stem alternate could be specified in the lexicon as above (and listed in the list of stem alternates and their respective feature sets in the Appendix of the present analysis), it may be more elegant in some cases to specify a phonological rule that selects the stem alternant on phonological, rather than morphosyntactic, bases. A phonological rule of some sort is especially useful for verbs with only two stem alternates, such as *pensar* or *volver*, which only have the stem alternants *piens/pens* and *vuelv/volv*, respectively. Since some stem alternants seem to follow generalized phonological rules, the choice of stem alternant for some forms could be based on the rule shown in Figure 5.3, which states that a mid-vowel in a verb stem is replaced by a diphthongized stem when the given vowel occurs in stressed position.

FIGURE 5.3. Possible phonological rule for diphthong/mid-vowel alternation in Spanish verbs.

Root vowel [+ Mid vowel, – Diph] → [+ Diph] / syllable [+ stress].

(That is, a monophthong mid vowel in the root is realized as a diphthong in a stressed syllable.)

For example, consider the following forms (with the diphthong/mid-vowel alternation shown in bold type, as above, and the stressed syllable marked with an accent mark for illustrative purposes, although not necessarily the conventional spelling):

pensár ('think', Infinitive) *piénso*, *piénsas*, *piénsa*, *pensámos*, *pensáis*, *piénsan*
 (all Present Indicative forms), *pensába*, *pensábamos* (Imperfect Indicative forms),
pensé, *pensó*, *pensáste* (Preterite Indicative forms), or
volvér ('return', Infinitive), *vuélvo*, *vuéives*, *vuélve*, *volvemos*, *volvéis*, *vuélven*,
 (all Present Indicative forms), *volvía*, *volvíamos* (Imperfect Indicative forms),
volví, *volvió*, *volviste* (Preterite Indicative forms).

The diphthongized variant for the above mentioned verbs occurs only in the syllables that are stressed. The normal stress patterns for all Spanish words are governed by a few simple stress rules applied throughout the language. The normal stress pattern in Spanish is the following:

- 1) The penultimate syllable is stressed if the word ends in a vowel or the consonant /n/ or /s/.
- 2) The last syllable is stressed if the word ends in any consonant other than /n/ or /s/.
- 3) Words that do not conform to the above normal stress pattern bear a written accent in the orthographic representation of the word.

The way the rule in Figure 5.3 is stated is the way this diphthong/mid-vowel alternation is traditionally viewed, which is that the a monophthongal root vowel is realized as a diphthong in an unstressed syllable for such 'stem-changing' verbs. However, without a way to determine which verbs are 'stem-changing' verbs, such a rule might produce incorrect forms for a verb like *aprender* 'learn', such as **apriendo*, **apriendes*, **apriende*, rather than *aprendo*, *aprendes*, *aprende*, etc, or **cuemo*, **cuemes*,

**cueme* for *come, comes, come* ‘eat’, and other similarly incorrect forms for other verbs with a monophthong mid vowel in the stem.

However, if we postulate the reverse of the rule in Figure 5.3, with the diphthong variant regarded as the base lexeme for the so-called ‘stem-changing’ verbs, then we would not have to account for the verbs that do *not* have the diphthongized variant. For such a rule, a phonological process would change the diphthong to a single mid vowel, instead of the other way around. In this view, a nonce verb such as **sostar* could be assumed to maintain its monophthong mid vowel throughout its inflected paradigm, while a nonce form such as **suecha* could be assumed to have a monophthong mid-vowel stem alternate in unstressed contexts. This alternate view with the diphthong as the base variant is particularly appealing in light of Bybee’s (1985: 50, 59-60) and Spencer’s (1997: 218) observations about the 3 Singular form being the default form in many languages, and since Spanish verbs with diphthong/mid-vowel alternation have the diphthong variant in the 3 Singular Present Indicative form.

If the phonological rule in Figure 5.3 is stated in reverse (that is, [+ Diph → – Diph]) instead of [– Diph → + Diph], as is typically presented in teaching texts, the rule could be represented as shown in Figure 5.4.

FIGURE 5.4. Alternate phonological rule for diphthong/mid-vowel alternation in Spanish verbs.

Root vowel [+ Diph] → [– Diph] / syllable [– stress].

(That is, a diphthong in the root is realized as a monophthong mid vowel in a stressed syllable.)

If a general phonological rule is going to be assumed, the ‘backwards’ version shown in Figure 5.4 is more economical, since not all verbs with a mid-vowel in the infinitive form have this diphthong variant. Bybee points out that “there are many mid vowels that do not alternate, e.g., *flotar* ‘to float’, and *aprender* ‘to learn’, which have mid vowels throughout their paradigms,” adding that “non-alternating diphthongs in verb stems are very rare,” occurring “only in a few verbs formed from nouns or adjectives with stressed diphthongs, e.g., *muébles* ‘furniture’, *amueblár* ‘to furnish; *viéjo* ‘old’, *aviejár* ‘to grow old’” (Bybee and Pardo 1981: 939). As Bybee points out, “the diphthong could be taken as underlying, since there are very few verbs with a nonalternating diphthong” (1988: 123). In this view, the base lexeme for *pensar*, for instance, would be /pyens/, since the rule states that the diphthong /ye/ would become the front mid-vowel /e/ whenever the syllable is unstressed, as in *pensámos* above, with the second syllable stressed. But the base lexeme for *aprender* would remain /aprend/ whether stressed or not, according to the rule in Figure 5.4. In other words, a stem (or stem alternant) with a mid vowel cannot be assumed to have a diphthongized alternate, but a stem alternant with a diphthong in most cases can be assumed to have a monophthong mid-vowel alternate. *Aprender* and *comer*, for example, maintain the monophthong mid vowel throughout their conjugations, as in *aprendo*, *aprendes*, *aprende*, *aprendemos*, *aprendéis*, *aprenden*.

However, in experiments with native Spanish speakers, Bybee and Pardo (1981: 124) tried unsuccessfully to get Spanish speakers to apply a generalized phonological rule such as that in Figure 5.3 or that in Figure 5.4 to nonce verbs. The experiment failed to consistently elicit diphthongized forms such as **suecha* from a nonce mid-vowel form such as **sochar*, as would be produced by the rule in Figure 5.3. The experiment was

equally unsuccessful in eliciting monophthong mid-vowel forms such as **sochar* from diphthongized forms such as **suecha*. This observation, if correct, would point more toward lexical specification of diphthong/mid-vowel stem alternation rather than a generalized phonological rule.

Ruiz' (1986) study of vocalic alternation in the speech of Spanish-English bilinguals also suggests lexicalization of the various alternants. Among such speakers, Ruiz' study found wide variation in the use of diphthongal stem variants vs. those with mid vowels (as in *ie ~ e* or *ue ~ o*), and similarly with high vowels vs. mid vowels (as in *i ~ e* or *u ~ o*). It is interesting to note that he found such variation in many nouns as well as verbs. The verbal variants were frequently regularized so that all forms of a paradigm exhibited the same alternant, such as diphthongizing the whole paradigm: *cuento, cuentas, cuenta, *cuentamos, *cuentáis, cuentan*, for Standard Spanish *cuento, cuentas, cuenta, contamos, contáis, cuentan*.

The patterns of these alternations were not consistent among the speakers in Ruiz' study, however. He cites a number of sociolinguistic factors that may account for such apparent instability or lack of systematicity in the language of such bilingual speakers. These sociolinguistic factors are not particularly relevant to the present analysis, but it is relevant to note that variation in the system, whether individual or collective, may be the norm rather than the exception. It is also relevant to note that, as Ruiz comments, "we can describe linguistic behavior within a system without providing insights into the actual processes underlying such behavior" (115).

The question of why certain forms became diphthongized in the diachronic development of the language is not particularly relevant to the present synchronic

analysis of Spanish verbs, just as it is not particularly relevant why there is such dialectal variation. However, the present analysis does need to account for this diphthong/mid-vowel alternation in the system as a whole and for any given speaker's knowledge of the fact that the alternation does not occur in all verbs with mid vowels in the stem, and it also must account for forms such as *tengo* that do not have a diphthong in the stressed syllable.

It could be argued that neither the mid vowel nor the diphthong variant is the base lexeme, but rather that the base lexeme is an abstract concept, just as a phoneme is an abstract concept and is arbitrarily represented by one of its allophones. For example, whether we represent the abstract phoneme in the middle of the word *haber* as /b/ or /β/ is essentially arbitrary, since we are using a concrete symbol of a physical sound to represent an abstract, non-concrete concept. Although it may not be necessary to specify a concrete base lexeme if we see the lexeme as an abstract concept, we may need to have a way of representing the base lexeme for notational purposes, just as we need to have a way to represent phonemes for notational purposes.

Although the infinitive form of the verb is the form listed in printed dictionaries and is the form typically presented as the base form in pedagogical contexts, there is no apparent psychological reason to assume that the infinitive is the base lexeme stored in the lexicon. In fact, Spencer (1997) comments that “In many inflectional systems, 3rd person singular is the default PN specification” (218). If this is the case, then the stem of a 3 Singular form like *piensa* or *muere* may be the base form, thus supporting some form of the rule in Figure 5.4, with the diphthong variant becoming the monophthong mid-vowel variant in stressed position. Furthermore, the infinitive in Spanish, like any other

form of a given verb, consists of a stem that is phonologically altered in some way, typically by adding an affix at the end of the stem. The monophthong variant occurs in the infinitive form, where the stem-containing syllable is not stressed, as in /pensár/, but this fact does not preclude the possibility that the base lexeme may be the diphthongal variant rather than the monophthongal one, or the possibility that all stems have equal status in the lexicon.

In Spencer's (1997) discussion of Anderson's Extended Word-and-Paradigm theory, he comments that "Anderson's most popular solution [to the problem of generating incorrect forms] seems to be to assume a phonological rule whenever feasible" (219). However, neither the rule in Figure 5.3 nor that in Figure 5.4 can be a generalized phonological rule, at least not without the stipulation that this change occurs only for lexically specified verbs. The Figure 5.3 rule cannot be generalized because does not apply to all verbs with a mid vowel in the stem, but neither can the Figure 5.4 rule be a generalized phonological rule, since it could produce forms such as **tieno*. The rule cannot simply state that 'a monophthong root vowel is realized as a diphthongal vowel in a stressed syllable', nor can it simply state that 'a diphthongal root vowel is realized as a monophthong vowel in an unstressed syllable'. Rather, the rule must state that 'a diphthongal root vowel is realized as a monophthong vowel in an unstressed syllable for lexically specified verbs', and, in addition, in some cases there must be other, more specific rules (or representations) that block application of the diphthongized stem, such as in *tengo*.

Since not all verbs with a mid-vowel in the stem exhibit this diphthong/mid-vowel alternation, a verb that does exhibit this alternation must be specified as such in the

feature sets for each of the alternants in its lexical stem set. In other words, this diphthong/mid-vowel alternation must be ‘lexically specified’. Bybee (1988) comments that “This diphthong/mid vowel alternation resulted from the diphthongization of lax mid vowels in stressed syllables, but this process is no longer phonetically conditioned in Spanish” (123). Bybee maintains that such alternations in Spanish, like English strong verb past tense forms, can only be explained as a process of “accessing the lexicon rather than by applying a rule to a base form to change the vowel” (125).

For instance, for the verb *pensar*, the stem alternants would be listed as follows (showing phonetic notation here rather than conventional spelling):

/pyens/ / syllable [+ stress]

/pens/ / syllable [– stress].

This lexical stem set, when fed into the proposed model, yields all the correct forms for the verb *pensar* (*think*), such as *pensár* itself (the Infinitive form), *pensába*, *pensábas*, *pensábamos*, *pensábais*, *pensában* (all Imperfect Indicative forms), *pensé*, *pensáste*, *pensó*, *pensámos*, *pensásteis*, *pensáron* (all Preterite Indicative forms), among others.

For a verb like *pensar*, which has only two alternants in its lexical stem set, the above listing of feature sets is not very complicated. A verb such as *tener*, however, is a more complex, since in the case of *tener* there are five stem alternates: /ten/, /tyen/, /teng/, /tuβ/, /tendr/. Either of the above rules, that in Figure 5.3 or the reverse version in Figure 5.4, might produce incorrect forms if the rule is applied as a generalized phonological rule rather than a phonological rule that is lexically conditioned, with more specific rules that precede and block diphthongization in forms that have other stem alternants. The rule in Figure 6.3 would diphthongize the front mid-vowel /e/ in *tengo*, thus generating

the incorrect form **tiengo*, unless there is stipulated a more specific rule that produces *tengo*, as in the proposed model. Likewise, a general phonological rule stipulating [– Diph → + Diph] would also have to apply to *tengo*, since the syllable in question is in stressed position and contains a monophthong mid vowel. Just as with the Figure 5.3 rule, the rule in Figure 5.4 would produce an incorrect form such as **tieno* if there is no further lexical specification, since this latter assumes the diphthongized stem to be the base variant.

In the proposed model, the lexical stem set for *tener* specifies */teng/* for [{+ Pres Subj; 1 Sg Pres Indic}], meaning that the stem */teng/* is specified for forms with the semantic/grammatical feature of Present Subjunctive or 1 Singular Present Indicative. The stipulation of */teng/* for [{+ Pres Subj; 1 Sg Pres Indic}] is more specific than the specification of */tyen/* for a form in which the stem is stressed, noted as */tyen/* / syllable [+ stress]. Thus, *tengo* can be regarded as a representation that is more specific than the more general rule/representation */X/* → */X + o/* for the 1 Singular Present Indicative found in Block I of the Basic Model, even though the latter is a more specific rule than the other rules for *e*-class verbs found in Block I. Therefore, the specification of */teng/* for [{+ Subj; 1 Sg Pres Indic}] blocks the more general rules that would otherwise produce incorrect forms.

Lexical specification for stem selection, such as:

/pyens/ / syllable [+ stress]; */pens/* / syllable [– stress], or

/muer/ / syllable [+ stress]; */mor/* / syllable [– stress],

makes the general phonological rules in Figure 5.3 or 5.4 unnecessary. For any form of the verb *pensar*, for example, the above stems feed into the first block of the model as the initial X variable, and the correct stem variant is then selected on a phonological basis, depending upon the phonological environment, that is, whether the stem is in stressed or unstressed position. A verb with a non-alternating mid vowel, such as *comer*, *flotar*, or *aprender*, would not have such a diphthongal variant in its lexical stem set, and, therefore, would not exhibit this diphthong/mid-vowel alternation. Thus, the stem alternate is selected on morphophonological bases, rather than phonological bases alone or morphological bases alone.

In keeping with the present model's whole-word, morpholexical view of the lexicon, the monophthongal variant is stored in the lexicon along with the diphthongal variant, with each variant specified (in its associated feature set, as in the example of /*pyens*/ and /*pens*/ above) as occurring when the stem syllable is stressed or unstressed, as the case may be. The feature sets include this stipulation along with the other semantic and grammatical features associated with the lexeme, just as subcategorization requirements may be included in the feature set of a given verb. The form-meaning association would only be made for a given form when the semantic and grammatical specifications are met, and in this case, when the phonological specifications are met. Since a verb such as *aprender* has only one stem in its lexical stem set, /*aprend*/ in this case, there is no possibility of producing an incorrect diphthongized form such as **apriendo*.

In the proposed verb model, the regular Spanish stress patterns described above apply to all forms unless noted otherwise by a written accent on the vowel in question.

Thus, a phonological string with an abnormally stressed vowel requires a separate listing in the model, in contrast with the normally unstressed version of the same sequence, and is indicated with a written accent mark.

The 2 Plural form *pensáis* may appear to pose a problem, as did *vivís* above, since it appears to break a pronunciation rule and thus require a written accent. However, even though typical pronunciation treats the /áy/ combination as a diphthong, it may be that psychologically this string is seen as two vowels in hiatus, as in the two /i/'s apparently produced by the model for the 2 Plural *i*-class verbs discussed earlier. Each vowel, both the /a/ and the /i/ in *pensáis*, could be regarded as having a distinct morphological identity, as in the case of the two /i/'s in *viví-is*. The /a/ of a form like *pensáis* is affixed as a theme vowel, while the /i/ is affixed in conjunction with the feature set [+ 2 Plural]. For this reason, then, we could postulate a sort of 'psychological hiatus' in spite of the diphthongized pronunciation of the two vowels together. Even if we view morphology as *a-morphous* in Anderson's terms, the above model shows that there are distinct feature sets associated with each of the vowels, with each of the vowels affixed by different rules in different blocks, with different feature sets. This separation of two vowels may account for the psychological hiatus in spite of the phonetic diphthongization of the two vowels, thus, as with *vivís*, making the stressed /a/ not really irregular after all.

5.7. Phonologically variant forms with mid-vowel raising.

In addition to verbs with lexically conditioned diphthong/mid vowel alternation based on syllable stress, as discussed the previous section, there are other verbs in Spanish with stems that alternate between a monophthong mid vowel and a

monophthongal high vowel, such as the front mid vowel /e/ and the front high vowel /i/ or the back mid vowel /o/ and the back high vowel /u/. This process is referred to as ‘mid-vowel raising’, since the mid vowel is ‘raised’ to a high vowel. This group includes verbs such as *pedir*, *seguir*, *servir*, *sentir*, *medir*, *vestir*, *reír*, *freír*, *morir*, all of which happen to be *i*-class verbs. For this group of verbs, it might seem at first glance that the general phonological rule could be stated as shown in Figure 5.5, modeled after the rule in Figure 5.4 above. (As before, the alternating vowels are indicated by bold type, and the stressed syllables are indicated by a written accent on the stressed vowel, although the stress mark may not be used in the standard orthographic representation.)

FIGURE 5.5. Hypothetical (but insufficient) phonological rule for mid-vowel/high-vowel alternation in Spanish verbs.

Root vowel [+ High vowel] → [+ Mid vowel] / syllable [– stress],

as in *pedír*, *pído*, *pídes*, *píde*, *pedímos*, *pedís*, *píden*.

(That is, a high root vowel is realized as a mid vowel in unstressed syllable position for lexically specified verbs.)

This rule is not sufficient to produce all the correct forms of the specified verbs, however, since this group of verbs, in addition to the stem alternation based on syllable stress as above, also exhibits the same vowel-raising in certain other forms: the 3 Singular and 3 Plural Preterite Indicative, all forms of the Imperfect Subjunctive, and the Present Participle. For example, the verb *pedir*, in addition to the forms given above in Figure 5.5, has the following forms in the Preterite Indicative:

<i>pedí</i>	<i>pedimos</i>
<i>pediste</i>	<i>pedisteis</i>
<i>pidió</i>	<i>pidieron</i>

In addition, the Present Participle is *pidiendo*, and the forms of the Imperfect Subjunctive are the following, all of which have the high vowel /i/ in unstressed position:

<i>pidiera</i>	<i>pidiéramos</i>
<i>pidieras</i>	<i>pidiérais</i>
<i>pidiera</i>	<i>pidieran</i>

All of the *i*-class verbs with mid-vowel/high-vowel alternation also have mid-vowel/high-vowel alternation in the 1 Plural and 2 Plural Present Subjunctive and the Present Participle, as well as in the 3 Singular and 3 Plural Preterite Indicative and in all forms of the Imperfect Subjunctive. Furthermore, there is another group, all *i*-class verbs, that exhibit both monophthong/diphthong alternation and mid-vowel/high-vowel alternation, such as *divertir*, *advertir*, *dormir*, *morir*. For this latter group, in addition to all of the mid-vowel/high-vowel alternations exhibited by the other *i*-class verbs discussed above, such as *pedir*, there is monophthong/diphthong alternation in stressed syllable position in the Present Indicative and Present Subjunctive as discussed in Section 5.6, except for the 1 Plural and 2 Plural Present Subjunctive. These latter two Present Subjunctive forms have mid-vowel/high-vowel alternation instead of monophthong/diphthong alternation. Thus, some verbs, like those listed above with *pedir*, have only mid-vowel/high-vowel alternation, but others, such as *divertir*, have a combination of diphthong/mid vowel alternation and mid-vowel/high-vowel alternation.

For example, the verb *divertir* has the following forms in the Present Indicative, exhibiting monophthong/diphthong alternation:

<i>diviérto</i>	<i>divertímos</i>
<i>diviértas</i>	<i>divertís</i>
<i>diviérte</i>	<i>diviérte</i>

In the Present Subjunctive, however, it has a combination of diphthongization and mid-vowel raising:

<i>diviérta</i>	<i>divirtámos</i>
<i>diviértas</i>	<i>divirtáis</i>
<i>diviérta</i>	<i>diviértan</i>

The Preterite Indicative maintains the mid vowel /e/ in all forms except the 3 Sg and 3 Pl, with mid-vowel raising in these latter two forms:

<i>divertí</i>	<i>divertímos</i>
<i>divertiste</i>	<i>divertísteis</i>
<i>divirtió</i>	<i>divirtiéron</i>

There is also mid-vowel raising in the Present Participle, *divirtiéndo*, but not in the Past Participle, *divertido*.

Lathrop (1980) gives a historical reason for the mid-vowel raising in the Preterite Indicative and the Imperfect Subjunctive. Lathrop contends that “The reason for this vocalic change lies in the effect exerted by a Vulgar Latin yod [the high glide or semi-vowel /y/] that existed in certain forms, and only in the *-ire* conjugation” (73).

According to this theory, the *yod* in the diphthong of the following syllable (the syllable with the affix) exerted a raising effect on the vowel in the stem, as in the Vulgar Latin paradigm shown in Figure 5.6.

FIGURE 5.6. Vowel-raising caused by *yod* in the evolution of Spanish from Vulgar Latin (Lathrop 1980: 74). (Grammatical labels for the Spanish forms have been added by the present author.)

VULGAR LATIN		SPANISH
<i>servierunt</i>	→	<i>sírvieron</i> [3 Pl Pret Indic]
<i>servierant</i>	→	<i>sírvieran</i> [3 Pl Imperf Subj]
<i>serviamus</i>	→	<i>sírvamos</i> [1 Pl Pres Subj]
<i>serviendu</i>	→	<i>sirviendo</i> [Pres Part]
<i>dormierunt</i>	→	<i>durmieron</i> [3 Pl Pret Indic]
<i>dormierant</i>	→	<i>durmieran</i> [3 Pl Imperf Subj]
<i>dormiamus</i>	→	<i>durmamos</i> [1 Pl Pres Subj]
<i>dormiendu</i>	→	<i>durmiendo</i> [Pres Part]

The table in Figure 5.6 includes the 1 Plural Present Subjunctive form, which in modern Spanish has a raised vowel in the stem in spite of having no *yod* in the following syllable. Lathrop (1980) explains this phenomenon as follows:

It should be noted that the *yod* was finally lost in the first person singular present subjunctive forms shown above (*serviamus* → *sírvamos*, *dormiamus* → *durmamos*). This was due to analogical pressure from the many verbs that had no *yod* in that form (CL [Classical Latin] *laudemus*, *ponamus*) (74).

Lathrop goes on to say that “The *yod* was not lost, however, until it had had its full phonetic effect on the preceding vowel” (74).

The 2 Plural Present Subjunctive form (such as *muráis*) also exhibits the same vowel raising, but Lathrop does not discuss this latter form. It is not clear whether this *yod* influence can explain mid-vowel raising in a synchronic analysis of the language, that is, whether the *yod* in the syllable following the stem might continue to condition the preceding vowel by raising it in the mind and speech of individual speakers of modern Spanish. Since *yod* is high on the vowel triangle, it is conceivable that it continues to exert a raising influence on the vowel in the preceding syllable. But even if this is true, a synchronic analysis cannot explain the vowel-raising seen in the 1 Pl Present Subjunctive forms for this class of verbs, such as *durmamos* or *muramos*. Lathrop says that these mid and high vowels were close to each other in Vulgar Latin (73). If this is still partly true for modern Spanish, then it could be that speakers are not clearly aware whether they are pronouncing an /u/ or an /o/. It would be interesting in the future to conduct an experiment to determine empirically whether this might be true. If so, then the mid-vowel raising may be more of an orthographic convention than a phonetic anomaly.

If, like Anderson, we assume that “the most popular solution [to the problem of generating incorrect forms] seems to be to assume a phonological rule whenever feasible” (Spencer 1997: 219), then it might be beneficial to empirically evaluate the possibility of a synchronic phonological rule that could fully explain these anomalies. However, the present analysis lacks concrete evidence for either of the above possibilities at the present time (the continuing influence of *yod* in forms where *yod* is no longer present, or the phonetic closeness of the mid and high vowels). For this reason, and because the present analysis takes a primarily lexical view of inflectional morphology, the present study proposes specifying the above discussed stem alternants in the lexicon, along with their

corresponding grammatical feature sets when such a specific set of features is needed to take precedence over the more general morphophonological rules, as in the following:

/mur/ / [{1Pl Pres Subj; 2Pl Pres Subj; 3Sg Pret Indic; 3Pl Pret Indic; Imperf Subj}]

/muer/ / syllable [+ stress]

/mor/ / syllable [– stress]

Since the first rule is more specific than the latter two, it applies first (when the feature sets match those of the given form), and thus blocks application of the other two more general rules. The latter two rules would apply in all contexts other than those lexically specified for */mur/*, which is the form with mid-vowel raising.

One possibility for this mid-vowel/high-vowel alternation is that it is due to a dissimilatory phonological process, whereby neighboring sounds become less similar in articulation. The mid vowel alternant */e/* occurs in the stem when the following syllable has vocalic */i/*, but when the following syllable does not contain this vocalic */i/*, the high vowel alternant */i/* occurs in the stem. For example:

<i>pedir, seguir</i>	<i>pido, sigo</i>
<i>pedía, seguía</i>	<i>pides, sigues</i>
<i>pedimos, seguimos</i>	<i>pidan, sigan</i>
<i>pedís, seguís</i>	<i>pidiendo, siguiendo</i>
<i>pedido, seguido</i>	<i>pidiera, siguiera</i>

In the case of forms such as *pidiendo, siguiendo, pidiera, siguiera*, although there is a graphemic <*i*> in the syllable following the stem syllable, this graphemic <*i*> is realized as */y/*, that is as a *yod*. Thus, it is a glide rather than a vocalic */i/*.

Another apparent ‘irregularity’ is seen in the Preterite forms of verbs like *leer* and *creer*, for which the grapheme <*i*> changes to <*y*> when it occurs between two vowels, as in *creyó, leyeron, creyeras, leyeran*, etc., instead of the expected **creió, *leieron*,

**creieras, *leyeran*. This vowel alternation is not really an alternation in pronunciation, however. It is merely an orthographic convention, since the forms would sound the same whether spelled with a <y> or an <i>.

In summary, the irregularities in the group designated as the phonologically variant group in the present analysis are treated as alternation of the stem, with the choice of alternant based primarily on phonological rules that are lexically conditioned, rather than as independent rules operating only in the phonological component. This means that a phonological rule such as that stated in Figure 6.4 only applies if such a diphthong/mid-vowel alternation is specified in the lexicon for a given verb. Thus, in the present analysis such phonological ‘irregularities’ in the stem are accounted for by the semantic/grammatical feature sets that accompany the stem alternates for a given verb, such as those listed in the Appendix. As we have seen earlier, Anderson (1995) suggests a “lexical stem set” (133) with various forms of a given lexical item already specified for certain morphosyntactic features (or in this case, morphophonological features) in the lexicon. As in Anderson’s model, when choosing the particular stem to undergo Word-Formation Rules, the one that matches the most features of the morphosyntactic (or morphophonological) representation is the one selected from the lexical stem set. The model proposed in the present analysis deals with stem allomorphy in general by employing Anderson’s notion of lexical stem sets and by selecting stem alternants based on matching features, as does Anderson’s model. And, consistent with Bybee’s view of morphology as lexical organization, the proposed model sees these stem alternations as “highly lexically restricted, that is, dependent upon the particular items to which they apply, and not extendable to new items” (1988: 124).

If the rule-and-feature sets are built up in the lexicon through analogical connections and collocations as in Bybee's analogical model of morphology (1985, 1988), and/or by establishing sets of 'cohorts' as in Marslen-Wilson's theory (1987, 1989, 1999), then it is likely that the complexity and 'accuracy' of the rule-and-feature sets varies from one person to another. This variance would depend to some degree upon the extent of the individual's exposure to the language, since a great many forms would have to be heard and processed in order to build up and strengthen the necessary connections or cohort sets.

5.8. Stem alternation and exceptional/regular forms.

As mentioned above, a verb such as *tener* is more complex than the above example of *pensar*, since in the case of *tener* there are five stem alternates: /ten/, /teng/, /tyen/, /tuβ/, /tendr/. If the above phonological rule is applied to the verb *tener*, there is the problem of knowing where the other stem alternates occur, that is, the stems other than /tyen/ and /ten/. This apparent problem of choosing from among the stem alternates is handled by the specification of particular contexts (meaning the specific TMA and PN features) for each stem alternate in their respective feature sets in the lexicon, such as those listed in the Appendix.

The stem /teng/, for instance, is lexically specified for [{+Pres Subj; +1Sg Pres Indic}]. This very specific rule, which selects the stem alternate that enters the first block of the model, feeds into the less specific rule that yields the 1 Singular Present Indicative form, and into the more general rule that yields the Present Subjunctive forms, yielding *tengo* (1 Singular Present Indicative) and the Present Subjunctive forms *tenga*, *tengas*,

tenga, tengamos, tengáis, tengan. The other stem alternates for *tener* are specified in the lexicon for other morphological contexts, such as /tuβ/, which is the stem specified for [+ Pret Indic; + Imperf Subj]. Thus the stem /tuβ/ feeds into the Basic Model as the X variable for Imperfect Subjunctive forms, and into the Preterite model (presented in Chapter 6 of the present study) for all forms of the Preterite Indicative. Such a stem alternate, although ‘exceptional’ and traditionally considered ‘irregular’, enters the first block of the model and can then proceed with a somewhat regular pattern through the rest of the blocks. For example, there is an entire group of verbs that function like *tener* in the Preterite; thus they are ‘exceptional/regular’ verbs. This group is discussed further in the section on Preterite forms later in this chapter.

The stem alternant list, with the respective feature sets associated with each stem, precedes the first block in the model, thus choosing the appropriate stem for a given verb form. The appropriate stem (the one that matches the feature sets for the given verb form) enters the first block as the X variable and then adds the appropriate phonological material as specified by all the matching feature sets, as in the case of /teng/ entering the model and yielding the correct forms for the 1 Singular Present Indicative form and all of the Present Subjunctive forms. The further usefulness of this treatment of stem alternation (‘allomorphy’ in traditional terms) will be seen later in the discussion of forms other than the ones produced by the Basic Model in Figure 5.2, such as the Preterite Indicative, Future, and Conditional forms.

5.9. Rules of referral.

For a few verbs, the allomorphy in the root could possibly be dealt with by a rule of referral, as in Zwicky (1985). ‘Rules of referral’ in general are rules that allow for a sort of metalinguistic awareness of syncretism or homonymy among forms in a paradigm, such as in many 1 and 3 Singular verb forms in Spanish. For example, a rule of referral effectively allows us to say that the 1 Singular and 3 Singular forms for the Imperfect Indicative are identical (as in *hablaba, bebía, escribía*) as are the 1 Singular and 3 Singular forms of the Imperfect Subjunctive as well (as in *hablara, bebiera, escribiera*). Similarly, a rule of referral would allow us to say that for most verbs the 3 Singular Present Indicative (such as *habla, bebe, escribe*) is identical to the Singular Familiar Imperative (as in *habla, bebe, escribe*). Spencer (1997) comments that “Any theory of inflection (or of morphology as a whole) which failed to make allowance for such phenomena would be sadly deficient” (194).

Rules of referral can also be cited for forms that have homonymous stems, such as Spanish Present Subjunctive, for which the stem is identical to that of the 1 Singular Present Indicative form. Spencer points out the following: “This Spanish case [of Present Subjunctive] can be handled by a rule of referral, provided it can refer to stem allomorphs marked with features such as [1sg. Pres. Indic.]” (194). Rules of referral in this case allow for one form to be based on another, that is, to ‘refer’ to another form. If we allow such a rule, referring to a specific whole form to form a new stem, we could formulate a rule such as that shown in Figure 5.7.

FIGURE 5.7. Rule of referral for Present Subjunctive forms.

$$\text{If } /X/ = \begin{bmatrix} + 1 \text{ Sg} \\ + \text{ Pres} \\ + \text{ Indic} \end{bmatrix} - /o/ \text{ for the feature set that includes } \begin{bmatrix} + \text{ Pres} \\ + \text{ Subj} \end{bmatrix}$$

then /X/ feeds into Rule 3 of Block I of the Basic Model.

(That is, for the feature set that includes Present Subjunctive, the variable /X/ is the 1 Sg Pres Indic form minus the final /o/, and it is this stem that enters Block I of the Basic Model.)

In the case of *poner* ‘put’, for example, if we accept as a stem /X/ the 1 Singular Present Indicative form *pongo* minus the final /o/, then this alternant stem variant /pong/ would feed into Rule 3 of Block I above, yielding *pongo*, *pongas*, *ponga*, *pongamos*, *pongáis*, *pongan*. This rule of referral would generate the small group of present subjunctive forms whose stem is the same stem allomorph as that of the 1 Singular Present Indicative form, such as *poner*, *hacer*, *tener*, *conducir*, and others.

A similar rule of referral can be formulated to produce some of the stem variants for the Imperfect Subjunctive, since the Imperfect Subjunctive forms have the same stem variant as the Preterite Indicative. The rule of referral for Imperfect Subjunctive could be represented as a rule such as that in Figure 5.8. Although such rules of referral may seem unnecessarily complex, it should be noted that in the case of the Imperfect Subjunctive forms, such a rule would allow bypassing the first two blocks entirely and proceeding directly to Block III.

FIGURE 5.8. Rule of referral for Imperfect Subjunctive forms.

$$\text{If } /X/ = \left[\begin{array}{l} + 3 \text{ Pl} \\ + \text{ Pret} \\ + \text{ Indic} \end{array} \right] - /n/ \text{ for the feature set that includes } \left[\begin{array}{l} + \text{ Imperf} \\ + \text{ Subj} \end{array} \right]$$

then /X/ feeds directly into Block III of Figure 6.2 above.

(That is, for the feature set that includes Imperf Subj,

the variable /X/ is the 3 Pl Preterite Indicative form minus the final /n/, and it is this stem that enters Block III of the Basic Model.)

One problem with this latter Rule of Referral (Figure 5.8) is that it assumes the pre-existence of a separate Preterite model. This could be the case, since Preterite forms are somewhat erratic and idiosyncratic and could, therefore, be represented as more specific and thus be applied early in a disjunctively ordered block of feature sets. However, since the Preterite forms are not entirely suppletive and do have some regular patterns, including the 1 Pl /*mos*/ and the 3 Pl /*n*/, it seems redundant and inefficient not to take these regularities into account. Therefore, further presentation of the proposed model will deal with the fact that the Preterite and the Imperfect Subjunctive have the same stem alternants.

It should be noted that only the /*ra*/ subjunctive forms, and not the forms with /*se*/, are included in the present verb model. However, in varieties of Spanish in which the /*se*/ forms are common, the same model would apply, but with /*se*/ in place of /*ra*/ for the feature sets that include the notation of Imperfect Subjunctive. In varieties of Spanish in which both forms are used but occur in distinct contexts, the particular alternant chosen would have to be specified in the semantic feature set.

5.10. An alternative to the first rule of referral.

It may be possible to treat some cases of allomorphy with a phonological rule. The grapheme *c* in *hacer* is realized as /θ/ in Castilian Spanish (although it is neutralized as /s/ in other varieties of Spanish), and never appears before the low back vowel /a/. This phonological constraint would rule out the normal application of Rules 2 and 3 of Block I in the Basic Model in Figure 5.2 above. Without the root allomorphy this block would yield **haca*, **hacas*, etc., but pronounced **/aθa/*, **/aθas/*, etc. with the disallowed /θ/ or /s/ in the middle, not /k/. If, however, this consonant is the /k/ version of the grapheme *c*, according to the phonological rule above (disallowing /θ/ before a low back vowel), then it is not far from the actual form *haga*, *hagas*, etc. So it may be that the ‘irregular’ form is really regular after all, in a sense, with a phonological rule changing the /θ/ (or /s/) to /k/, then intervocalic voicing changing it to /g/.

From a diachronic viewpoint, however, the reverse may actually be the case, and the form with /g/ could be the base lexeme (if we wish to assume one stem as more ‘basic’ than another). Since Spanish *hacer* /aθer/ or /aser/, is from the Latin *facere*, with the graphemic *c* pronounced as /k/ before *e* or *i*, unlike modern Spanish and English, the variant with /k/ (pronounced */fakere/*) instead of /s/ or /θ/ could be considered the base lexeme, at least from a diachronic viewpoint. According to Lathrop (1980),

The CL [Classical Latin] *c-* before *e* or *i* was pronounced *k* in Classical Latin, but became [ts] in Vulgar Latin. This latter sound remained throughout Old Spanish, but later simplified to a spirant based on [t], giving the modern [θ], and in the south of Spain, the [ts] simplified to the second part of the cluster [s]” (p. 114).

this same consonantal phoneme would be realized as [s] or [θ] before a front vowel.

If the /g/ form is the base variant but is actually an intervocalic voiced /k/, then perhaps the /s/ alternant naturally occurs before a front vowel, such as in *hacer*, realized as /aser/ in Latin American pronunciation. The alternation between /k/ and /s/ is seen in cognates among many languages in the Indo-European family, and is in fact the basis for the classification of Indo-European languages according to the so-called ‘*centum-satem* split’, based on alternate developments of the Indo-European velar /k/ and palatal /k/ phonemes. The two words *centum* and *satem* are the Latin and Avestan words, respectively, for ‘one hundred’, with the velar stop /k/ as the initial consonant in the former and the alveolar fricative /s/ as the initial consonant in the latter (Pyles and Algeo 1993: 66-67).

Nevertheless, although such a historical analysis is interesting and relevant to the diachronic development of the stem alternants, the present model selects stem alternants based on lexical specifications in the feature sets for each of the phonologically variant stems in the lexical stem set for a given verb. For the verb *hacer*, for example, the lexical stem set includes the following:

hag (/ag/) / [Pres Subj; 1 Sg Pres Indic]

hic/hiz (/is/ or /iθ/) / [Pret Indic; Imperf Subj]

/har/ (/ar/) / [Fut; Cond]

hac (/as/ or /aθ/) / elsewhere

5.11. Idiosyncratic/suppletive and exceptional/regular forms.

If allomorphy in the stem is accounted for, the above blocks of feature sets and their associated rules describe virtually all of the forms for all verbs in the following tenses and moods: Present Indicative, Present Subjunctive, Imperfect Indicative, and Imperfect Subjunctive, except for a few idiosyncratic/suppletive ones. Even some of the suppletive ones are consistent with Block III if one of the suppletive stems (one of the /X/ variables) is fed into this block. The suppletive stems, and a few entire forms, are lexical variants of a given verb and are treated as members of the verb's lexical stem set. These variants, whether whole forms themselves or just parts of whole forms, feed into the blocks under certain specified conditions, with certain feature sets. Considering that the rules in the blocks of the proposed model apply from most specific to least specific, the idiosyncratic/suppletive forms would be listed first in a separate block and would, of course, apply before the later more general rules.

Let's look at an example for *ser*, one of the 'be' verbs in Spanish:

Present Indicative: *soy, eres, es, somos, sois, son*;

Present Subjunctive: *sea, seas, sea, seamos, sedáis, sean*;

Imperfect Indicative: *era, eras, era, éramos, érais, eran*;

Imperfect Subjunctive: *fuera, fueras, fuera, fuéramos, fuerais, fueran*.

In the feature sets associated with each rule specifically pertaining to *ser* in the proposed model, the notation *be/ser* is used as a shorthand notation to represent all the semantic notions associated with Spanish *ser*, but it is important to remember that this is only an abbreviated, concrete notation of a complex abstract concept. Both the English *be* and the Spanish *ser* are used in this shorthand notation to show that the given form is lexically

specified for the verb *ser*, but these forms are not directly transferable to all contexts in which a form of *be* would be used in English. Just as stated earlier regarding the use of grammatical notations in the feature sets, this *be/ser* notation is just an abbreviation for the many semantic values possible for *ser* depending upon the context.

In the case of *soy* (*be/ser*, 1 Singular Present Indicative), the disjunctive ordering of rules in each block precludes the application of Rule 7 in Block I of the model shown in Figure 5.8, as well as any others that might seem contradictory, since a more specific rule has already applied in Rule 1 of Block I. Instead of completely renumbering the blocks, the new block added on at the beginning is referred to here as ‘Pre-Block I’, so that the other blocks in this set can maintain the same numbering as the blocks presented above in Figure 5.2 for categorical/regular verbs.

The idiosyncratic/suppletive model for *be/ser* presented below in Figure 5.8 is the same as the ‘regular’ model presented above as the Basic Model in Figure 5.2, except that the feature sets for *be/ser* are superimposed on the Basic Model, including the very specific suppletive lexical stems and/or forms that must apply before the less specific feature sets of the other forms may apply. For instance, Rule 5 in Block I precludes the application of Rule 10 of the same block, even though *ser* is a [+ *e/i*-class] verb as specified in the feature set for Rule 10. As before, phonemic representations are used for the constants in the model. In the model presented in Figure 5.9, however, the phonemic notation for all of the constants happens to also be the orthographic representation of the same, except for /*ye*/, which represents orthographic <*ie*>, and /*βa*/, which represents <*ba*>.

FIGURE 5.8. Idiosyncratic/suppletive model with *be/ser* added to basic model, yielding Present Indicative, Present Subjunctive, Imperfect Indicative, and Imperfect Subjunctive forms for categorical/regular verbs and for idiosyncratic/suppletive *be/ser*.

Pre-Block

$\left[\begin{array}{l} + be/ser \\ + Pres \\ + Subj \end{array} \right]$
/X/ = /se/

Block I

1) $\left[\begin{array}{l} + be/ser \\ + 1 Sg \\ + Pres \\ + Indic \end{array} \right]$	2) $\left[\begin{array}{l} + be/ser \\ + 2 Sg \\ + Pres \\ + Indic \end{array} \right]$	3) $\left[\begin{array}{l} + be/ser \\ + 3 Sg \\ + Pres \\ + Indic \end{array} \right]$	4) $\left[\begin{array}{l} + be/ser \\ + Pres \\ + Indic \end{array} \right]$	5) $\left[\begin{array}{l} + be/ser \\ + Imperf \\ + Indic \end{array} \right]$	6) $\left[\begin{array}{l} + be/ser \\ + Imperf \\ + Subj \end{array} \right]$
/X/ = /soy/	/X/ = /ere/	/X/ = /es/	/X/ = /so/	/X/ = /era/	/X/ = /fue/

7) $\left[\begin{array}{l} + 1 Sg \\ + Pres \\ + Indic \end{array} \right]$	8) $\left\{ \left[\begin{array}{l} + Pres \\ + Subj \\ + a\text{-class} \end{array} \right] \right\}$ + e/i-class	9) $\left\{ \left[\begin{array}{l} + Pres \\ + Subj \\ + e/i\text{-class} \end{array} \right] \right\}$ + a-class	10) $\left[\begin{array}{l} + Imperf \\ + Indic \\ + e/i\text{-class} \end{array} \right]$
/X/ → /X + o/	/X/ → /X + e/	/X/ → /X + a/	/X/ → /X + ia/

11) $\left[\begin{array}{l} + Imperf \\ + Subj \\ + e/i\text{-class} \end{array} \right]$	12) $\left[\begin{array}{l} + \{ 1 Pl / 2 Pl \} \\ + i\text{-class} \end{array} \right]$
/X/ → /X + ye/	/X/ → /X + i/

Block II

1) $\left[\begin{array}{l} + \text{Imperf} \\ + \text{Indic} \\ + a\text{-class} \end{array} \right]$	2) $\left[\begin{array}{l} + \text{Imperf} \\ + \text{Subj} \end{array} \right]$
$/X/ \rightarrow /X + \beta a/$	$/X/ \rightarrow /X + ra/$

Block III

1) $\left[+ 2 \text{ Sg} \right]$	2) $\left[+ 1 \text{ Pl} \right]$	3) $\left[+ 2 \text{ Pl} \right]$	4) $\left[+ 3 \text{ Pl} \right]$
$/X/ \rightarrow /X + s/$	$/X/ \rightarrow /X + mos/$	$/X/ \rightarrow /X + is/$	$/X/ \rightarrow /X + n/$

It is important to note here the difference between the symbols = and \rightarrow in the above model. The first, of course, means ‘equals’, while the second means ‘becomes’. For categorical/regular verbs, the stem $/X/$ entering the model is the verb minus the infinitive ending, such as $/abl/$, $/kant/$, $/kom/$, $/\beta i\beta/$ (*habl*, *cant*, *com*, *viv*), although, as noted above, there is not necessarily any psychological reality in considering the infinitive as any sort of base form. It is mentioned here only as a reference point. In the case of idiosyncratic/suppletive forms, the unique whole forms (such as *soy*) require a separate block (pre-Block I), which, complying with the conjunctive ordering stipulation, applies before proceeding to the next block. Block I, the next block after the Pre-Block, contains some idiosyncratic/suppletive forms, which may subsequently feed into later blocks. The ‘equals’ symbol (=) shows that these idiosyncratic/suppletive forms are themselves the $/X/$ variables for the specified verb rather than the regular lexical stem that enters for most verbs.

Other 1 Singular Present Indicative forms with /oy/ instead of just /o/, such as *doy*, *estoy*, *voy*, would each appear as a separate rule in Block I, as does *soy*, and this more specific rule would preclude the application of the rule that adds /o/ for the 1 Singular Present Indicative. There is another group of ‘exceptional/regular’ verbs that have velar insertion in the 1 Singular Present Indicative and in all forms of the Present Subjunctive, such as *tengo*, *vengo*, *traigo*, *conozco*, *parezco*, *ofrezco*, *plazco*, *conduzco* (all 1 Singular Present Indicative) and *tenga*, *venga*, *traiga*, *conozca*, *parezca*, *ofrezca*, *plazca*, *conduzca* (all 1 Singular and 3 Singular Present Subjunctive). The stems for these ‘irregular’ forms would need to be specified in a Pre-Block, as in Figure 5.9, according to their respective feature sets as listed in the Appendix. It should be noted that all of the stems in the lexical stem set for a given verb are actually included in the Pre-Block, along with their respective feature sets, thus determining what stem feeds into the first block for any given form.

Figure 5.9 shows a sample Pre-Block for the verb *tener* ‘have, possess’, with velar insertion in the stem in Rule 1 for the Present Subjunctive and the 1 Singular Present Indicative forms. Since the notation in the first rule-and-feature set includes curly brackets and square brackets, this feature set means that the resulting form will be Present tense and, in addition, must also be either 1 Singular Indicative or any form of the (Present) Subjunctive, since regular brackets mean that all enclosed features must apply, while curly brackets signify mutually exclusive features. The other stems variants for *tener* are also shown here, since they are the other possible stem variants that can enter the first regular block of the model. Rules 4 and 5, since they are more general than the Rules 1, 2, and 3, only apply in forms for which the other stems do not apply, since the

rules within each block are disjunctively ordered. Note that the infinitive stem /ten/ is just one of the five stem variants for *tener*. However, there is no need to specify [+ Infin] in the feature set for /ten/, since this stem applies to all forms other than the more specific contexts specified in Rules 1, 2, and 3, with the morphophonological stipulation that it applies only when the stem syllable is not stressed.

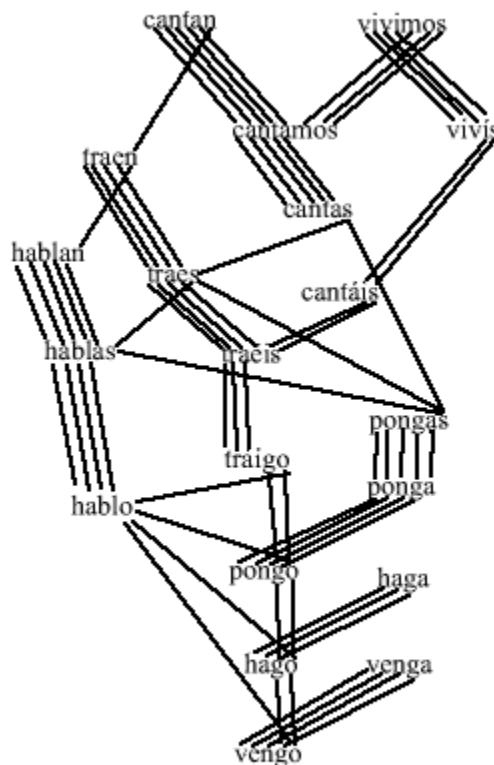
FIGURE 5.9. Pre-Block for the verb *tener*, with velar insertion in one stem alternant.

1) $\left. \begin{array}{l} + \text{ Pres} \\ \left\{ \begin{array}{l} \text{1 Sg Indic} \\ \text{Subj} \end{array} \right\} \\ + \end{array} \right\}$	2) $\left. \begin{array}{l} \left\{ \begin{array}{l} + \text{ Imperf} \\ + \text{ Subj} \end{array} \right\} \\ \left\{ \begin{array}{l} + \text{ Pret} \\ + \text{ Indic} \end{array} \right\} \end{array} \right\}$	3) $\left\{ \begin{array}{l} + \text{ Fut} \\ + \text{ Cond} \end{array} \right\}$	4) Syllable [+ stress]	5) Syllable [-stress]
/X/ = /teng/	/X/ = /tuβ/	/X/ = /tendr/	/X/ = /tyen/	/X/ = /ten/

According to Bybee's (1985) analogical model, as explained in Chapter 3 of the present analysis, all of the verbs with velar insertion would be stored together in the lexicon, with links between their analogous forms. This group includes the 1 Singular Present Indicative forms *tengo*, *pongo*, *traigo*, *hago*, *vengo*, as well as all the Present Subjunctive forms with these same stems, as in *tengas*, *pongas*, *traigamos*, *hagáis*, *vengan*. These 1 Singular Present Indicative forms with /go/ would also have semantic and phonological links to other 'regular' 1 Singular Present Indicative forms with /o/, such as *hablo*, *canto*, *vivo*. The 1 Singular Present Indicative forms with /oy/ would not have identical phonological links to the /o/ and /go/ forms, but they would still have semantic links to all 1 Singular Present Indicative forms, that is, with all forms that share

that same semantic/grammatical feature of 1 Singular Present Indicative. Figure 5.10 shows a small sample of such semantic/phonological connections among verb forms.

Figure 5.10. Sample analogical connections among verb forms, including some with velar insertion in the stem (*traigo, pongo, hago, vengo*). The lines between forms show semantic and phonological identity for each connected phoneme/grapheme. (Conventional spelling is used here for the sake of easy recognition of the forms; therefore, the <h> remains in spite of being unpronounced, and the <c> and <v> are used instead of /k/ and /β/.)



In the proposed model all verbs with idiosyncratic/suppletive forms are treated in essentially the same manner as *ser* or *tener* above. Any unique forms (such as *soy*) that do not conform to any pattern exhibited by other forms of the given verb must be presented in a separate rule in the first block. However, a stem alternant that can feed into the rules in the blocks may require a separate Pre-Block, such as */se/* (as in Figure 5.8) or */teng/* (as in Figure 5.9), since these forms serve as stems for several different forms instead of being a single, unique, complete suppletive form. It is for this reason that such verbs can be referred to as ‘exceptional/regular’.

A full version of Block I for these idiosyncratic/suppletive verbs would include, in addition to all the rules already present in Block I for categorical/regular verbs, any suppletive forms that exhibit some regularity in their irregularity, thus accounting for the ‘exceptional/regular’ verb group. For example, if we posit */X/ = /fue/* for the feature set that includes [*be/ser Imperf Subj*], then this variable */X/* can proceed to Block II, where it will add the regular Imperfect Subjunctive affixes as per Rule 2 of that block:

/X/ → /X + ra/, thus yielding */fuera/*.

This form then exits Block II and proceeds to Block III, where it adds the regular PN affixes to produce all of the complete imperfect subjunctive forms:

For 2 Sg, */X/ → /X + s/*, producing *fueras*,

For 1 Pl, */X/ → /X + mos/*, producing *fuéramos*

For 2 Pl, */X/ → /X + is/*, producing *fuérais*,

For 3 Pl, */X/ → /X + n/*, producing *fueran*.

It should be noted that for 1 Singular and 3 Singular */fuera/*, the final form that exits the blocks is $/X/ \rightarrow /X + \emptyset/$, meaning that the form exits the block exactly as it entered, with nothing added, although these forms are not explicitly listed. They need not be listed because, as stated in the initial explanation of the model, the form that enters any given block is the form that exits that block if there are no applicable feature sets in that block. In traditional terms, the 1 Singular and the 3 Singular forms have ‘zero morphs’, but there is no need to specify a feature set for them in this last block. The $/X/$ variable that enters the block, *fuera* in this case, exits the block just as it entered.

5.12. An a-morphous, morphololexical approach.

If we treat morphemes according to a structuralist, word-analysis centered, agglutinative view of morphology, called the Item-and-Arrangement (IA) theory by Hockett (1958), then the above blocks would roughly correspond to the following ordering of verbal morphemes, though not exactly in every case:

root + class marker + TMA marker + PN marker,

where TMA = tense, mood, and aspect, and P/N = person and number.

This is the order of Bull’s three-part analysis of Spanish verbs (1965: 112-120), which is essentially an Item-and-Arrangement approach. Since there is some mixing between the class marker level and the TMA marker level (see Blocks I and II), thus complicating a straight Item-and-Arrangement approach, this proposed blocked model is a better alternative descriptive analysis of Spanish verbs. Bull’s analysis, however, is intended as a pedagogically useful description of Spanish verbs and is still a useful analysis for teachers. Since Spanish inflectional morphology is mostly linear, albeit fusional, there

seems to be no need for a truly nonconcatenative model such as that proposed by McCarthy (1982) as there may be for other languages such as Arabic. There is indeed a need to deal with the fusional aspects of Spanish, however, and the present model does so.

In most languages the grammatical functions of tense, mood, aspect, person, and number are not easily broken down into separable formal units with one-to-one correspondence between the signified and the signifier. Spanish, for example, is a fusional synthetic language, rather than being strictly agglutinative, since there is not always a one-to-one correspondence between form and function as there would be in a hypothetical, strictly agglutinating language. In other words, there is not always an entirely separate phonological string for each grammatical feature to be represented. There are instances of multiple exponence, where one form represents several different morphological categories (in other words, where there is a one-to-many ratio between form and function), such as in *hablara*, in which the /ara/ string represents verb class and mood as well as tense, or as in preterite *habló*, where /ó/ represents tense, mood, person, and number all at once. A strictly agglutinating language would have entirely separate (and separable) phonological strings for each grammatical function associated with a given concept, such as person, number, tense, aspect, and mood.

Spanish also exhibits syncretism, which means using a single form to represent two or more different functions, but not at the same time. This also represents a one-to-many ratio between form and function, although slightly different from multiple exponence explained above. Spanish *habla*, for example, signifies both 3 Singular

Present Indicative, and 2 Singular Imperative. Similarly, /s/ in English is the 3 Singular verb marker and the noun possessive marker, as well as the noun plural marker.

We also see some overlapping exponence in Spanish, meaning that there is a many-to-one correspondence between form and function in some cases. In overlapping exponence, there may be more than one form or morpheme (whether simply a phonological string, or a process) that is the signifier for a given grammatical category. In other words, a single category may be realized in more than one way within a word. For instance, there is a group of so-called irregular Preterite verbs in which there is a special stem used both in the Preterite Indicative and the Imperfect Subjunctive, such as in *andar*, shown in Figure 5.11 below. There are others of these irregular Preterite verbs, such as *tener*, *poner*, *saber*, and others, which have alternate Present Indicative forms as well, as discussed in previous sections. A Pre-Block for *tener*, for example, is shown in Figure 5.10 above, and Figure 5.12 below shows the stem alternant that is found in both the Preterite Indicative and the Imperfect Subjunctive.

FIGURE 5.11. Forms for the verb *andar*, showing the stem alternate /*anduβ*/ found in Preterite Indicative and Imperfect Subjunctive forms only.

PRES	PRET INDIC	IMPERF SUBJ
ando , andas, anda, andamos, andais, andan	anduve , anduviste, anduvo, anduvimos, anduvisteis, anduvieron	anduviera , anduvieras, anduviera, anduviéramos, anduviérais, anduvieran

FIGURE 5.12. Forms for the verb *tener*, showing the stem alternate /tuβ/ found in Preterite Indicative and Imperfect Subjunctive forms only.

PRES	PRET INDIC	IMPERF SUBJ
tengo, tienes, tiene, tenemos, tenéis, tienen	tuve, tuviste, tuvo, tuvimos, tuvisteis, tuvieron	tuviera, tuvieras, tuviera, tuviéramos, tuviérais, tuvieran

For the Preterite forms there is essentially a two-to-one ratio between form and function, since there are two phonological indicators of past tense meaning. The special stem signifies preterite indicative function at the same time that the respective Preterite Indicative endings also do so. In fact, these forms actually exhibit a many-to-many ratio, since the affixes in the form represent person and number as well as tense. In the case of the alternate special stem for these verbs, syncretism is also seen, since the alternate special stem signifies both preterite indicative and imperfect subjunctive functions.

One reason that an a-morphous, or morpheme-less, grammar works better than a traditional Item-and-Arrangement model is because of the cases of multiple exponence and fusion, syncretism, and overlapping exponence. The model proposed in the present study, like Anderson's model, regards the morpheme as a process, rather than as a discrete phonological string as seen in the traditional definition of the morpheme as the minimal linguistic sign. However, the morpheme in the a-morphous sense, as a process, could still be a minimal unit of meaning, hence a minimal linguistic sign. If there is an abstract level for the concept of 'morpheme', as discussed above regarding the abstract

nature of the lexeme, then it should not be a problem to accept an abstract process as a minimal linguistic sign as well. In other words, the process, represented by the rules in the above model, from which *hablas* is derived for the feature set that includes 2 Singular Present Indicative, could be a minimal unit of meaning just the same as any concrete phonological string might be. If this is so, then this process-as-morpheme approach could account for nonlinear changes such as those in the stem-changing verbs above or other changes in the base lexeme, such as the alternate stem for the irregular Preterite forms discussed above.

Treating morphology in this a-morphous, morpheme-less, process-oriented way might seem to be just an Item-and-Process (IP) theory of morphology, but there are differences. Spencer (1997) points out that in an IP model, although morphemes are regarded as processes as described above, there must, in addition, also be a distinction between underlying base forms and the forms that are derived by the application of certain processes, such as in the irregular Preterite verb forms above. (50.) In IP theory, there needs to be an underlying form, meaning that one of the alternate stems needs to be designated as the underlying form upon which the processes operate to derive the other forms. In fact, Bybee (1988) states that “The defining characteristic of an IP model is that it sets up one underlying form for alternating allomorphs and derives the surface forms by applying feature-changing rules to the underlying form” (119). For example, /ten/ might be designated as the underlying form for *tener*, and the alternates (traditionally called allomorphs) /teng/ [{Pres Subj; 1 Sg Pres Indic}], /tyen/ [{2 Sg, 3 Sg, 3 Pl} Pres Indic], /tuβ/ [{Pret Indic; Imperf Subj}], and /tendr/ [{Fut; Cond}] would be the surface forms along with /ten/ (the default form for all other cases besides the ones

given above for the alternate stems), or else there may be even be an underlying form that has no surface form at all in an IP approach (Spencer 1997: 102; Bybee 1988: 120).

For Anderson (1995), a ‘lexical stem’ can be regarded as “an idiosyncratic association of phonological, syntactic, and semantic properties,” although there may be more than one “lexicalized set of phonological properties” associated with the same set of semantic and syntactic properties (122). Thus, according to Anderson’s model, which is essentially a generativist model, the lexicon provides a single stem form that underlies all of its phonological variants. In Anderson’s view, there is one lexical stem, which may have several suppletive variants, although some may be only partially suppletive. Anderson refers to this set of alternate forms of a lexical entry (classically viewed as ‘allomorphs’) as a ‘lexical stem set’, defined by Anderson as “a group of phonologically distinct stems ... with the same syntactic requirements and semantic interpretation, each associated with its own (partial) set of morphosyntactic properties” (133).

The present analysis, like an Item-and-Arrangement (IA) model, does not regard any one stem as an underlying form, nor does it see any one form as basic over the other stem alternants. All of the stem alternants in the present model are regarded as equal, with none taking precedence over the other except in the sense that a more specific rule takes precedence over a more general rule. One stem may also be more frequently occurring than another, but even this fact does not make one stem more ‘basic’ than any other. However, although the present analysis treats all stems as equal with none considered as ‘underlying’, the proposed model is not an IA model, because it does not regard morphemes as ‘items’. The present model is a Word-and-Paradigm (WP) model,

because it “chooses the word (rather than the stem) as the unit of lexical representation and includes in the lexicon a full listing of paradigms” (Bybee 1988: 120).

Bybee disagrees with the IP assumption that “rules and representations are discrete and distinct elements of the grammar” (123). The present model, like Bybee’s, regards rules and representations as essentially the same thing. Rules and representations are seen as a continuum, rather than as two independent entities. Along this continuum, the completely suppletive representations are at one end, while the highly regular, generalized patterns (which can be seen as ‘rules’ of a sort) are on the other end. In between along the continuum are varying degrees of ‘regularity’, including forms that are considered ‘exceptional/regular’ in Skousen’s terms (1989: 6). Some Spanish Preterite forms would be in this exceptional/regular group, since they require special stems but are analogous with each other in their inflectional affixes. The ‘rules’ in the present model, then, are not true rules in the sense of being independent from the representations, but rather are just descriptions of the representations themselves, that is, descriptions of the forms found in the lexicon. The ‘continuum’ nature of the proposed model, consistent with Bybee’s view, is evident in the fact that suppletive forms are treated as ‘most specific’ rules/representations that apply before the more general, or less specific, rules that apply only when the more specific rules cannot apply.

Bybee (1988) contends that “part of the representation has to be built into the rule. And of course, so-called rules governing suppletion are nothing more than representations” (123). For example, Rule 1 of Block I in Figure 5.8, with the feature set [+ *be/ser*, + 1 Sg + Pres + Indic] is really just a description of the suppletive form /*soy*/, since the rule is written as /X/ = /*soy*/. Even a rule such as /X/ → /X + a/ for the feature

set [+ Pres + Subj + *e/i*-class] (Rule 3 of Block I in Figure 5.2 above) is really just a description of a whole group of verb forms that end in /a/ and have a semantic/grammatical value of [+ Pres + Subj + *e/i*-class]. Another way of stating this assertion is that the rule /X/ → /X + a/ for the feature set [+ Pres + Subj + *e/i*-class] is a description of the analogical connections among words that end in /a/ and that have a semantic/grammatical value of [+ Pres + Subj + *e/i*-class]. The feature set shows the semantic connection among such forms, and the /a/ ending shows the phonological connection.

5.13. Summary and conclusions.

Although the model proposed in the present analysis is formally based on Anderson's Word-Formation Rules (1995), there are theoretical differences. The proposed model is also a-morphous, like Anderson's model, since it does not recognize separate and separable morphemes as 'items' in the lexicon. But unlike Anderson's model, the proposed model does not regard the rules in the model as rules in a generative sense, nor does it regard these rules as taking place outside of the lexicon, independent from the representations themselves. Concurring with Bybee, the model proposed in the present study "does not have a lexicon and a morphological component as separate components of the grammar. Rather the model has only a lexicon" (Bybee 1988: 125). In the present model, as in Bybee's model (explained further in Chapter 3 of the present analysis), "morphological rules and lexical representations are not separate from one another. Rather, morphological and morphophonemic rules are patterns that emerge from the intrinsic organization of the lexicon" (125). For Bybee and for the present model, the

lexicon consists of networks of connections “among stored elements of knowledge and the ability to register the frequency of individual items and patterns,” which Bybee refers to respectively as “lexical connection and lexical strength” (125). Like Bybee’s model, the proposed model is morpholexical and analogical, rather than extra-lexical and generative, because it regards rules as patterns of organization in the lexicon, with rules and representations seen as essentially the same thing.

This chapter has presented a detailed introduction to the Basic Model for categorical/regular verbs in the Present Indicative, Present Subjunctive, Imperfect Indicative, and Imperfect Subjunctive. It has also discussed the proposed model’s treatment of irregular verbs, including the categories designated respectively as phonologically variant, exceptional/regular, and idiosyncratic/suppletive.

CHAPTER 6

FURTHER CONSIDERATIONS

6.0. This chapter deals first with lexical gaps and defective verbs and then presents the models for nonfinite forms (Infinitive; Present and Past Participles) and Future and Conditional, including phonologically variant Past Participle, Future, and Conditional forms. The Basic Model for categorical/regular Preterite is then presented, followed by the exceptional/regular Preterites and idiosyncratic/suppletive Preterites. The present analysis designates a group of categorical/regular Preterites as ‘allo-class’ based on their having ‘allo’-stems (that is, allomorphs). There is also a subset of this allo-class, which is designated as ‘*j*-allo-class’ because verbs in this group contain the grapheme <*j*> in their respective ‘allo’ stems. Compound forms are then considered, with models given for three-part as well as two-part compounds. The Appendix following the final chapter presents a list of some of the lexical stem sets (‘allomorphs’ in traditional terminology) that feed into the proposed model, along with the respective feature sets that determine which alternant is chosen for a given inflected verb.

6.1. Lexical gaps.

Nonexistent, unattested forms, referred to as lexical gaps, are forms that for one reason or another do not appear (or at least are not common) in any dialect or register of the language even though they are grammatically possible and might hypothetically be produced by analogy to existing forms in the lexicon. If Word-Formation Rules were to apply completely independent of the semantic component of the grammar, then all of

these unattested forms would be a problem. However, since each of the word formation rules is bound to a set of semantic features that represent the *signifié* part of the two-part linguistic sign, then some of these lexical gaps are not a problem. The ‘allowed’ forms are specified in the lexical entry in the sense that in the blocks of the model there are no semantic feature sets that contain the disallowed grammatical functions. This means that for those verbs that have lexical gaps for semantic reasons, there are fewer rules in the blocks that apply to them.

The present analysis divides these verbs into three groups: 1) weather verbs, 2) tense-restricted verbs, and 3) true defective verbs. This latter group, which might be considered the only ‘true’ lexical gaps, refers to verbs for which there is clearly no semantic reason for the missing forms. For verbs in this latter group, there may be forms that are semantically possible, such as the 1 Singular Present Indicative form of *abolir* ‘abolish’ but that do not exist in the actual language, or at least are not well attested. Thus, in such a case a form may exist at the *signifié* level but not at the *signifiant* level. This type of lexical gap will be discussed further in Section 6.3 below.

6.2. Weather verbs.

Weather verbs, when used in the impersonal, literal semantic sense, only occur in the third person singular, as in *llueve* (‘it rains’, 3 Singular Present Indicative), *llovió* (‘it rained’, 3 Singular Preterite Indicative), *llovía* (‘it rained/was raining’, 3 Singular Imperfect Indicative), *lloverá* (‘it will rain’, 3 Singular Future Indicative). Other examples of unipersonal verbs (in their literal sense) are *nevar* ‘snow’ and *helar* ‘freeze’. Although it is grammatically possible to produce 1 Singular, 2 Singular, 1 Plural, 2

Plural, and 3 Plural forms of this verb (as in Pres Indic *lluevo, lluevas, lloremos, llovéis, llueven*), these forms make no sense in the literal, impersonal senses of the words.

Common sense tells us that ‘to rain’ and ‘to snow’ in their literal sense have meaning only in an impersonal sense, thus only in the 3 Singular form, although metaphorically they may be used with other semantic values.

A weather verb such as *llover* may be used in a metaphorical sense, as in “El techo de don José se llueve cada vez que hay tormenta” (‘Don José’s roof rains every time there is a storm’) (Ruiz 2000). In this sense, however, Ruiz points out that the verb *llover* is used in a figurative sense to mean “se pasa, tiene goteras” (‘it leaks, it has leaks’). Ruiz goes on to mention more figurative uses and ‘semantic displacements’, such as in the ‘raining’ of bullets, pine needles, or even starlight.

Entonces, obviamente se trata de desplazamientos semánticos, un tanto figurados, que también nos permiten decir "De pronto nos llovieron balas de todas partes" para decir que hubo una gran "balacera" (shooting) dirigida a nosotros. (‘Then obviously, this is a question of semantic displacements, rather figurative, that allow us to say “Soon bullets were raining on us from all sides”, meaning that there was a profuse shooting directed at us’.) (Ruiz 2000) [The translation is the present author’s.]

Ruiz points out further that such verbs are not true lexical gaps, since in their literal, basic meaning they only make sense in the third person singular. However, allowing for the metaphorical dimension, nothing keeps us from saying, for example, that “al despertarse, el inmenso pino que nos cubría llovió sobre las cabezas soñolientas de Anita y Josefina; (‘upon awakening, the immense pine tree above us rained over Anita’s and Josefina’s

sleepy heads’). [The translation is the present author’s.] We could also say that “era una lluvia clara que no se oponía a que las estrellas siguieran lloviendo su luz amarilla y pálida sobre las piedras donde debían apoyarse los descalzos pies de las rehenes...” (‘it was such a clear rain that it didn’t prevent the stars from continuing to rain their pale yellow light over the rocks where the hostages rested their bare feet’) (Ruiz 2000). [The translation is the present author’s.]

The rule-and-feature blocks in Figure 6.1 represent the literal semantic sense of these impersonal weather verbs, although, as Ruiz makes clear, there is nothing that necessarily precludes the metaphorical use of these usually nonexistent forms. Such metaphorical usage, however, would be specified in the feature set associated with a given form. Although the blocks that produce these weather verbs are separately listed below in Figure 6.1 for illustrative purposes, they need not be represented separately in the entire model as a whole, since the given semantic feature sets would include or preclude certain forms and not others.

FIGURE 6.1. Weather-verb model: Present Indicative, Present Subjunctive, Imperfect Indicative, and Imperfect Subjunctive.

Block I

1) $\left\{ \begin{array}{l} \left[\begin{array}{l} +\text{Pres} \\ +\text{Subj} \\ +a\text{-class} \end{array} \right] \\ +e/i\text{-class} \end{array} \right\}$	2) $\left\{ \begin{array}{l} \left[\begin{array}{l} +\text{Pres} \\ +\text{Subj} \\ +e/i\text{-class} \end{array} \right] \\ +a\text{-class} \end{array} \right\}$	3) $\left[\begin{array}{l} +\text{Imperf} \\ +\text{Indic} \\ +e/i\text{-class} \end{array} \right]$	4) $\left[\begin{array}{l} +\text{Imperf} \\ +\text{Subj} \\ +e/i\text{-class} \end{array} \right]$
$/X/ \rightarrow /X + e/$	$/X/ \rightarrow /X + a/$	$/X/ \rightarrow /X + ía/$	$/X/ \rightarrow /X + ye/$

Block II

1) $\left[\begin{array}{l} + \text{ Imperf} \\ + \text{ Indic} \\ + a\text{-class} \end{array} \right]$	2) $\left[\begin{array}{l} + \text{ Imperf} \\ + \text{ Subj} \end{array} \right]$
/X/ → /X + ba/	/X/ → /X + ra/

Block III

(There is no Block III, which would add PN markers, because the above blocks yield the 3 Singular forms by default, as per the Elsewhere principle.)

6.3. Tense-restricted verbs.

The verb *solear* ‘to be accustomed to, to be in the habit of, to have the custom of’ occurs in all six persons, but some grammar texts list *solear* as occurring only in certain tenses. In this sense, *solear* could be regarded as a tense-restricted verb. Kendrix (1990) lists *solear* forms only for the Simple Present, the Imperfect Indicative, the Present Subjunctive, and the Present Perfect Indicative, but not for the other four simple tenses or the other six compound tenses. Noble and Lacasa (1997), however, also show Preterite Indicative and Imperfect Subjunctive forms for the verb *solear*. A semantic reason seems somewhat plausible for some of the tense gaps, such as future, but not for others. The very fact that Noble and Lacasa list Imperfect Subjunctive forms for this verb but Kendrix does not shows that there is disagreement over what forms are ‘grammatical’. The only real test of the existence of these or any other forms would be to check the frequency of their occurrence in a linguistic corpus of actual Spanish usage.

Like the suppletive forms above whose semantic features specify certain conditions, it may be that the forms of the verb *sober* would have restrictions specified in their associated semantic feature sets. In other words, there may be potential *signifiants*, such as future forms, that do not exist because there is no *signifié*, meaning that such a concept does not exist because it is not semantically feasible.

If we were to accept Kendrix's tense restrictions as accurate for the verb *sober*, the model would have to eliminate any and all feature sets containing the feature Imperfect Subjunctive for such a tense-restricted verb. For Noble and Lacasa's version, however, there would be no need to alter the basic model at all, but the models for the compound tenses and for the Future and Conditional would have to be modified for Noble and Lacasa's version, as they would for Kendrix's version as well. However, rather than specifying a separate model, a more plausible and economical way of handling such tense-restricted verbs would be to present the entire model as for any other verb, but to regard the unattested forms as blocked by the semantic feature sets.

6.4. True defective verbs.

All of the examples of lexical gaps given so far have been verbs for which semantic reasons may account for the gaps, and those gaps can be explained by semantic feature sets that preclude the unattested forms. However, as in most languages, there are in Spanish verbs that lack certain forms for no apparent semantic reason. According to Albright (2000: 39), these true defective verbs fall into two categories:

- 1) Verbs that avoid forms in which the stress would fall on the root (1, 2, 3 Singular, and 3 Plural of the Present Indicative, and all forms of the Present Subjunctive).

- 2) Verbs that are lacking only the 1 Singular Present Indicative, and all forms of the Present Subjunctive.

The first of these groups contains the verb *abolir* ‘to abolish’, which happens to have only those forms containing the high front vowel /i/ or the high front glide /y/, as in: *abolimos, abolís* (the only attested Present Indicative forms); *abolía, abolías, abolía, abolíamos, abolíais, abolían* (Imperfect Indicative); *aboliera, abolieras, aboliera, aboliéramos, abolierais, abolieran* (Imperfect Subjunctive). All forms of the Preterite Indicative, the simple Future, the Conditional, and all compound tenses happen to contain the grapheme <i> as well (hence the phoneme /i/ or /y/), and are likewise among the attested forms. In the Imperative forms, only the 2 Plural form is attested, and it happens to be the only imperative form containing the grapheme <i>: *abolid*.

A similar verb is *embair* ‘trick’, which also has only those forms containing /i/ or /y/. For this verb we cannot say that we will include only those forms with the grapheme <i>, however, because of the spelling change in the some forms, as in *embayó, embayeron* (3 Singular and 3 Plural Preterite Indicative) and *embayera, embayeras, embayera, embayéramos, embayerais, embayeran* (Imperfect Subjunctive) and *embayendo* (Present Participle). The phoneme /i/ is still present in all of these forms, however, even though it becomes a glide when it occurs before another front vowel (as it does in *aboliera*, above) and has a spelling change when it occurs in unstressed form between two vowels (as in *embayeron*).

As for these verbs in which the only attested forms are those containing /i/ or /y/, it could be proposed that there is some phonological restriction specified in the feature

set, but there may be a simpler explanation. It may be pure coincidence that the attested forms all contain /i/, and this fact may, therefore, be nothing more than descriptive rather than causal. There are other verbs, such as *balbucir* ‘stammer’, that are similar to *abolir* but do have some forms that do not contain /i/, such as *balbucés*, *balbuce*, *balbucimos*, *balbucís*, *balbucen* (2 Singular Present Indicative, 3 Singular Present Indicative, 1 Plural Present Indicative, 2 Plural Present Indicative, 3 Plural Present Indicative, respectively). We should, therefore, perhaps look for another common thread that may prove more explanatory, and Albright does just that in his proposal of two groupings for true defective verbs.

All of the gaps in the first group above involve deciding whether or not to diphthongize the mid vowel. Shall we say **abuelo* or **abolo* for the 1 Singular Present Indicative form? For *agredir* ‘assault’, is it **agriado* or **agredo*? It may simply be that speakers are hesitant to use these forms because they are not sure which form to use, and consequently, none of them is well attested. Albright (2000) suggests a ‘gradient effect’, with “a whole array of reluctance between complete refusal to answer [i.e., provide a specified ‘gap’ form when given an attested form] (as for *asimos* → ___) and completely automatic responses (as for *cantamos* → *canto*)” (42).

The second group of gaps or defective verb paradigms given above involves a similar phonological dilemma, that of deciding whether or not to insert a velar consonant in the 1 Singular Present Indicative form and in all of the Present Subjunctive forms. This velar consonant insertion occurs in such common verbs such as *conocer* ‘know, be acquainted with’: *conozco* (1 Singular Present Indicative) and *conozca*, *conozcas*, *conozca*, *conozcamos*, *conozcáis*, *conozcan* (Present Subjunctive). But what shall we do

with *balbucear* or *asir*? Is it **balbuzco* for 1 Singular Present Indicative? **Asgo* or **aso* for *asir* ‘grasp’? If the speaker is not sure, maybe because these verbs are not very common in usage, then he/she may avoid them altogether, thus reinforcing their ‘defective’ nature in their overall usage. A different word might substitute for the problematic one, maybe a similar word without the phonological dilemma. For example, *balbucear* ‘stammer’ is listed in most dictionaries as well as the problematic *balbucir*. A speaker might choose the alternative with the /ear/ infinitive ending to avoid having to decide whether or not to insert the velar consonant. The 1 Singular Present Indicative form of *balbucear* would simply be *balbuceo*, which presents no phonological dilemma.

It is interesting to note that verb lists and dictionaries are not consistent in their listing of verbs as defective or not, nor are they consistent in what forms should be included and what forms should not. Albright’s (2000: 39-40) experiments did elicit some of these so-called nonexistent forms from native speakers, and, according to Albright, these forms can be found to some degree in a corpus such as LEXESP as well. Perhaps what this points out is that there is not a clear binary distinction between existent and nonexistent forms, or between grammatically correct and grammatically incorrect forms. Thus we cannot say with absolute certainty what verbs are genuinely defective, nor can we state certainly what and where the gaps actually are.

Do these so-called defective verbs need a separate model, then? Perhaps not. It is possible that the models may exist in the mind of the speaker, but their existence is dependent upon the feature sets that have been built up in the given speaker’s mind. If there are no associated feature sets for the given word, meaning that the speaker is not familiar with that word, then the speaker would probably not generate any of the forms

for that word. If an artificial experiment attempted to elicit the forms of an unknown or nonce verb, the speaker would probably make analogous connections to a phonologically similar verb and then use some of the feature sets associated with that verb to produce the unknown forms. But since there is no semantic reason to explain these lexical gaps, then there is no reason to make absolute statements about the defective nature of any given verb.

In fact, all of the verb models presented here may actually be different in the minds of different speakers, although the differences may be small and may have varying degrees of significance. Since it is highly unlikely that two people have exactly the same set of semantic features associated with a word, it follows that there may also be differing subcategorization specifications or differing specifications regarding when to use what allomorph or stem alternant. These variations, and the cause of their existence (or lack thereof), would be a question of discussion for dialectologists or sociolinguists.

Ruiz (1986) comments that “It is unquestionable that no two physical sounds are identical” and that “the concept of a static sound, or for that matter, the notion of a synchronic analysis as something entirely static is only a convenient although necessary fiction” (112). A linguistic analysis such as the present study of Spanish verbs looks at generalizations in the system as a whole, but we cannot ignore the very real fact that there is variation in the system, and that variation is an inherent part of the system. Language change and variation are the norm; homogeneity and maintenance are the exception. As Milroy (1992) says, “Languages are not in reality completely stable or uniform, and there is absolutely no reason why they should be” (4).

6.5. Basic Model for Infinitive, Present Participle, Past Participle, Future, and Conditional forms.

This section presents the blocks for a model representing nonfinite forms and the Conditional forms for Spanish verbs, although as noted earlier, Alvar and Pottier (1983) maintain that nonfinite forms are not true verbs. These authors comment:

El lexema *cant-* unido al morfema *-ar* constituye la forma tradicionalmente llamada «presente del infinitivo», pero esta forma ... carece de especificidad verbal y funciona como sustantivo (*el cantar, los cantares*); entonces *-ar* es un elemento disponible por cuanto pertenece al paradigma verbal, pero no es estrictamente verbo (no tiene tiempo ni persona). ('The lexeme *-cant* joined to the morpheme *-ar* yields the form traditionally called 'present infinitive', but this form ... lacks verbal specificity and functions as a noun ('singing/song, singings/songs'); *-ar*, then, is an available element inasmuch as it belongs to the verbal paradigm, but it is not strictly a verb (it has neither tense nor person.) (215). [The translation is the present author's.]

Alvar and Pottier's comment reinforces the earlier mentioned difficulty in distinguishing derivational and inflectional processes in morphology, as shown in Bybee's (1985: 12) continuum of lexical vs. syntactic expression. Bybee regards inflection and derivation not as separate categories, but as part of a continuum, with full lexical expression on one end, syntactic expression at the other end, and derivation and inflection in between. Bybee explains lexical expression as "the most highly fused means

of expression,” and syntactic or periphrastic expression as “the most loosely joined” means (1985: 12), as shown in the following continuum:

lexical --- derivational --- inflectional --- free grammatical --- syntactic
 < -----
 greater degree of fusion

Infinitive and Participle forms would be more toward the left end of this continuum, since Infinitives function essentially as nouns, Past Participles as adjectives, and Present Participles as adverbs.

Nevertheless, the present analysis includes these nonfinite forms, since they are traditionally included with verb paradigms. The blocks of rule-and-feature sets shown in Figure 6.2 produce all of the Infinitive, Present Participle, Past Participle, Future, and Conditional forms for categorical/regular *a*-class, *e*-class and *i*-class verbs. The first two blocks below are numbered I A and I B, respectively, in order that Block III is always the block with the PN features in all of the models presented in this analysis.

FIGURE 6.2. Basic model for Infinitive, Past Participle, Present Participle, Future, and Conditional forms

Block I A

1) $\left\{ \begin{array}{l} \text{Infin} \\ \text{Past Part} \\ \text{Pres Part} \\ \text{Fut} \\ \text{Cond} \\ + a\text{-class} \end{array} \right\}$	2) $\left\{ \begin{array}{l} \text{Infin} \\ \text{Fut} \\ \text{Cond} \\ + e\text{-class} \end{array} \right\}$	3) $\left\{ \begin{array}{l} \left\{ \begin{array}{l} \text{Infin} \\ \text{Fut} \\ \text{Cond} \end{array} \right\} \\ + i\text{-class} \\ \left\{ \begin{array}{l} + \text{Past Part} \\ + e/i\text{-class} \end{array} \right\} \end{array} \right\}$	4) $\left[+ \text{Pres Part} \right]$
$/X/ \rightarrow /X + a/$	$/X/ \rightarrow /X + e/$	$/X/ \rightarrow /X + i/$	$/X/ \rightarrow /X ye/$

Block I B

$\left[+ \left\{ \begin{array}{c} \text{Fut} \\ \text{Cond} \\ \text{Infin} \end{array} \right\} \right]$
$/X/ \rightarrow /X + r/$

Block II

1) $\left[\begin{array}{c} + \text{Fut} \\ + \{1 \text{ Sg}, 1 \text{ Pl}, 2 \text{ Pl}\} \end{array} \right]$	2) $\left[+ \text{Fut} \right]$	3) $\left[+ \text{Cond} \right]$
$/X/ \rightarrow /X + \acute{e} /$	$/X/ \rightarrow /X + \acute{a} /$	$/X/ \rightarrow /X + \acute{ia} /$

Block III

1) $\left[+ 2 \text{ Sg} \right]$	2) $\left[+ 1 \text{ Pl} \right]$	3) $\left[+ 2 \text{ Pl} \right]$	4) $\left[+ 3 \text{ Pl} \right]$
$/X/ \rightarrow /X + s /$	$/X/ \rightarrow /X + \text{mos} /$	$/X/ \rightarrow /X + \text{is} /$	$/X/ \rightarrow /X + n /$

5) $\left[+ \text{Pres Part} \right]$	6) $\left[\begin{array}{c} + \text{Past Part} \\ + \text{to-class} \end{array} \right]$	7) $\left[+ \text{Past Part} \right]$
$/X/ \rightarrow /X + \text{ndo} /$	$/X/ \rightarrow /X + \text{to} /$	$/X/ \rightarrow /X + \text{do} /$

Rule 3 in Block IA may need some clarification, although it should be clear from the previous explanation of the function of regular brackets vs. curly brackets. What Rule 3 means is that the resulting form must be either an *e/i*-class Past Participle or one of the following: an *i*-class Infinitive, an *i*-class Future form, or an *i*-class Conditional form.

For an *a*-class verb such as *amar*, the stem /am/ (the only stem in the lexical stem set for *amar*) might enter the first block in the model, specifying the feature [+ Pres Part]. Since Rule 1 of Block I, stating that /X/ → /X + a/, has the matching features [+ Pres Part] and [+ *a*-class], the form exiting Block IA would be /ama/, and this same form will enter Block IB as the new X variable. Block IB has no matching feature sets, nor does Block II. Therefore, the form /ama/, still the X variable at this point, now enters Block III, where it finds a matching feature set in Rule 5. The rule /X/ → /X + ndo/ yields the final form /amando/, represented in conventional orthography as *amando*. It should be noted that although Rule 4 of Block I specifies [+ Pres Part], that rule is blocked for *amar* by the more specific Rule 1, which specifies [+ *a*-class] as well as [+ Pres Part].

Let's look at another example, a 1 Plural Future form for *asistir* 'attend', an *i*-class verb. The stem /asist/ enters the first block, where it finds a feature match in Rule 3: /X/ → /X + i/. This rule results in the form /asisti/, which then exits the block and enters Block IB as the new X variable. There is a match with the feature set for the only rule in Block IB, so that the rule /X/ → /X + r/ results in /asistir/ as the new X variable to enter Block II. Rule 1 of Block II applies, with the rule /X/ → /X + é/ yielding /asistiré/, which now enters Block III as the new X variable. Since Rule 2 in Block III provides a feature match, the rule /X/ → /X + mos/ produces /asistirémos/ as the final form, conventionally spelled *asistiremos*.

The model for Future forms represented in Figure 6.2 reflects somewhat the diachronic development of Spanish future verb forms. In medieval Spanish, the future tense forms consisted of two parts: an inflected form of the auxiliary verb *haber* plus the infinitive of the main verb, as in *hemos comprar* (Bull (1965: 115). The verb *haber* has

lost its stem in modern Spanish, except for the 2 Pl form *habéis*; thus the modern forms are *he, has, ha, hemos, habéis, han*. Modern Spanish also puts the main verb first, with the truncated form of *haber* attached to the infinitive of the main verb, with the silent *h* dropped in the orthographic representation of the form, and the entire stem *hab* is dropped in the 2 Pl form. The resulting forms for *comprar* ‘buy’ are: *compraré, comprarás, comprará, compraremos, compraréis, comprarán*. These forms are historically equivalent to *comprar he, comprar has, comprar ha, comprar hemos, comprar habéis, comprar han*, respectively. In the model in Figure 6.2, Blocks IA and IB form the infinitive of the given verb; then Block II adds /é/ or /a/ (equivalent to *he* and *ha* from *haber*), and Block III adds the PN affixes. In the case of Future and Conditional forms, traditional analyses using morphemes seems descriptively accurate. However, as seen in the idiosyncratic/suppletive forms discussed in the previous chapter and in the Preterite forms discussed later in the present chapter, an a-morphous approach works better. For a theory or model of inflectional morphology to be descriptively accurate for a given language, it needs to be capable of handling all of the cases of verbal inflection, rather than just some cases. Therefore, a theory or model that can handle all tenses, moods, and aspects is more economical, even though it may seem unnecessary for some individual forms or tenses.

6.6. Phonologically variant Future and Conditional forms.

As with the phonological irregularities discussed above for the Basic Model, the phonologically variant forms for future and conditional functions have regular patterns in their ‘irregularity’ as well, and could be treated as cases of stem alternation. There are

thirteen of these phonologically variant Future and Conditional forms, plus the compounds formed with prefixes attached to them, such as *sobresalir*, *suponer*, *proponer*, *convenir*, *rehacer*, although *predecir* does not follow this irregular pattern (*predecir/predeciré/predeciría*).

Figure 6.3 shows the thirteen basic phonologically variant Future and Conditional forms as listed by Ranson (1995), although the grouping shown here is the present author's. Since these variants exhibit generalized phonological patterns, they are grouped here according to the phonological similarities in their patterns.

FIGURE 6.3. Thirteen phonologically variant Future and Conditional forms as listed by Ranson (1995), grouped by the present author according to phonological similarities.

	Infinitive	Future 1 Sg	Cond 1 Sg
1)	caber haber saber poder querer	cabré habré sabré podré querré	cabría habría sabría podría querría
2)	poner tener valer salir venir	pondré tendré valdré saldré vendré	pondría tendría valdría saldría vendría
3)	hacer satisfacer decir	haré satisfaré diré	haría satisfaría diría

The first of these groups loses the /e/ that is present in the infinitive form, possibly because of a natural tendency in languages to delete unstressed syllables. Lathrop (1980) explains that “In Vulgar Latin, pretonic internal vowels (except *a*) disappeared in most

cases” (71), citing examples such as VL (Vulgar Latin) *fab(u)las* → MS (Modern Spanish) *hablas*; VL *an(i)ma* → MS *alma*; VL *sem(i)ta* → MS *senda*; VL *com(pu)tare* → MS *contra*. If we consider this general historical tendency in the development of Spanish, this ‘irregularity’ could be seen as simply an orthographic convention. The unstressed /e/ is deleted in pronunciation, and at some point in the diachronic development of the language has come to be deleted in the orthographic representation as well. However, it should be noted that the so-called variants would most likely be pronounced the same as above, at least in rapid speech, even if we didn’t have this orthographic rule.

The same syncopation of an unstressed vowel is seen in the second group as well, but there is one additional factor. In this group we also see the insertion of the dental consonant /d/ between the final consonant of the stem and the /r/ of the infinitive marker. However, as in the group discussed in the preceding paragraph, this could be seen not as a true irregularity but simply as a spelling convention. If one pronounces **poneré* or **veniré* in casual, rapid speech, for instance, not only does the unstressed vowel of the infinitive ending tend to disappear, but the combination of the alveolar /n/ or /l/ and the alveolar tap /r/ naturally produces an alveolar stop, that is a [d] sound, in between the two consonants. It is difficult if not impossible to pronounce such words otherwise.

Lathrop cites evidence for such a phenomenon in the diachronic development of Spanish, maintaining that “Sometimes a consonant cannot properly adjust to the point of articulation of the following consonant and instead, a third consonant is generated between the two” (99). In fact, *poner* (with its ‘irregular’ Future/Conditional stem /*pondr/*) and several others from the second group above are directly cited in Lathrop’s

examples: VL *pon(e)r * → MS *pondr *; VL *sal(i)r * → MS *saldr *; VL *ten(e)r * → MS *tendr *; VL *val(e)r * → MS *valdr *; VL *ven(i)r * → MS *vendr * (99). The consonant inserted in these examples is [d], because [d] has the same point of articulation as the preceding /l/ or /n/. According to Lathrop, “If the first consonant is alveolar (*l,n*) a *d* is generated; if it is bilabial (*m*), a *b* is generated” (99), as in VL *hum(e)ru* → MS *hombro*; VL *trem(u)lare* → MS *temblar*.

There are certainly words in Spanish with the sequence /nr/, as in *enriquecer* ‘enrich, make rich’ or *enrollar* ‘wind, coil, roll up’, with no [d] insertion in the spelling as seen in the conventional spelling convention of the above group of Future/Conditional stems. However, in *enriquecer* and *enrollar* as well, there is a sort of [d] insertion in that an /r/ following an /n/ cannot be a single tap, but must be pronounced as a multiple trill, which is roughly equal to inserting a [d] sound in between the /n/ and the /r/ (Lathrop 1980: 98). Since orthography is arbitrary to a certain degree in any language, it may be just spelling convention that puts an orthographic *d* in the above verbs but not in words like *enriquecer* or *enrollar*.

A similar articulatory phenomenon is manifest in the English words *dance* and *prince* (Aitchison 1991: 130). The words *prince* and *prints* are homonyms in English, and the words *dance* and *pants* are minimal pairs, although the latter word in each of these pairs is spelled with a *t* between the /n/ and the /s/, and the former is not. The spelling difference is totally arbitrary, however, since the [t] sound is actually in both words, regardless of the spelling. This tendency toward insertion is due to a similar articulatory problem as that given above regarding the insertion of [d] in the Spanish infinitives. In the case of *dance* and *pants*, it is difficult if not impossible to pronounce a

pure alveolar fricative /s/ after alveolar /n/ without inserting a [t], thus making it a double consonant [ts] instead of just [s].

For the above two groups of phonologically variant Future and Conditional forms, the present model accepts the above phonological explanation, especially since an alternate stem such as /*pondr*/ (for *poner*) would produce doubly marked forms in the model, as in **pondreré* instead of the correct form *pondré*. The small third group is the only one of the three groups that is difficult to account for by simple phonological patterns or orthographic conventions. The stem alternates used in the future and conditional of these verbs would indeed have to be specified in their semantic feature sets, and there would also have to be some mechanism for bypassing Blocks IA and IB of the model (in Figure 6.2). These special stem alternates would only be acquired (along with the rest of the associated feature sets) through extensive exposure to these forms. Upon hearing (and comprehending) *diré* in a future context, for example, the alternate stem /*dir*/ would then be stored along with the semantic and grammatical features associated with that particular context. Unless and until this ‘correct’ form is heard in the input, however, a speaker would be likely to produce the ‘incorrect’ form **deciré*.

It is interesting to note that in some regional varieties of Spanish in which /s/ is used instead of /θ/ (as in /*desir*/ as opposed to /*de θ ir*/, the /s/ is realized as an aspirated sound, as /h/, or even deleted in some phonetic contexts (Lipski 1985). In such varieties of the language, the /s/ in the regularized **deciré* might naturally be deleted, along with the unstressed /i/, yielding [*deré*]. If the front vowels /e/ and /i/ are indeed phonetically close as Lathrop (1980) says they were in Latin (73), speakers in the aspirated /s/

varieties of Spanish might well produce *diré* without having to process the form as having a special stem alternant.

This third group of phonologically variant Future and Conditional forms demonstrates the interesting observation that the most irregular verbs in any language are also among the most common verbs (Bybee 1988: 132), as are *decir* ‘say’ and *hacer* ‘do, make’ in Spanish. Irregularities in uncommon verbs would not be able to survive in the diachronic development of a language, since there would not be sufficient exposure to the irregular forms to consistently build the associated feature sets in the minds of the speakers.

In the list of verb stem sets presented in the Appendix of the present analysis, the above ‘irregular’ stems are included, nevertheless. Future research possibilities include empirical studies to see whether literacy has any effect on the phonemic perception of the ‘irregular’ Future/Conditional forms. There may be a separate listing for such ‘irregular’ forms in the mental lexicon of educated, literate speakers of the language, who may ‘see’ the form in their minds, but not for those who are not aware of the orthographic convention for writing these forms. It may be that careful pronunciation of these verb forms by non-literate Spanish speakers would produce regularized forms like **poneré*, **saliré*, **teneré*, **venire*. Such an empirical experiment might validate the argument that *pondré* may be indistinguishable in pronunciation from the ‘regular’ (but incorrect in written form) **poneré*, or at least that the mind of some speakers may not perceive the difference.

6.7. Phonologically variant Past Participle forms.

There is a significant number of Past Participle forms based on /to/ instead of the more common /do/ ending; therefore, the model includes a designation for ‘to-class’ verbs. Because of the application of the Elsewhere principle, and because there are far more Past Participles with /do/ than /to/, the rule for to-class verbs would block the application of the more general past participle rule that results in forms with /do/. After hearing enough examples of to-class Participles, the speaker would begin to identify (even if subconsciously, of course) these verbs as belonging to this separate class. Without sufficient input to make this analogical connection and thus form this more specific rule, the more general rule would probably apply and the speaker would generate the ‘incorrect’ regularized form with the /do/ ending, such as **escribido* instead of the correct form *escrito*, as many second language learners do. Evidence for this regularization was seen earlier in the example of the young child who insisted that “¡La palabra es *morido*! ¡No es *muerto*!” It should be noted that these to-class forms also have a different stem alternant, which is typically a shortened version of the infinitive stem, as in *escrito* above, where the syllable /βi/ has been deleted before adding the /to/ ending.

Figure 6.4 lists the common to-class verbs and their Past Participles. There are many other verbs that are compounds of these, and these compounds also belong to this class, such as *descubrir*, *devolver*, *suponer*, *recubrir*, *envolver*, *revolver*, *reescribir*, *posponer*, *deponer*, *reponer*.

FIGURE 6.4. *To*-class verbs and their Past Participles.

Infinitive	Past participle
abrir	abierto
cubrir	cubierto
escribir	escrito
morir	muerto
poner	puesto
resolver	resuelto
volver	vuelto

Since these Participles, including the many compounds derived from them, have such irregular stem alternates, these forms are probably stored and processed as completely separate lexical items, thus specified early in the model among the most specific of the rule-and-feature sets in the blocks. However, after sufficient exposure to these forms and familiarity with a significant number of them, there would be some sort of mental organization grouping them together and associating them analogically with each other in what is called here a *to*-class. The stem alternate would have to be specified in the feature set in the lexicon, of course, along with the specification of *to*-class. In addition, for some of these forms there would be semantic-phonological links with other forms that have the same initial strings of phonemes and the same basic semantic content, such as: *cubierto-cubre-cubrimos-cubrió*, or *muerto-muere-mueren-muera-mueras*.

In addition to this *to*-class group, there are two additional irregular Participles, plus all the compounds derived from them. As is typical of irregular forms, these Participles belong to two very common verbs: *decir* ‘say’ and *hacer* ‘do, make’. Their Past Participles are *dicho* and *hecho*, respectively. These suppletive Participles would have to be stored as separate lexical items, although it is conceivable that they might also

be stored along with the group of compounds formed from them, so that they would all generate participles with *dicho* and *hecho*.

6.8. Categorical/regular Preterite added to Basic Model.

The model shown in Figure 6.5 is an expansion of the basic model presented in Figure 5.2 in Chapter 5. In addition to producing all of the Present Indicative, Present Subjunctive, Imperfect Indicative, and Imperfect Subjunctive forms for categorical/regular verbs, as does the Basic Model, the expanded version of the model in Figure 6.5 also produces all the categorical/regular Preterite Indicative forms.

FIGURE 6.5. Categorical/regular Preterite Indicative added to Basic Model.

(Note that to save space in the written representation of the model, the notation [Pret] implies [Preterite *Indicative*].)

Block I

1) $\left[\begin{array}{l} + 1 \text{ Sg} \\ + \text{Pres} \\ + \text{Indic} \end{array} \right]$ $/X/ \rightarrow /X + o/$	2) $\left\{ \left[\begin{array}{l} + \text{Pres} \\ + \text{Subj} \\ + a\text{-class} \end{array} \right] \right\}$ $+ e/i\text{-class}$ $/X/ \rightarrow /X + e/$	3) $\left\{ \left[\begin{array}{l} + \text{Pres} \\ + \text{Subj} \\ + e/i\text{-class} \end{array} \right] \right\}$ $+ a\text{-class}$ $/X/ \rightarrow /X + a/$	4) $\left[\begin{array}{l} + \text{Imperf} \\ + \text{Indic} \\ + e/i\text{-class} \end{array} \right]$ $/X/ \rightarrow /X + ía/$
5) $\left[\begin{array}{l} + \left\{ \begin{array}{l} \text{Imperf Subj} \\ [3 \text{ Pl Pret}] \end{array} \right\} \\ + e/i\text{-class} \end{array} \right]$ $/X/ \rightarrow /X + ye/$	6) $\left[\begin{array}{l} + 1 \text{ Sg} \\ + \text{Pret} \\ + a\text{-class} \end{array} \right]$ $/X/ \rightarrow /X + é/$	7) $\left[\begin{array}{l} + 3 \text{ Sg} \\ + \text{Pret} \\ + a\text{-class} \end{array} \right]$ $/X/ \rightarrow /X + ó/$	8) $\left[\begin{array}{l} + 3 \text{ Sg} \\ + \text{Pret} \\ + e/i\text{-class} \end{array} \right]$ $/X/ \rightarrow /X + yo/$

9) $\left[\begin{array}{l} + \{1\text{Sg}/2\text{Sg}/1\text{Pl}/2\text{Pl}\} \\ + e/i\text{-class} \\ + \text{Pret} \end{array} \right]$	$\left[\begin{array}{l} + \{1\text{ Pl} / 2\text{ Pl}\} \\ + \text{Pres Indic} \\ + i\text{-class} \end{array} \right]$
/X/ \rightarrow /X + i/	

Block II

1) $\left[\begin{array}{l} + \text{Pret} \\ + \{2\text{ Sg}, 2\text{ Pl}\} \end{array} \right]$	2) $\left[\begin{array}{l} + \text{Pret} \\ + 3\text{ Pl} \end{array} \right]$	3) $\left[\begin{array}{l} + \text{Imperf} \\ + \text{Indic} \\ + a\text{-class} \end{array} \right]$	4) $\left[\begin{array}{l} + \text{Imperf} \\ + \text{Subj} \end{array} \right]$
/X/ \rightarrow /X + ste/	/X/ \rightarrow /X + ro/	/X/ \rightarrow /X + β a/	/X/ \rightarrow /X + ra/

Block III

1) $\left[\begin{array}{l} + 2\text{ Sg} \\ - \text{Pret} \end{array} \right]$	2) $\left[+ 2\text{ Pl} \right]$	3) $\left[+ 1\text{ Pl} \right]$	4) $\left[+ 3\text{ Pl} \right]$
/X/ \rightarrow /X + s/	/X/ \rightarrow /X + is/	/X/ \rightarrow /X + mos/	/X/ \rightarrow /X + n/

To test the model in Figure 6.5, we could start with a verb like *cantar* ‘sing’, which has only one stem in its lexical stem set: /kant/, which is specified in its feature set as [+ a-class]. Suppose we want the form for the semantic/grammatical value of 1 Plural Imperfect Indicative. The stem /kant/ enters Block I of the model. There is only one feature set in Block I that matches, the one for Rule 3, specified as [+ e/i-class Pres Subj] or [+ a-class], whose rule/representation is /X/ \rightarrow /X + a/, yielding /kanta/. This /kanta/ form then enters Block II as the new X variable, where Rule 3 is the only one that can apply, with a feature set of [+ a-class Imperf Indic]. The corresponding rule is /X/ \rightarrow /X

+ βa /, thus yielding /*kanta β a*/, which then enters Block III as the new X variable, with the corresponding rule /X/ \rightarrow /X + mos / producing the final form /*kantá β amos*/, spelled conventionally as *cantá β amos*.

Regarding syllable stress (which for Spanish is always vowel stress), the only time vowel stress is specified in the rules in the present model is when the given vowel departs from the normal stress pattern in Spanish, which is explained above in Section 6.6. For a form in which a final vowel is stressed, such as /i/ in *escribí*, the stress on the vowel must be specified with a written accent mark in the rule representation. If the stress on the given vowel does not contradict the normal stress pattern for other forms based on the given form with the stressed vowel, such as *escribimos* or *escribiste*, there is no need to specify a separate rule with no written accent, even though the written form would not require an accent mark. If the stressed vowel does contradict the normal stress pattern of other forms that might be generated from the stressed form, then another rule with the vowel unspecified for stress must be included. For example, the /é/ with specified stress in Rule 6 of Block I in Figure 6.5 would contradict the natural stress for a Present Subjunctive form such as *cante* /*kánte*/, with the stress on the first syllable. For this reason, there must be a separate rule, Rule 2 of Block I, for forms with natural a stress pattern that is inconsistent with the stress specified for the /é/ in Rule 6. In other words, the general rule here, as in all of the model, is that the more specific always takes precedence over the more general. There is no need to specify stress on a vowel in the model if the final form follows the normal Spanish stress pattern. If a given form departs from normal stress patterns, then a written accent mark is used, but that accent mark does

not adversely affect any forms that would otherwise normally carry stress on that vowel, such as in *escribimos* or *escribiste* as discussed above.

It should be mentioned here that there is one stress problem that the present model has not as yet been able to completely account for. In the 1 Plural Imperfect Indicative forms for *a*-class verbs, the antepenultimate syllable is stressed, requiring a written accent mark, since it departs from the normal syllabic stress patterns, as in *cantábamos*. The /ta/ syllable is stressed in order to maintain the stress on the same syllable throughout the verb paradigm, so that /áβa/ is always stressed on the first syllable, even when that syllable is not penultimate. Since the model itself does not account for this antepenultimate syllable stress (as in *cantábamos*), it is possible that analogy with other Imperfect forms with /áβa/ might account for this stress pattern.

Rule-and-feature set 9 in Figure 6.5 looks somewhat complicated, since it has several layers of square brackets and curly brackets. Upon careful examination, however, one can easily discern that the set must include: either Preterite *e/i*-class with a 1 Singular, 2 Singular, 1 Plural, or 2 Plural value, *or* Present Indicative *i*-class with a 1 Plural or 2 Plural value. After adding the appropriate strings in Block II and III for the specified feature sets, this rule would yield forms such as *comí*, *comiste*, *comimos*, *comisteis* (all *e*-class Preterite forms, with PN values of 1 Singular, 2 Singular, 1 Plural, 2 Plural, respectively), or *vivimos*, *vivisteis* (both Present Indicative forms, with PN values of 1 Plural and 2 Plural, respectively).

An important note must be added here. So far, in the previous models before the one in Figure 6.5, all of the features in the feature sets have been designated with [+] values. Positive values are used whenever possible in the model, always implying a

negative value for any complementary features such as opposing tenses, moods or PN values. One reason that the notation [+] has been used at all in the model, however, is to allow for the possibility that it may be necessary at times to specify a negative value for some features. Such is the case with Rule 1 of Block III in Figure 6.5, which specifies a negative value for the Preterite. For the rule stated as /X/ → /X + s /, the corresponding feature set is specified as [+ 2 Sg, – Pret], disallowing the application of the rule to any form with a semantic/grammatical value of [+ Preterite].

This particular rule, and the need to specify a negative value for the first time in the model, is particularly interesting. This anomaly in the model could account for the common, widespread use of ‘incorrect’ 2 Singular Preterite forms, even among educated speakers, such as **hablastes*, **escribistes*, **hicistes*, or **fuistes*, in which an extra /s/ is added to the form. This tendency to add /s/ to the 2 Singular Preterite forms can be explained by analogy, referring in this case to the tendency of items that are similar in meaning to become similar in form, sometimes referred to as ‘analogical leveling’ or ‘paradigm leveling’. All other 2 Singular forms in Spanish, in all other tenses and moods, end in /s/, as in *hablas*, *hables*, *hablabas*, *hablarás*, *hablarías*, *hablaras*. It makes sense, then, that given the same partial semantic/grammatical value of 2 Singular, the 2 Singular Preterite forms would have a natural tendency to add the same /s/ by analogy. This anomaly in the present model, requiring special negative notation to prevent the incorrect but well-attested forms, helps to corroborate the power of analogy in language change and the natural propensity to pattern regularization.

6.9. Exceptional/regular Preterite added to Basic Model.

In addition to the designation of *a*-class, *e*-class, and *i*-class verbs, a small group of exceptional/regular Preterite forms is now added to the model (Figure 6.6) and is designated in the present analysis as ‘allo-class’. The present study adopts the term ‘allo-class’ because of the meaning of *allo* from Greek *allos* ‘other’, as it is used in the terms ‘allophone’ and ‘allomorph’. The allo-class group includes the following verbs: *tener*, *andar*, *estar*, *haber*, *poder*, *poner*, *saber*, *venir*, *querer*, *hacer*. This group of verbs, which have special stem alternants for Preterite and Imperfect Subjunctive forms, are also somewhat irregular in their inflectional affixes. However, consistent with Skousen’s ‘exceptional/regular’ classification (1989: 8), there are regularities in their irregularities. They have similar patterns, and can thus be grouped together for organizational simplicity and economy. Just as illustrated above with the special Future and Conditional stem alternants specified in their respective feature sets, there are certain stem alternants for the Spanish Preterite Indicative and Imperfect Subjunctive forms that have to be specified in the feature sets as well. These stem alternates, listed in a Pre-Block such as that shown earlier in the idiosyncratic/suppletive model, feed into the first block of the model and are specified for the appropriate semantic-grammatical contexts.

Another exceptional/regular group is a subset of the specially designated allo-class, and is designated in the present analysis as ‘*j*-allo-class’. This second ‘allo’ group, which we will call the ‘*j*-allo-class’, is a subset of the allo-class and includes the following: *decir*, *traer*, *conducir*, *traducir*, *producir*, *reducir*. (The <*j*> of this *j*-allo-class is actually /x/ phonetically, but the graphemic <*j*> is used in the notation *j*-allo-class for ease of pronunciation in oral discussions or explanations of the model.) This latter group has

essentially the same patterns as the allo-class, except that the forms in this group all have the grapheme *j* (the phoneme /x/) as the final letter of the stem. When this /x/ appears before what would otherwise be a glide, as in the /ye/ in Rule 5 of Figure 6.6, the glide /y/ is deleted, resulting in a single vowel sound rather than a diphthong, such as in *trajeron*, *dijeron*, *condujeron*, as opposed to *tuvieron*, *anduvieron*, *hubieron*, for instance. Compounds based on these verbs exhibit the same stem alternation and thus would also be included in these groups.

It should be remembered that since the *j*-allo-class is a subset of the larger allo-class, any feature set that includes allo-class would also include *j*-allo-class, unless, of course, there is another rule in the same block that specifies *j*-allo-class. In the latter case, the more specific rule including a specification for *j*-allo-class would take precedence over a more general rule for allo-class verbs. Although any notation for allo-class includes the subset *j*-allo-class, the notation *j*-allo-class does not include allo-class. Another important note here is that although this group of allo-class verbs (along with its subset *j*-allo-class) is introduced in Figure 6.6, labeled ‘Exceptional/regular Preterite Indicative added to Basic Model’, the stems of these allo-class verbs also appear in Imperfect Subjunctive forms of their respective verbs. The model takes this into account, however, as in Rules 5 and 6 of Block I, where the resulting form can be either a Preterite form or an Imperfect Subjunctive form.

The model shown in Figure 6.6 is a further expansion of the basic model presented earlier in Figure 5.2. In addition to producing all of the Present Indicative, Present Subjunctive, Imperfect Indicative, and Imperfect Subjunctive forms for categorical/regular verbs, as does the Basic Model, the expanded version of the model

presented in Figure 6.6 also produces all of the Preterite Indicative forms, including the above mentioned exceptional/regular forms (designated here *allo-class* and *j-allo-class*). It is important to remember that the exceptional/regular group is so designated because the verbs in this group exhibit regular patterns and are thus not completely irregular or suppletive, although each of these verbs does have more than one member in its respective lexical stem set.

It is important to remember that these stem alternates feed into the first block and can be represented as rules in a Pre-Block, as in the *be/ser* model presented above. As in the idiosyncratic/suppletive *be/ser* model, the rules that represent stem alternates feeding into Block I of the model would be written with an equals sign, as */X/ = /the given stem/*, rather than */X/ → /the given form/*, since in the former case the X variable is the final form for the given block in the respective model, rather than the */X/* variable becoming a new form with additional material added onto it. For instance, for the *allo-class* verb *andar* ‘walk’, a Pre-Block would include the stem alternates */and/* and */anduβ/*, written as */X/ = /and/* and */X/ = /anduβ/*, respectively. The feature set for */anduβ/* would specify [+*allo-class*], but there would be no need to specify any grammatical features for the stem */and/*, since */and/* would apply everywhere except where [+*allo-class*] is specified in the feature set for a given rule. Just like everywhere else in the model, a general rule applies whenever it is not precluded by a more specific rule.

FIGURE 6.6. Exceptional/regular Preterite Indicative added to Basic Model. (Also included are the previously presented Present Indicative, Present Subjunctive, Imperfect Indicative, Imperfect Subjunctive, and categorical/regular Preterite.)

Block I

1) $\left\{ \begin{array}{l} + \{1 \text{ Sg Pres Indic}\} \\ + \{3 \text{ Sg Pret allo-class}\} \end{array} \right\}$	2) $\left\{ \begin{array}{l} [\text{Pres Subj } a\text{-class}] \\ + \\ [1 \text{ Sg Pret allo-class}] \\ [e/i\text{-class}] \end{array} \right\}$	3) $\left\{ \begin{array}{l} + \text{Pres} \\ + \text{Subj} \\ + e/i\text{-class} \\ + a\text{-class} \end{array} \right\}$	4) $\left\{ \begin{array}{l} + \text{Imperf} \\ + \text{Indic} \\ + e/i\text{-class} \end{array} \right\}$
$/X/ \rightarrow /X + o/$	$/X/ \rightarrow /X + e/$	$/X/ \rightarrow /X + a/$	$/X/ \rightarrow /X + ia/$

5) $\left\{ \begin{array}{l} + \{ \text{Imperf Subj} \} \\ + \{ \text{Pret 3Pl} \} \\ + \{ e/i\text{-class} \} \\ + \{ \text{allo-class} \} \end{array} \right\}$	6) $\left\{ \begin{array}{l} [1 \text{ Sg Pret } a\text{-class}] \\ + \\ [3 \text{ Pl } j\text{-allo-class Pret}] \\ [j\text{-allo-class Imperf Subj}] \end{array} \right\}$	7) $\left\{ \begin{array}{l} + \text{Pret} \\ + 3 \text{ Sg} \\ + a\text{-class} \end{array} \right\}$	8) $\left\{ \begin{array}{l} + \text{Pret} \\ + 3 \text{ Sg} \\ + e/i\text{-class} \end{array} \right\}$
$/X/ \rightarrow /X + ye/$	$/X/ \rightarrow /X + é/$	$/X/ \rightarrow /X + ó/$	$/X/ \rightarrow /X + yó/$

9) $\left\{ \begin{array}{l} + \{1\text{Sg}/2\text{Sg}/1\text{Pl}/2\text{Pl}\} \\ + \text{Pret} \\ + \{e/i\text{-class}/\text{allo-class}\} \\ + \\ + \{1 \text{ Pl} / 2 \text{ Pl}\} \\ + \text{Pres Indic} \\ + i\text{-class} \end{array} \right\}$
$/X/ \rightarrow /X + i/$

Block II

1) $\left[\begin{array}{l} + \text{ Pret} \\ + \{2 \text{ Sg}, 2 \text{ Pl}\} \end{array} \right]$	2) $\left[\begin{array}{l} + \text{ Pret} \\ + 3 \text{ Pl} \end{array} \right]$	3) $\left[\begin{array}{l} + \text{ Imperf} \\ + \text{ Indic} \\ + a\text{-class} \end{array} \right]$	4) $\left[\begin{array}{l} + \text{ Imperf} \\ + \text{ Subj} \end{array} \right]$
$/X/ \rightarrow /X + ste/$	$/X/ \rightarrow /X + ro/$	$/X/ \rightarrow /X + ba/$	$/X/ \rightarrow /X + ra/$

Block III

1) $\left[\begin{array}{l} + 2 \text{ Sg} \\ - \text{ Pret} \end{array} \right]$	2) $\left[+ 2 \text{ Pl} \right]$	3) $\left[+ 1 \text{ Pl} \right]$	4) $\left[+ 3 \text{ Pl} \right]$
$/X/ \rightarrow /X + s/$	$/X/ \rightarrow /X + is/$	$/X/ \rightarrow /X + mos/$	$/X/ \rightarrow /X + n/$

Now let's check the above model with a specific feature set. Let's say we want the form that corresponds to the semantic value represented by the verb *poner* 'put', with a TMA value of Preterite Indicative and a PN value of 3 Plural. As before, it should be noted that the infinitive form is given here as a point of reference only, simply as a symbol for all the abstract concepts associated with the verb *poner*. The mental lexicon lists the lexical stem set along with all the semantic features representing the abstract concept of *poner* and with the semantic/grammatical features specified for each stem alternant. The stem set includes */pon/*, */pong/*, */pus/*, *pondr/*, all specified as [*e/i*-class; allo-class], with the stem */pus/* also specified for [+{Pret Indic; + Imperf Subj}]. For the feature set that includes [+ 3 Pl] and [+ Pret Indic], Rule 5 in Block I applies, stated as $/X/ \rightarrow /X + ye/$. Since the associated feature set for Rule 5 is more specific than any other set in the block for the feature set [+ 3 Pl, + Pret Indic], the associated rule, Rule 5, precludes the application of any other rule that might otherwise apply. So if we apply the rule $/X/ \rightarrow /X + ye/$ to the stem */pus/*, the form exiting Block I is */pusye/*. We then proceed to Block II with */pusye/* as the new X variable, and we see that Rule 2 is $/X/ \rightarrow$

/X + ro/, specified for [+3 Pl, + Pret], thus yielding /pusyero/. The new X variable /pusyero/ then enters Block III, where we see that the only rule with a matching feature set (which is [+ 3 Pl]) is Rule 4, stated as /X/ → /X + n/. The final form exiting the last block is thus /pusyeron/, which is represented graphemically as *pusieron*.

6.10. Idiosyncratic/suppletive Preterite forms.

The few idiosyncratic/suppletive Preterite forms, such as the very common *ir* ‘go’ and *ser* ‘be’, must be specified as such in the lexicon in the feature sets of the *signifié* associated with the corresponding *significant*, just as are the stems for the *allo*-class and its subset *j-allo*-class verbs discussed in the previous section. The forms for the common verbs *ir* and *ser* happen to have identical Preterite forms: *fui*, *fuiste*, *fue*, *fuimos*, *fuisteis*, *fueron*. However, even in these idiosyncratic/ suppletive forms there are familiar patterns, such as the 2 Singular /ste/, 1 Plural /mos/, 2 Plural /is/, and 3 Plural /ron/. As a matter of fact, if /fu/ is designated as a stem alternant for the Preterite forms of *ser* and *ir*, then the only form that is not consistent with the *i-class* endings is the 3 Singular /fue/. This form, then, is the only one that would have to have a separate block (as being most specific) for a Preterite model. (This /fue/ has already been introduced in the model in Figure 5.8 in Chapter 5, where it resulted in all of the Imperfect Subjunctive forms for the verb *ser*.) The other forms with /fu/ would be consistent with regular *e/i-class* verbs for the Preterite, as long as the /fu/ stem alternant is specified as the /X/ variable entering the respective block.

FIGURE 6.7. Basic verb model with Preterite Indicative, but also including an idiosyncratic/suppletive form: *ser* ‘be’. (The model below includes Present Indicative, Present Subjunctive, Imperfect Indicative, Imperfect Subjunctive, categorical/regular Preterite, and exceptional/regular Preterite, as well as the corresponding forms for *ser*.)

Pre-Block

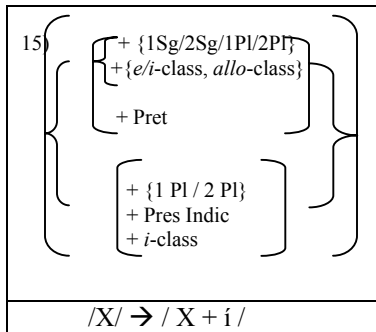
1) $\left[\begin{array}{l} + be/ser \\ + Pres \\ + Subj \end{array} \right]$	2) $\left[\begin{array}{l} + be/ser \\ + Pret \end{array} \right]$
/X/ = /se/	/X/ = /fu/

Block I

1) $\left[\begin{array}{l} + be/ser \\ + 1 Sg \\ + Pres \\ + Indic \end{array} \right]$	2) $\left[\begin{array}{l} + be/ser \\ + 2 Sg \\ + Pres \\ + Indic \end{array} \right]$	3) $\left[\begin{array}{l} + be/ser \\ + 3 Sg \\ + Pres \\ + Indic \end{array} \right]$	4) $\left[\begin{array}{l} + be/ser \\ + Pres \\ + Indic \end{array} \right]$	5) $\left[\begin{array}{l} + be/ser \\ + Imperf \\ + Indic \end{array} \right]$	6) $\left[\begin{array}{l} + be/ser \\ + Imperf \\ + Subj \\ [+ 3 Pret] \end{array} \right]$
/X/ = /soy/	/X/ = /ere/	/X/ = /es/	/X/ = /so/	/X/ = /era/	/X/ = /fue/

7) $\left\{ \begin{array}{l} + [1 Sg Pres Indic] \\ + [3 Sg Pret allo-class] \end{array} \right\}$	8) $\left\{ \begin{array}{l} + [Pres Subj a-class] \\ + [1 Sg Pret allo-class] \\ + [e/i-class] \end{array} \right\}$	9) $\left\{ \begin{array}{l} + \left[\begin{array}{l} Pres \\ Subj \\ e/i-class \end{array} \right] \\ + a-class \end{array} \right\}$	10) $\left[\begin{array}{l} + Imperf \\ + Indic \\ + e/i-class \end{array} \right]$
/X/ → /X + o/	/X/ → /X + e/	/X/ → /X + a/	/X/ → /X + ía/

11) $\left\{ \begin{array}{l} + \left[\begin{array}{l} Imperf Subj \\ 3 Pl Pret \end{array} \right] \\ + \left[\begin{array}{l} e/i-class \\ Pret allo-class \end{array} \right] \end{array} \right\}$	12) $\left\{ \begin{array}{l} + [Pret 1 Sg a-class] \\ + [3 Pl j-allo-class Pret] \\ + [j-allo-class Imperf Subj] \end{array} \right\}$	13) $\left[\begin{array}{l} + 3 Sg \\ + Pret \\ + a-class \end{array} \right]$	14) $\left[\begin{array}{l} + 3 Sg \\ + Pret \\ + e/i-class \end{array} \right]$
/X/ → /X + ye/	/X/ → /X + é /	/X/ → /X + ó/	/X/ → /X + yó /



Block II

1) $\left[\begin{array}{l} + \text{Pret} \\ + \{2 \text{ Sg}, 2 \text{ Pl}\} \end{array} \right]$	2) $\left[\begin{array}{l} + 3 \text{ Pl} \\ + \text{Pret} \end{array} \right]$	3) $\left[\begin{array}{l} + \text{Imperf} \\ + \text{Indic} \\ + a\text{-class} \end{array} \right]$	4) $\left[\begin{array}{l} + \text{Imperf} \\ + \text{Subj} \end{array} \right]$
/X/ → /X + ste /	/X/ → /X + ro /	/X/ → /X + βa /	/X/ → /X + ra /

Block III

1) $\left[\begin{array}{l} + 2 \text{ Sg} \\ - \text{Pret} \end{array} \right]$	2) $\left[+ 2 \text{ Pl} \right]$	3) $\left[+ 1 \text{ Pl} \right]$	4) $\left[+ 3 \text{ Pl} \right]$
/X/ → /X + s /	/X/ → /X + is /	/X/ → /X + mos /	/X/ → /X + n /

Another common verb, *ver* ‘see’, is presented in textbooks and grammar manuals as having irregular Preterite forms: *vi, viste, vio, vimos, visteis, vieron*. If we allow a single consonant to be a stem, the Preterite forms of this verb are not really irregular at all, however. The only irregularity is the lack of the written accent marks normally used in the 1 Singular and 3 Singular Preterite forms, and this spelling convention is due to the fact that the written accents are not necessary in one syllable words like *vi* and *vio*, just as they are equally not needed (or used) in *fui* and *fue*. If we accept the single consonant /β/ as the stem (since the present analysis, as explained earlier, uses /β/ to represent the

phoneme for the graphemes <*b*> and <*v*>), then the endings are completely regular and are consistent with other categorical/regular *e*-class verbs.

The common verb *dar* ‘give’ is also presented as irregular in the Preterite: *di, diste, dio, dimos, disteis, dieron*. However, the only thing irregular about this paradigm is that the verb switches classes in its Preterite Indicative and Imperfect Subjunctive forms. In all other forms, this verb functions as an *a*-class verb, but in the Preterite Indicative and Imperfect Subjunctive, the endings are exactly the same as any categorical/regular *e*-class or *i*-class verb. In fact, the Preterite paradigm is the same as for that of *ver* in the preceding paragraph, except for the single consonant /*d*/ as the stem instead of the single consonant /*β*/. Since the verb *dar*, though, typically functions as an *a*-class verb in other contexts, this class switch for the Preterite Indicative and Imperfect Subjunctive has to be specified in its associated feature sets in the lexicon.

The rule of referral mentioned earlier in Chapter 5 might serve useful for the small allo-class of Preterite forms (including the *j*-allo subclass), along with *ver* and *dar*, since for this group of verbs the same stem alternant is used for the Preterite Indicative and for the Imperfect Subjunctive. However, since this rule of referral is only necessary for a small group of verbs, it may be more economical for the system as a whole to specify these special cases in the basic model itself, as we have done above in adding the Preterite to the basic model. Both of the above mentioned rules of referral – this one and the one for forming all of the Present Subjunctive based on the 1 Singular Present Indicative form – are still probably useful for pedagogical purposes, but for the simplest, most economical organization of the system as a whole, the model proposed in the present study may work better than the rules of referral.

6.11. Compound forms.

Thus far this analysis has covered all of the simple tenses in Spanish in both the indicative and subjunctive moods, along with the three nonfinite, noninflected verb forms: the Infinitive, the Past Participle, and the Present Participle. The present analysis will also consider the compound verb tenses, although in a strict sense the compound forms might be considered in the syntactic, rather than morphological, realm. While the nonfinite forms themselves (Infinitive, Past Participle, and Present Participle) are more toward the ‘lexical expression’ end of Bybee’s continuum (1985: 12), as discussed in Section 6.5 above, the compound forms are more toward the ‘syntactic expression’ end of the continuum. Compound forms include an inflected verb form followed by one or two nonfinite forms, such as *ha trabajado* ‘has worked’ or *ha estado trabajando* ‘has been working’. The present model includes the following:

- 1) The compounds based on *haber* + Past Participle: Present Perfect Indicative (as in *he cantado*), Past Perfect (Pluperfect) Indicative (as in *había cantado*), Future Perfect Indicative (as in *habrá cantado*), Conditional Perfect (as in *habría cantado*), Present Perfect Subjunctive (as in *haya cantado*), Past Perfect (Pluperfect) Subjunctive (as in *hubiera cantado*), and perhaps the not-so-common Preterite Perfect Indicative (as in *hubo cantado*) as well;
- 2) The compounds based on *estar* + present participle (or sometimes other verbs instead of *estar*, such as *seguir*, *ir*, *andar*, *continuar*): Present Progressive Indicative (as in *está cantando*), Present Progressive Subjunctive (as in *esté cantando*), Past Progressive Indicative (as in *estaba cantando*), and Past Progressive Subjunctive (as in *estuviera cantando*);

- 3) The multiple compounds that are a combination of the above, such as the Present Perfect Progressive Indicative (as in *ha estado cantando*), Past Perfect Progressive Indicative (as in *había estado cantando*), Future Perfect Progressive (as in *habrá estado cantando*), and Conditional Perfect Progressive (as in *habría estado cantando*).

For a compound verb model, we have to think three-dimensionally and imagine a model that includes nonfinite forms (Infinitives and Participles) (Figure 6.2) superimposed on the Basic Model for finite forms (Figure 5.2 or the expanded Figure 6.7). There is an /X/ variable, the so-called content-bearing part of the main verb, feeding into the nonfinite form dimension, and another semi-variable /Y/, which is the stem of either *haber* or *estar* or one of the *estar* substitutes, feeding into the finite form dimension. The /Y/ stem picks up the finite part of the compound, while the /X/ stem picks up the nonfinite part of the compound. The output, meaning the final form, would then be a concatenation of the outputs of the two different dimensions, with the order specified at the outset by the general rule shown in Figure 6.8.

FIGURE 6.8. Rule for the formation and ordering of a two-part compound form.

$/X/ \rightarrow /Y \text{ finite}/ + /X \text{ nonfinite}/,$

where Y = the specified stem alternate of *haber* or *estar* (or *estar* substitute);

X = the content-bearing stem;

finite = the finite inflected output of the finite model dimension;

nonfinite = the nonfinite output (i.e., Present or Past Participle) of the nonfinite model dimension.

The rule in Figure 6.8 means that a given verbal lexeme /X/ is transformed into a compound form by concatenating in this specific sequence: a finite form of a /Y/ semi-variable (*haber*, *estar*, or an *estar* substitute) with a nonfinite form of the given /X/ variable, which is the ‘content-bearing’ part of the given verb. The feature sets associated with this rule, and with the more specific rules within the blocks themselves, would specify the abstract *signifié* associated with the given compound tense, such as ‘ongoing action in the past’ (labeled Past Progressive) or ‘completed action in the past having bearing on the present’ (labeled Present Perfect).

In the case of multiple compounds, such as the Past Perfect Progressive Indicative form *había estado trabajando* (*had been working*), the entire *signifié* would be specified in the feature set (such as the notion of an *ongoing action previous to a given past reference point, from an objective viewpoint*), and there would be multiple dimensions to the model with two nonfinite levels, as in the rule in Figure 6.9.

FIGURE 6.9. Rule for the formation and ordering of a three-part compound form.

/X/ → /Y finite/ + /Z nonfinite/ + /X nonfinite/,

where Z = the specified stem alternate of the second auxiliary.

Since a three-part compound form in Spanish will necessarily have a form of *haber* in the first (finite) position and a form of *estar* in the second (nonfinite) position, such as *ha estado cantando* (Present Perfect Progressive Indicative), we could rewrite Rule 6.8 to accommodate a three-part compound form, with curly brackets indicating a mutually exclusive choice between two outcomes for the rule. Note that the first option

in this rule (shown in Figure 6.10) would allow for either *haber* or *estar* (or a substitute for *estar*) in the finite position, yielding a form such as *hemos cantado* (Present Perfect Indicative) or *estamos cantando* (Present Progressive Indicative). The second option, however, specifies the choice and order for the *haber* and *estar* forms.

FIGURE 6.10. Rule for the formation and ordering of a two- or three-part compound form.

$$/X/ \rightarrow \begin{cases} /Y \text{ finite}/ + /X \text{ nonfinite}/ \\ /haber \text{ finite}/ + /estar \text{ (or substitute) Past Part}/ + /X \text{ Pres Part}/, \end{cases}$$

where Y = the specified stem alternate of *haber* or *estar* (or *estar* substitute);

X = the content-bearing stem;

finite = the finite inflected output of the finite model dimension;

nonfinite = the nonfinite output (i.e., Present or Past Participle) of the nonfinite model dimension.

It should be noted that in the basic model for nonfinite forms (Figure 6.2), the abbreviations ‘Pres Part’ and ‘Past Part’ are used for the Present and Past Participle forms, respectively, but these abbreviations are really just shorthand for the entire concept these grammatical designations represent, as are all of the labels used in the model. The abstract concept, the *signifié*, associated with the Present Participle would be something like ‘ongoing action or state’ (with no implication of time as there is in the misnomer ‘Present Participle’), while the *signifié* of the Past Participle would be something like ‘completed action or state’.

As discussed in Chapter 4, the use of grammatical terms (or abbreviations thereof) throughout the model is really just a shortcut, an abbreviated way of representing part of the whole abstract mental concept (the *signifié*) associated with the whole inflected physical word form (the *signifiant*). As noted before, these grammatical designations are in a sense not significantly different from the so-called content-bearing part of the *signifié*, but rather are part of the overall *signifié* of a given word. They are represented as constants in the model because they have analogical links to other forms with the same or similar phonological strings but with different ‘content-bearing’ stems, which are the initial variables in the model. Some of these constants in the model may be identical to the phonological strings traditionally identified as morphemes, but they are not necessarily so.

6.12. Conclusion.

The preceding chapter, Chapter 5, introduced the proposed model, including analyses of the model’s treatment of categorical/regular forms, phonologically variant forms, exceptional/regular forms, and idiosyncratic/suppletive forms for the Present Indicative, Present Subjunctive, Imperfect Indicative, and Imperfect Subjunctive. This chapter, Chapter 6, has expanded upon the Basic Model and discussed lexical gaps, defective verbs, nonfinite forms, Future and Conditional forms, irregular Past Participle, Future, Conditional, regular and irregular Preterite, and compound forms. The model presented in this chapter and the preceding chapter, along with the Appendix listing of the lexical stem sets for some common Spanish verbs, as far as we know, accounts for all forms of all Spanish verbs. By seeking common denominators and analogical

organizational patterns, the model has condensed the inflectional morphology of Spanish verbs into a set of representational rules that describe as concisely as possible the patterns of organization of those verbs in the mental lexicon.

This proposed model corroborates Anderson's Extended Word-and-Paradigm model of morphology, as well as his a-morphous morphological theory, although the present model departs from Anderson in regarding the Word-Formation Rules as descriptive patterns in the lexicon rather than as true generative rules that occur independent from the lexicon. Thus the proposed model is morpholexical, and like Bybee's (1985, 1988) model, this model regards rules and representations as essentially the same thing, as occurring on a continuum rather than as independent entities. The present analysis supports Bybee's analogical model of morphology in that it sees Word-Formation Rules as descriptive patterns that show analogical connections in the lexicon among forms that are similar in semantic and/or phonological features. The proposed model is also consistent with Marslen-Wilson's (1987, 1989, 1999) cohort model of lexical processing, since a cohort theory of word recognition or word formation may facilitate faster mapping of the appropriate phonological forms onto their corresponding feature sets.

PART III

CONCLUSIONS

AND

PEDAGOGICAL IMPLICATIONS

CHAPTER 7

ORGANIZATION AND LABELING OF VERB PARADIGMS

7.0. This chapter presents an overview of the traditional ways of organizing and labeling verb paradigms for analytical and pedagogical purposes, comparing Spanish with several other languages in this regard. This overview serves as background and justification for the grammatical notations used in the model of Spanish inflectional verbal morphology proposed in the present analysis. This discussion also points out the variability, arbitrariness, and language specificity of such paradigms and their labels and organization. Bybee's (1985) observations and comments about the psychological validity of paradigms are considered, contrasting with the generativist view of abstract underlying representations for verb paradigms. Constants and variables in different kinds of verb paradigms are also considered, including Bybee's analogical model of morphology, thus establishing a basis for the constants and variables used in the proposed model.

7.1. How many sets of forms?

When linguists, grammarians, and textbook writers present or discuss the many verb forms of a language, they use specific terminology to refer to the various verb forms and to the paradigms, or sets of inflected forms. Although the terminology may differ somewhat from one source to another, the manner of grouping the forms into sets or paradigms is fairly consistent for Spanish. A brief overview of Spanish textbooks and grammar manuals, whether the labels and explanations are presented in Spanish or in

English, shows a high level of consistency in the grouping of the verb forms into paradigms. Textbooks for beginning and intermediate Spanish, as well as more advanced grammar review texts, typically present at least thirteen paradigms, with each set of inflected forms representing a different tense, mood, aspect (TMA) or some combination of these three. The thirteen basic tense/mood/aspect-based paradigms presented in most texts and grammar reviews are: Present Indicative, Imperfect Indicative, Preterite Indicative, Future, Conditional, Present Subjunctive, Imperfect Subjunctive, Present Perfect Indicative, Past Perfect Indicative, Future Perfect, Conditional Perfect, Present Perfect Subjunctive, and Past Perfect Subjunctive. Each of these paradigms comprises six forms, representing person and number (PN): first, second and third person singular (1 Sg, 2 Sg, 3 Sg), and first, second, and third person plural (1 Pl, 2 Pl, 3 Pl) of the given verb.

In addition to the above thirteen sets of forms, some texts include the Antepreterite (such as *hube hablado*/I had spoken), also known as Preterite Perfect, which translates essentially like the Past Perfect. Many modern texts do not include the Antepreterite forms, however, because “it is used only rarely, primarily after conjunctions of time” as in “Así que hubieron terminado la comida... (As soon as they had finished the meal...)” (Da Silva and Lovett 1965: 28). The obsolete Future Subjunctive forms, both the simple Future Subjunctive (*hablare*) and compound Future Perfect subjunctive (*hubiere hablado*) add two more possible sets if they are included, as does *Langenscheidt’s Pocket Bescherelle Spanish Verbs* (1999). The alternate Imperfect Subjunctive forms with /se/, as in *hablase* as opposed to /ra/ as in the more common *hablara*, and the Past Perfect Subjunctive *hubiese hablado*, presented in some but not all

textbooks, add two more sets of forms, although the /se/ forms have essentially the same meaning as the /ra forms. Dozier and Iguina (1999) note that “In most of Latin America, the -ra forms predominate” (164), and da Silva and Lovett (1965: 103-104) point out that the /se/ forms are more limited in their usage, in addition to regional preferences for one or the other. From a functional standpoint, the /ra/ and /se/ forms differ more in style than in meaning. For this reason, the /se/ Imperfect Subjunctive forms, rather than presented as a separate category with a separate label, are usually just listed along with the /ra/ forms (as in Kendris 1990, Lloréns-Camp 1998, Seco 1996, Zatarain et al 1998, and others). The Imperative forms may also be considered a separate paradigm of sorts, although compared to the other sets, the Imperative forms are an incomplete paradigm due to the obvious lack of a first person singular form.

In addition, if we add the Progressive *estar* forms (as in *estoy hablando* ‘I am speaking’) to the total number of sets of forms inflected for person and number, the total number of inflected paradigms, not counting Imperative forms, would double, since all of the above mentioned sets of forms potentially have a Progressive counterpart. (It is even conceivable that the Imperatives could have a Progressive counterpart, for example: *Esté trabajando*, as in ‘Be working’ [while I am out of the room].) However, from a strictly morphological standpoint, we might not want to include any such periphrastic forms at all, since they might be considered more of a syntactic issue than a morphological one. Furthermore, Progressive forms may use verbs other than *estar* in the finite position of the compound, such as in: *va creciendo*, *anda buscando*, *sigue hablando*. Because of this, *estar*, as part of the Progressive paradigms, may not be a true auxiliary verb as is *haber* in Perfective forms like *he hablado*, *había conocido*, *habré escrito*.

It is important to remember that the above discussed way of arranging, organizing, and labeling verb paradigms is not universal, but rather is language specific to a great degree. For example, it is interesting to note that Beyer's *501 English Verbs* (1998) lists English verbs in a different way than what seems to be typical for the Romance languages. Although the forms for the six grammatical 'persons' mentioned above (1 Sg, 2 Sg, 3 Sg, 1 Pl, 2 Pl, 3 Pl) are all given in Beyer's verb charts, they are not listed in columns as in Spanish, nor are they necessarily in the order that is conventional for Spanish. All of the personal pronouns that correspond to a given English verb form are listed together with that form for the given tense and mood, as in Figure 2.1, which shows all of the active voice forms for the Indicative mood. Here, the Passive forms for each tense of each verb are also given, as are the two Active and Passive Imperative forms and the Active and Passive forms for the Present, Past, and Future tenses of the Subjunctive mood. Although there is in most cases a literal Spanish equivalent of these English Passive voice forms, Spanish verb lists do not include Passive forms. Passive forms in Spanish are not synthetic, one-word forms as were their Latin predecessors. Thus, the inclusion or non-inclusion of Passive as well as Active voice in verb conjugation charts is very language specific.

FIGURE 7.1. Beyer's conjugation of English verbs (1998: 24, 271).

[The use of boldface type is Beyer's.]

INDICATIVE MOOD		
TENSE	<i>be</i>	<i>walk</i>
Present	I am we, you, they are he, she, it is	I, we, you, they walk he, she, it walks
Present progressive	I am being we, you, they are being he, she, it is being	I am walking we, you, they are walking he, she, it is walking
Present intensive		I, we, you, they do walk he, she, it does walk
Future	I, he, she, it, we, you they will be	I, he, she, it, we, you, they will walk
Past	I, he, she, it was we, you, they were	I, he, she, it, we, you, they walked
Past progressive	I, he, she, it was being we, you, they were being	I, he, she, it was walking we, you, they were walking
Past intensive		I, he, she, it, we, you, they did walk
Present perfect	I, we, you, they have been He, she, it has been	I, we, you, they have walked he, she, it has walked
Past perfect	I, he, she, it we, you they had been	I, he, she, it, we, you, they had walked
Future perfect	I, he, she, it, we, you, they will have been	I, he, she, it, we, you, they will have walked

Aspectual differences in verbs may also be included in verb conjugation charts, as in Spanish Preterite Indicative versus Imperfect Indicative. Russian also distinguishes between Perfective and Imperfective forms of verbs, and, in fact, the contrast between the two aspects in Russian in some ways parallels the distinction between Imperfect and Preterite in Spanish. Although the rationale behind the different usage of the Perfective versus Imperfective aspect of verbs in Russian is similar in many ways to that of Spanish, the Perfective vs. Imperfective distinction in Russian has broader usages and is obligatory

in all tenses of the given verb except the Present, and is used even in the Infinitive form, Participles, and Imperatives. In other words, there are two completely different sets of paradigms for all Russian verbs for all tenses, one for Perfective and one for Imperfective.

There is also a vast difference in the complexity of the respective forms in the Russian and Spanish. Imperfect Indicative forms in Spanish are almost entirely regular and are formed by the simple addition of the suffix /aβa/ for *a*-class verbs and /ía/ for *e/i*-class verbs. In Russian, however, Perfective forms may be derived from Imperfective forms, or Imperfective forms may be derived from the Perfective. The aspectual distinction is made by adding a prefix, dropping a syllable, changing the theme vowel in the Infinitive (thus changing the verb class), changing the verbal suffix, or by suppletion, using a completely different stem (Fayer 1959: 156). For example, the verb *читать* /čitat'/ 'read' is Imperfective and has forms in all tenses, while the verb *прочитать* /pročitat'/ is Perfective 'read'. The Perfective form likewise has forms in all tenses, except that there is no true 'present' sense for Perfective verbs, and the Present forms for Perfective verbs have a future meaning. In the case of *читать* /čitat'/ and *прочитать* /pročitat'/, the latter of these forms (the Perfective) is identical to the Imperfective form, except for the addition of the prefix /pro/. In other cases, there may be a change in vowel class (hence, conjugation class), as in Imperfective *объяснять* /obyasnyát'/ 'explain' vs. Perfective *объяснить* /obyasnyít'/ 'explain'. The aspectual distinction yields two parallel sets of forms for most verbs and might therefore be considered in the realm of derivational rather than inflectional morphology.

As discussed in Chapter 3 of the present analysis, Bybee (1985: 29-37) illustrates from cross-linguistic evidence that derivational and inflectional morphological processes are not clearly delineated and separate, but rather are part of a continuum with lexical expression on one end and syntactic expression on the other. In other words, what is expressed by a completely separate word in one language may be expressed by a periphrastic or syntactic expression in another language, or by some morphological process somewhere in between on the continuum. In fact, in Bybee's cross-linguistic study of fifty languages, she found that aspect markers are closer to the stem than markers for other inflectional categories in most cases, and that aspect is more likely to condition changes in the verb stem than any other inflectional category. Anderson (1995) distinguishes between derivational and inflectional morphology by stating that "where two or more suffixes are involved, inflectional ones come after derivational ones" (77). On Bybee's continuum, then, aspect is more likely to be over towards the lexical end of the continuum, thus tending more toward derivation than inflection. And in the case of aspectual differences in Russian, where aspect is represented by two different verb forms in every tense, this tendency toward the lexical end of the continuum is especially clear.

In other ways, however, Russian verbal morphology is less complex than that of Spanish. After choosing the appropriate Perfective or Imperfective stem for a Russian verb, there only remains a choice among three tenses: Present, Past, and Future, along with an additional set of forms for Conditional and one for Imperative mood. The past form does not inflect for first, second, and third person, but rather for gender, with separate (but very simple) forms for masculine singular, feminine singular, and neuter singular, and one plural form for all three genders. The Future form for imperfective

verbs is a compound using a finite form of the verb *быть* /bit'/ 'to be' plus an infinitive, and the Future for Perfective verbs is what would otherwise be the Present-tense conjugated form, except that the Perfective aspect is not compatible with the true Present tense. The Conditional is simply the Past form followed by the particle *бы* /bi/. The Imperative forms are also fairly simple and regular. The complexity of the Russian verbal system lies primarily in the variability of stems based on aspectual distinction and derivational variants of verbs, such as distinctions based on direction and extent of movement. The complexity of the Spanish verbal system is in the stem allomorphy throughout the thirteen or more tense/mood/aspect paradigms. Once the stem allomorphy and derivational processes are accounted for, however, both Spanish and Russian are relatively regular in their inflectional suffixes.

English, Latin, and German texts list the 'principle parts' of a given verb, as in *is, being, was/were, been* and *walks, walking, walked, walked* for the English verbs given in Figure 7.1. The principle parts for English are the 3 Singular Present form, the Present Participle, the Past form, and the Past participle. For English these three forms, along with the Infinitive, represent all of the inflected forms of a given verb, since Modern English has lost most of the verbal inflections that older forms of English had. For German (Jannach 1980), the principle parts are the Infinitive, the Past form, the Past Participle, and the Present form, although, unlike English, German verbs are further inflected for person and number in the Past tense. For example, the Past form listed with the principle parts for the verb *sein* (*be*) is *war*, which is the 3 Singular form, but there are additional inflected Past forms: *warst* (2 Sg Familiar), *waren* (1 Pl, 3 Pl, 2 Pl Formal), and *wart* (2 Pl). Latin texts such as Moreland and Fleischer (1977) list the principle parts

of a given verb as the 1 Singular Present Indicative form, the Infinitive, the 1 Singular Perfect form (which is a simple, rather than compound, form in Latin), and the Past Participle, for example: optō, optāre, optāvī, optātus. The custom of listing ‘principal parts’ for German and Latin is apparently due to the fact that these particular forms (the ‘principal parts’) serve as a basis for the generation of other related forms, since they show stem allomorphy, conjugation class, and/or theme vowel. It is also interesting and relevant to note here that Latin paradigms include Passive voice forms, since Latin, unlike English and Spanish, has simple synthetic (rather than compound) forms for the Passive voice.

All of the above contrasts among languages illustrate the variability in the complexity of verbal morphology in human languages and the variability in the way linguists and grammarians organize and label verb paradigms. This variability also points out the arbitrary nature and the language specificity of the way verb paradigms are organized and labeled.

7.2. Internal organization of verb paradigms.

Although the way of grouping and labeling verb forms may vary considerably among different languages and among different sources, the paradigms in Spanish verbal morphology do not vary significantly in their internal grouping, meaning the way they are grouped within the tense-based sets. The only variation in this regard seems to be whether the individual paradigms are grouped into two columns, one singular and one plural, as in *501 Spanish Verbs* (Kendris 1990) and *Encuentros* (Spinelli and Rosso-O’Laughlin 1997), as in Figure 7.2, or, alternatively, in a single column, as in *Dos*

Mundos (Terrell et al 1998), *Mundo 21* (Samaniego 2001), and *Larousse de la conjugación* (Zatarain et al 1998), illustrated in Figure 7.3.

FIGURE 7.2. Two-column six-member PN-based paradigm for Spanish verb *hablar*.

	SINGULAR	PLURAL
1 ST PERSON	<i>hablo</i>	<i>hablamos</i>
2 ND PERSON	<i>hablas</i>	<i>habláis</i>
3 RD PERSON	<i>habla</i>	<i>hablan</i>

FIGURE 7.3. One-column six-member PN-based paradigm for Spanish verb *hablar*.

1 ST PERSON SINGULAR	<i>hablo</i>
2 ND PERSON SINGULAR	<i>hablas</i>
3 RD PERSON SINGULAR	<i>habla</i>
1 ST PERSON PLURAL	<i>hablamos</i>
2 ND PERSON PLURAL	<i>habláis</i>
3 RD PERSON PLURAL	<i>hablan</i>

This difference in internal grouping of the paradigms may be a minor difference, although some might argue that there are pedagogical advantages to one or the other of these two ways of grouping the individual paradigms. The significant thing, however, is that the arrangement of the inflected verb forms into these six-member groups seems to prevail in virtually every Spanish textbook examined for this investigation, as well as in the French textbooks examined. Some beginning level texts, such as *¿Qué tal?* (Dorwick and Girones 1995) or the French text *Qu'est-ce qu'on dit?* (Manley et al 1994) do present isolated members of the paradigm at first, such as just *soy, eres, es*, the Spanish singular forms for *ser* 'be' or French *Je suis/je ne suis pas* 'I am/I am not' in the Chapitre

Préliminaire of the above French text. However, the grammar review sections of these and other such texts present all of the paradigms organized in the traditional way as above, whether in two columns or one.

These sets of six forms (1, 2, 3 Singular and Plural) each are then typically grouped according to tense, aspect, and mood, roughly based on the notions of present, past, or future time frames, and on indicative mood vs. subjunctive mood vs. imperative mood, and sometimes conditional mood (Lloréns-Camp 1998). The present investigation found no textbook grammar that organizes paradigms based on person, for example, rather than based on tense and mood. An example of a person-based organization is illustrated in Figure 7.4.

FIGURE 7.4. Hypothetical third-person-based paradigm.

THIRD PERSON FORMS FOR *HABLAR* (TO SPEAK):

	SINGULAR	PLURAL
PRESENT INDICATIVE	<i>habla</i>	<i>hablan</i>
IMPERFECT INDICATIVE	<i>hablaba</i>	<i>hablaban</i>
PRETERITE	<i>habló</i>	<i>hablaron</i>
PRESENT PERFECT INDICATIVE	<i>ha hablado</i>	<i>han hablado</i>
PLUPERFECT INDICATIVE	<i>había hablado</i>	<i>habían hablado</i>
ANTEPRETERITE	<i>hubo hablado</i>	<i>hubieron hablado</i>
SIMPLE FUTURE	<i>hablará</i>	<i>hablarán</i>
FUTURE PERFECT	<i>habrá hablado</i>	<i>habrán hablado</i>
PRESENT CONDITIONAL	<i>hablaría</i>	<i>hablarían</i>
CONDITIONAL PERFECT	<i>habría hablado</i>	<i>habrían hablado</i>
PRESENT SUBJUNCTIVE	<i>hable</i>	<i>hablen</i>
IMPERFECT SUBJUNCTIVE	<i>hablara</i>	<i>hablaran</i>

PRESENT PERFECT SUBJUNCTIVE	<i>haya hablado</i>	<i>hayan hablado</i>
PLUPERFECT SUBJUNCTIVE	<i>hubiera hablado</i>	<i>habían hablado</i>
IMPERATIVE	<i>hable</i>	<i>hablen</i>

7.3. Person/number-based sets.

Since there is evidence that the 3 Singular form is typically the first form acquired by a child in a Spanish-speaking environment and might be considered the basic or default form (Bybee 1985: 56, 59-60), it would not be unreasonable to consider the above organization of verb paradigms based on person rather than tense. This grouping of forms based on person instead of tense and mood is, of course, considerably larger than the six-member sets conventionally used. We could even group the forms according to number, but those sets would be even larger, since in Spanish there are only two possible groups based on number, that is, singular and plural. The point here is that the conventional arrangement of verb paradigms is somewhat arbitrary, and there is no reason to assume that it is necessarily the best from a pedagogical standpoint.

If the 3 Singular Present Indicative is the default form and the first form acquired, as Bybee contends, then another possible grouping (albeit a rather large potential group) would be simply 3 Singular Present Indicative for all possible lexemes. Although this is perhaps an unreasonably large group for the purposes of listing in a textbook, it is not unreasonable to conceive of a mental paradigm organized in this way, if paradigms do indeed have psychological reality. With this type of organization, a child or adult learner would simply add more lexemes to the paradigm through experience, but might possibly

group all known 3 Singular Present Indicative forms together for organization and retrieval purposes alike.

Dos Mundos (Terrell et al 1998) presents several 3 Singular Present Indicative forms in the Paso A (preliminary lesson): *es, habla, lleva, saluda, tiene, está, se llama*. *Dos Mundos* presents grammar explanations (in English) in the ‘blue page’ sections that follow each lesson and are suggested by the authors for use as reference and home self-study. The Paso A blue pages, following the regular white page section of the lesson, present a few more forms of the traditional paradigm for *ser* ‘be’: *soy, es, somos, son*) and for *llevar* ‘wear, carry’: *llevo, lleva, llevamos, llevan*, but they omit the second person forms of both of these verbs (*ser* and *llevar*). The grammar section of Paso B of *Dos Mundos* then presents the full traditional Present Indicative paradigm for *ser*, and the Paso 3 blue pages present the full traditional Present Indicative paradigm for *tener*, as well as the full traditional paradigm for so-called ‘regular -ar verbs’, including *hablar*.

Manley’s beginning French text (1994) also employs this sort of initial grouping according to one person only in his Chapitre Préliminaire, but with the 1 Singular Present Indicative forms rather than the 3 Singular Present Indicative. The preliminary chapter of *Qu’est-ce qu’on dit?* (Manley et al, 1994) presents *je suis, j’habite, je parle, je pense, je travaille*, all 1 Singular Present Indicative forms of various lexemes. The text presents other forms a little at a time in functional contexts in successive lessons, even presenting a 1 Singular Conditional form in Chapitre 2: *Je voudrais* ‘I would like’. In Chapitre 2 the text presents the full traditional Present Indicative paradigm for ‘regular verbs having infinitives ending in -er’. In virtually all modern language textbooks since the advent of Communicative Language Teaching in the 1980’s, many advanced forms of verbs, such

as French *voudrais* (Conditional) or Spanish *quisiera* (Imperfect Subjunctive), are used from the beginning in receptive (as opposed to productive) activities, such as readings, dialogues, instructions, videos, and various forms of realia.

Many beginning level textbooks and courses, including the two above mentioned, use Imperative forms very early, nearly always right from the beginning. They may be actively presented in the end-of-chapter vocabulary section, as in *Qu'est-ce qu'on dit?*, they may be used in written instructions, and/or they may be used by the instructor in class in TPR-type activities, as suggested by Terrell et al in *Dos Mundos* (1998) and other texts based on Terrell's Natural Approach (Krashen and Terrell 1983). TPR is 'Total Physical Response', which is Asher's technique of teaching through the use of commands and physical activity (Asher 1982). From a functional viewpoint, it seems natural to present commands early in the course sequence, at least for receptive purposes, whether TPR per se is actively used or not. It is useful, for example, to be able to ask students to open their books, take out pen and paper, close the door, answer a given question, repeat something, say something, read something, sit down, etc. Likewise, young children learning their first language are likely to be exposed to Imperative forms very early as parents tell children to come here, go to sleep, eat your dinner, sit down, stop jumping, pick up your toys, etc.

In light of Bybee's (1985) observation that the 3 Singular Present Indicative is usually the default form in Spanish and the first to be acquired (56; 59-60), and because of the high functional usage of command forms in both first and second language acquisition, it is interesting to note that in Spanish, unlike English, the 3 Singular Present

Indicative and the affirmative Familiar Imperative forms are generally identical forms, as seen in the examples in Figure 7.5 below.

Even the eight so-called ‘irregular’ affirmative familiar commands are mostly just truncated forms of what would be the ‘regular’ form, that is, a form identical to the 3 Singular form. Figure 7.6 lists all of the irregular affirmative familiar command forms, of which the first six are just truncated forms of the 3 Singular form. The first three (*haz, pon, sal*) simply delete the final vowel; the fourth form (*di*) deletes the entire final syllable, that is, the final vowel and the consonant preceding it; the fifth and sixth delete the final vowel and monophthongize the diphthong in the preceding syllable (the stem): *ten* and *ven*, respectively. All six of the truncated forms are possibly a result of natural diachronic changes, since lenition and deletion of final syllables are common phenomena of sound change in languages around the world. The last two, *ve* and *sé* are, therefore, the only ones that genuinely depart from the typical form that happens to be identical, or at least similar, to the 3 Singular form.

FIGURE 7.5. Spanish 3 Sg Present Indicative and Affirmative Familiar Imperative forms.

INFINITIVE	3 SG PRES INDIC	FAMILIAR COMMAND (AFFIRMATIVE)
<i>hablar</i> (speak)	habla	habla
<i>cantar</i> (sing)	canta	canta
<i>tomar</i> (take, drink)	toma	toma
<i>lavar</i> (wash)	lava	lava
<i>beber</i> (drink)	bebe	bebe
<i>comer</i> (eat)	come	come
<i>traer</i> (bring)	trae	trae
<i>seguir</i> (follow, continue)	sigue	sigue
<i>dormir</i> (sleep)	duerme	duerme

FIGURE 7.6. Spanish Irregular Affirmative Familiar Imperative forms compared with 3 Singular Present Indicative forms for the given verbs.

INFINITIVE	3 SG PRES INDIC	FAMILIAR COMMAND (AFFIRMATIVE)
1. <i>hacer</i> (do, make)	hace	haz
2. <i>poner</i> (put, place)	pone	pon
3. <i>salir</i> (leave)	sale	sal
4. <i>decir</i> (say, tell)	dice	di
5. <i>venir</i> (come)	viene	ven
6. <i>tener</i> (have)	tiene	ten
7. <i>ir</i> (go)	va	ve
8. <i>ser</i> (be)	es	sé

The conventional verb organizational system based on tense/mood/aspect with six-member person/number-based sets for any given lexeme is pedagogically useful and is a convenient way of organizing verb forms for reference purposes. This type of organization is also consistent with Bybee and Pardo's hypothesis of a hierarchical relationship among the surface forms of verb paradigms. Contrary to the generativist notion of an abstract underlying representation for a verb paradigm, in which the surface forms "do not have any official relation to one another" (Bybee and Pardo 1981: 947), Bybee and Pardo found evidence in empirical investigations with speakers that there is psychological validity to such paradigms and that there are indeed relations among surface forms. These researchers found strong evidence that "the person/number forms within one tense or aspect are more closely related to one another than they are to forms outside that tense or aspect" (947). This finding is consistent with Bybee's (1985) contention that aspect and tense are more semantically related to the stem than are person and number (35-37).

Nevertheless, there may be other ways of organizing verbal inflectional forms. The way we organize these forms for descriptive, analytical or pedagogical purposes may or may not be consistent with the psychological reality, that is, the actual mental organization of verb forms. Furthermore, it is conceivable that the way speakers mentally organize such forms for storage purposes may not be the same as the way the way they organize them for processing or retrieval purposes.

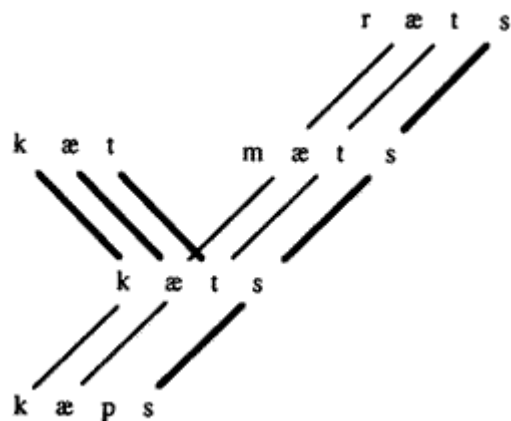
7.4. Constants and variables in verb paradigms.

The above-mentioned ways of organizing verb forms, both the conventional grammar textbook way and the hypothetical third-person-based paradigm or the absurdly large number-based paradigm, are all based on the lexical stem as the constant in the paradigm. All of the above paradigms assume that we are talking about the same verbal lexeme throughout the set of paradigms, where the variants are the phonological strings that represent person, number, tense, mood, and aspect, that is, the phonological strings traditionally viewed as bound grammatical morphemes. There may, however, be different ways of organizing the forms with different constants and variables. In Anderson's model (1995), as in Bull's (1965) three-part verb analysis and in the model proposed in the present study, the constants are the PN/TMA (person/number and tense/mood/aspect) markers, and the variables are the lexical stems, or, in traditional terminology, the content morphemes.

Bybee (1988) also presents a different kind of mental organization, one based on analogical connections. These analogical connections link similar forms based on similarities of their phonological and semantic components. Therefore, in this view, there is an overlapping of paradigms. There is a mental grouping of all forms with similar stems, as in the conventional paradigms, for example in the Preterite and Imperfect Subjunctive forms for *tener* 'have, possess': *tuve, tuviste, tuvo, tuvimos, tuvisteis, tuvieron, tuviera, tuvieras, tuviese, tuvieses*, etc. However, in Bybee's model, there is also a mental grouping of (or at least an analogical connection between) all forms of all verbs that end in unstressed /o/, for example, as in *hablo* ('speak', 1 Singular Present Indicative) and *escribo* ('write', 1 Singular Present Indicative). This grouping,

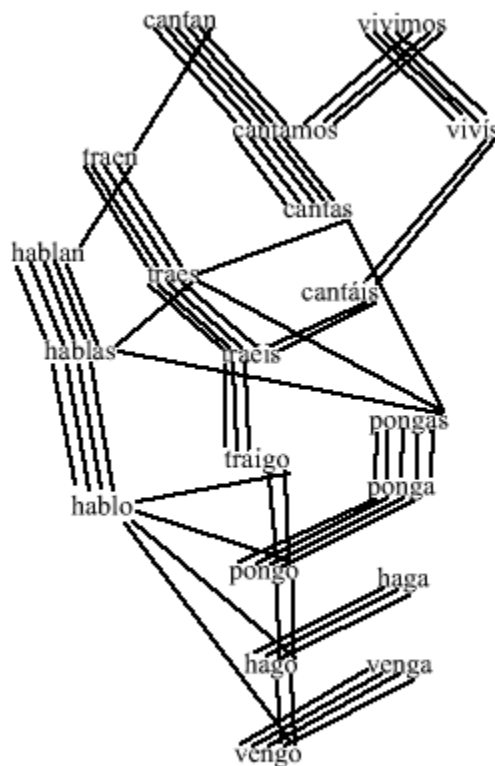
then, is based not on the base lexeme or even on the stem alternate, but rather on person and number alone, like the above mentioned third-person paradigm. Bybee's (1988) model, then, is multi-dimensional and represents overlapping paradigms based on the strengths of the connections between the semantic values and between the various phonological strings that represent the entire inflected word. It is thus an overlapping of paradigms based on arrangement by so-called grammatical markers or inflectional affixes as well as by lexical stem. For example, Figure 7.7 shows the connections between the English singular noun form *cat*, the plural form *cats* and other regular English noun plurals (Bybee (1988: 127)). The word *cat* is analogically connected to its plural form *cats*, due to shared semantic and phonological features. In addition, the word *cats* shares connections with other regular English noun plural forms, since they share the semantic feature *plural* as well as the word-final voiceless fricative.

FIGURE 7.7. Semantic and phonological connections among the English singular noun form *cat*, the plural form *cats* and other English noun plurals (Bybee (1988: 127)).



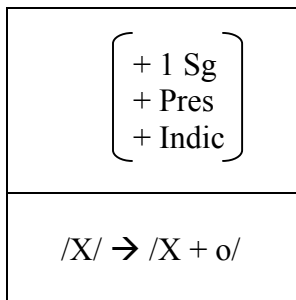
Likewise, there exist connections among Spanish verbs based on shared semantic and phonological features, such as the sample connections illustrated in Figure 7.8 for the forms of *traer*, *cantar*, *hablar*, *hacer*, *poner*, *venir*, *vivir*. This illustration is just a small sample of such analogical connections. As Bybee notes, “the number of relations quickly multiplies and defies visual representation, except perhaps as something that resembles a bowl of spaghetti” (Bybee, 1988: 126). Thus to comprehend the whole concept illustrated in Figures 7.7 and 7.8, one must mentally extend the connections shown in the diagram so that they connect almost endlessly with other inflected words, so that the resulting image is indeed spaghetti-like.

FIGURE 7.8. Semantic and phonological connections among Spanish verb forms.



The model presented in this investigation, although based in general form on Anderson's Extended Word-and-Paradigm model of morphological organization, is theoretically consistent with Bybee's analogical model of morphology, since the constants and variables are different from the constants and variables of traditional verb paradigms. The constants in the present model are the phonological strings representing grammatical information, meaning the strings signifying person, number, tense, mood, and aspect, while the variables are phonological strings signifying the so-called content part of the whole inflected word. For instance, in the following rule-and-feature set (Figure 7.9) from one of the blocks in the proposed model, the X variable might be /aβl/ 'speak' or it might be /teng/ 'have'; /βiβ/ 'live'; or /ag/ 'do,make', but the constant in the paradigm is the second part of the string, /o/, signifying a semantic-grammatical value of first person singular present indicative all in one simple vowel sound.

FIGURE 7.9. Sample Word-Formation Rule from the proposed verb model.

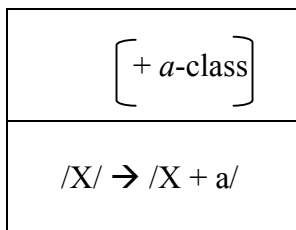


Although the vowel /o/ happens to be a suffix attached to what appears to be some lexical stem, we are clearly not dealing with strictly agglutinative morphology here, since this one vowel signifies at least four functions all at once: person, number, tense, mood, and possibly aspect. This block-style graphic could possibly be seen as just a different

way of illustrating Bybee’s (1985, 1988) multi-dimensional model, where all forms with /o/ are grouped together on one level. In the proposed model, which is presented in detail in Chapters 5 and 6, there are actually several rule-and-feature sets (also known as Word-Formation Rules) in each block in the model, not just one as in the example in Figure 7.9. At this point, one set from one of the blocks in the model is shown by itself simply for illustration purposes.

In the proposed model, what is viewed traditionally as the lexical stem is the variable, at least in the beginning, at the outset of each of the sequences of blocks of feature sets with their corresponding word formation rules. However, the X variable adds more phonological material as the blocks progress in a conjunctively ordered manner. For example, the above rule-and-feature set produces a final form /aβlo/. It is a final form in that it adds no more phonological strings; in other words, it produces an entire inflected word, *hablo*. Another rule-and-feature set in the same block that produced *hablo* is the set shown in Figure 7.10, which produces *habla* /aβla/.

FIGURE 7.10. Another sample rule-and-feature set from the proposed verb model.

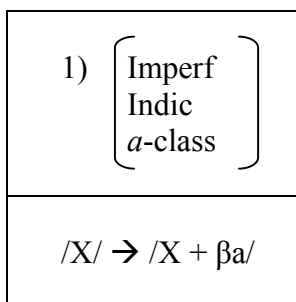


Note that there is no PN or TMA designation in the top half of this set as there is in the previous set that produced *hablo* /aβlo/. Consequently, the form that exits the block as a result of this set – *habla* /aβla/ – is a sort of default form, containing what is

commonly referred to as a ‘zero morph’. Since in the proposed model the blocks operate on a most-specific to most-general basis, this leaves /aβla/ as the form that is left after more specific forms are otherwise specified in the other sets in the block, such as the above *hablo* /aβlo/. Therefore, if in later blocks there are no PN or TMA features that specify 3 Singular Present Indicative, then the /aβla/ form, as a sort of default form, remains as the 3 Singular Present Indicative form without having to explicitly specify it as such. This notion of 3 Singular forms as default forms in the proposed model is interesting in light of the fact that Spanish-speaking children acquire the 3 Singular Present Indicative first and then use it as a default form until they acquire other forms, as discussed earlier (Bybee 1985: 50, 59-60). Bybee also points out that in Spanish this form, which also happens to be the 2 Singular Familiar Imperative, such as *habla, canta, bebe, come*, is “The basic, or least marked form” and that “the semantically basic and most frequent form, the 3s of the Present Indicative, can serve as a basis for the derivation of all other forms” (60).

Furthermore, since the /aβla/ form exits the first block with no specifiers for PN or TMA, then this form is now the X variable entering the next conjunctively ordered block. This next block includes the set shown in Figure 4.10.

FIGURE 7.11. Sample rule-and-feature set from Block II of the Basic Model.



The form produced by this block, therefore, is *hablaba* /aβlaβa/. This form now enters the next block (shown in Figure 7.12) as the new X variable, where it will add the appropriate phonological strings corresponding to the specified PN features.

FIGURE 7.12. Block III of the Basic Model, with four rule-and-feature sets.

1) $\left[+ 2 \text{ Sg} \right]$	2) $\left[+ 1 \text{ Pl} \right]$	2) $\left[+ 2 \text{ Pl} \right]$	3) $\left[+ 3 \text{ Pl} \right]$
/X/ → /X + s/	/X/ → /X + mos/	/X/ → /X + is/	/X/ → /X + n/

Since the rules within a given block proceed from most specific to most general, there is no need here to specify a separate rule and feature set for 1 Singular or 3 Singular. Just as the above case of *habla* as the final form, which remained unspecified for PN and TMA, the form *hablaba* exits the block as a default form, leaving it as the final form for the unspecified 1 Singular and 3 Singular of the Imperfect Indicative for *a*-class verbs.

The block presented above in Figure 7.12 is largely consistent with the third part of Bull's (1965: 112-121) three-part verb analysis. Spanish verbs, according to Bull, are made up of a stem plus two suffixes. The first part, which Bull considers the constant, is the verb stem, or the so-called content part of the verb. The second part, which he labels the variable, is the TMA slot, and the third part, also a constant, is the PN slot. He sums up the third slot (PN) as follows: "singular subject is marked by \emptyset or *s*; plural subject by *mos*, *n*, and *is*" (Bull 1965: 113). The above block of rule-and-feature sets essentially says the same thing, although the zero morph is not specified by any set in the block, but

rather is the default form because it is not specified. Also, the PN/TMA markers are the constants in the proposed model, and the content-bearing stem is the variable, at least at the outset.

7.5. External organization of verb paradigms.

The above discussion of the internal organization of verb paradigms referred to the person/number organization within the tense/mood/aspect based sets. External organization, then, means how those six-member sets are grouped based on tense, mood, and aspect. The paradigms presented in conventional Spanish textbooks always comprise six forms each, that is, their internal organization, whether they are shown in one or two columns as discussed above. These sets of forms, however, may be grouped externally in different ways. The sets are sometimes grouped according to mood, as in Dozier and Iguina (1999); Zatarain, Zatarain, and Romero (1998); Seco (1996); and Lloréns Camp (1998). Others, such as Kendris (1990) and Noble and Lacasa (1997), choose to organize according to whether the forms are simple or compound forms.

Dozier and Iguina (1999: 147-174) present first the Indicative mood forms, then the forms of what they refer to as the Conditional ‘mood’, then those of the Subjunctive mood, labeling these forms as shown in Figure 7.13 below.

FIGURE 7.13. Dozier and Iguina’s presentation of finite verb forms (1999: 147-174).

INDICATIVE MOOD
1. <i>present indicative</i> , as in <i>hablo</i> (I speak/am speaking/do speak)
2. <i>imperfect indicative</i> , as in <i>hablaba</i> (I spoke/was speaking/used to speak)
3. <i>preterite</i> , as in <i>hablé</i> (I spoke)

<ol style="list-style-type: none"> 4. <i>present perfect indicative</i>, as in <i>he hablado</i> (I have spoken) 5. <i>pluperfect indicative</i> (called <i>past perfect indicative</i> in some texts), as in <i>había hablado</i> (I had spoken) 6. <i>simple future</i>, as in <i>hablaré</i> (I will speak) 7. <i>future perfect</i>, as in <i>habré hablado</i> (I will have spoken)
CONDITIONAL MOOD
<ol style="list-style-type: none"> 1. <i>present conditional</i>, as in <i>hablaría</i> (I would speak) 2. <i>conditional perfect</i>, as in <i>habría hablado</i> (I would have spoken)
SUBJUNCTIVE MOOD
<ol style="list-style-type: none"> 1. <i>present subjunctive</i>, as in <i>hable</i> (I speak/may speak/might speak) 2. <i>imperfect subjunctive</i>, as in <i>hablara</i> (I spoke/might speak) 3. <i>present perfect subjunctive</i>, as in <i>haya hablado</i> (I have spoken) 4. <i>pluperfect subjunctive</i>, as in <i>hubiera hablado</i> (I had spoken)

Dozier and Iguina then list the Imperative mood forms (Figure 7.14), and then the non-finite forms (Figure 7.15), which include the Present Infinitive, the Perfect Infinitive, the Present Participle/Gerund, and the Past Participle, but they do not list the Perfective Participle form *habiendo hablado*. The Imperative mood forms might be considered an incomplete paradigm compared with the others listed above, since there is, of course, no first person Imperative for obvious semantic reasons, as in the example in Figure 7.14 for the verb *hablar*.

FIGURE 7.14. Spanish Imperative forms listed as a six-member PN-based paradigm.

IMPERATIVE MOOD (DOZIER AND IGUINA 1999: 147-174)		
	SINGULAR	PLURAL
1 ST PERSON	(no form)	<i>hablemos</i> ‘let’s speak’
2 ND PERSON	<i>habla/ no hables</i> ‘speak/don’t speak’	<i>hablad</i> ‘speak’
3 RD PERSON	<i>hable/no hable</i> ‘speak/don’t speak’	<i>hablen/no hablen</i> ‘speak/don’t speak’

FIGURE 7.15. Dozier and Iguina’s presentation of non-finite verb forms (1999: 147-174).

NON-FINITE FORMS (DOZIER AND IGUINA 1999: 147-174)
<ol style="list-style-type: none"> 1. <i>present infinitive</i>, as in <i>hablar</i> ‘to speak’ 2. <i>perfect infinitive</i>, as in <i>haber hablado</i> ‘to have spoken’ 3. <i>present participle</i> or <i>gerund</i>, as in <i>hablando</i> ‘speaking’ 4. <i>past participle</i>, as in <i>hablado</i> ‘spoken’

Zatarain et al (1998) and Seco (1996) also organize the paradigms based on mood. However, neither Zatarain nor Seco list Conditional as a separate mood, but rather as forms of the Indicative mood, labeled ‘Pospretérito’ and ‘Antepospretérito’. (The functional significance of labels such as these is discussed in Chapter 4 of the present study.) Kendris (1990), on the other hand, organizes the paradigms based on the ‘seven simple tenses’ and the ‘seven compound tenses’, plus the set of Imperative forms. Noble and Lacasa (1997) also use the *simple* vs. *compound* criterion (‘*tiempos simples*’ y ‘*tiempos compuestos*’), plus a separate grouping for Imperatives and non-finite forms (‘*formas imperativas, gerundios, y participios*’). Bybee (1985), however, groups the non-

compound, or non-periphrastic, Spanish verb paradigms according to ‘present’, ‘past’, and ‘subsequent’. Under the heading ‘Present’ are the Indicative and Subjunctive forms of the Present tense. Under ‘Past’ we find Preterite, Imperfect, and Past Subjunctive. Bybee’s ‘subsequent’ category includes Future and Conditional. It should be noted that Bybee (1985), however, unlike the other sources named above, is not a grammar reference text or a beginning language textbook, but rather a non-language-specific psycholinguistic study of morphology and the relation between meaning and form.

7.6. Variations in terminology.

In addition to this variation in the organization of the six-member sets of forms, there is considerable variation in the terminology used for each of the tenses and moods. Figure 7.16 shows a sampling of the various terms used in some of the texts consulted for the present investigation. The source for the two Real Academia Española columns is from Manuel Seco Reymundo of the Real Academia Española (1996: 268).

FIGURE 7.16. Sample nomenclature of Spanish verb forms (with the verb *cantar* ‘sing’).

SAMPLE VERB FORM (FIRST PERSON SINGULAR)	BULL 1965; DOZIER & IGUINA 1999 (English version, used in the present analysis and model)	DOZIER AND IGUINA 1999 (their Spanish version of terms)	SECO 1999; ZATARAIN ET AL 1998	REAL ACADEMIA ESPAÑOLA 1931	REAL ACADEMIA ESPAÑOLA 1973
<i>1. canto</i>	Present indicative	Presente, modo indicativo	Presente, modo indicativo	Presente, modo indicativo	Presente, modo indicativo

2. <i>cantaba</i>	Imperfect indicative	Imperfecto, modo indicativo	Copretérito, modo indicativo	Pretérito imperfecto, modo indicativo	Pretérito imperfecto, modo indicativo
3. <i>canté</i>	Preterite	Pretérito, modo indicativo	Pretérito, modo indicativo	Pretérito indefinido, modo indicativo	Pretérito perfecto simple, modo indicativo
4. <i>cantaré</i>	Future	Futuro, modo indicativo	Futuro, modo indicativo	Futuro imperfecto, modo indicativo	Futuro, modo indicativo
5. <i>cantaría</i>	Conditional	Presente, modo condicional	Pospretérito, modo indicativo	Potencial simple o imperfecto, modo indicativo	Condicional, modo indicativo
6. <i>cante</i>	Present subjunctive	Presente, modo subjuntivo	Presente, modo subjuntivo	Presente, modo subjuntivo	Presente, modo subjuntivo
7. <i>cantara/cantase</i>	Imperfect subjunctive	Imperfecto, modo subjuntivo	Pretérito, modo subjuntivo	Pretérito imperfecto, modo subjuntivo	Pretérito imperfecto, modo subjuntivo
8. <i>cantare</i>	(not included)	(not included)	Futuro, modo subjuntivo	Futuro imperfecto, modo subjuntivo	Futuro, modo subjuntivo
9. <i>ha cantado</i>	Present perfect indicative	Present perfecto, modo indicativo	Antepresente, modo indicativo	Pretérito perfecto, modo indicativo	Pretérito perfecto compuesto, modo indicativo
10. <i>había cantado</i>	Pluperfect indicative	Pluscuamperfecto, modo indicativo	Antecopretérito, modo indicativo	Pretérito pluscuamperfecto, modo indicativo	Pretérito pluscuamperfecto, modo indicativo
11. <i>hubo cantado</i>	(not included)	(not included)	Antepretérito, modo indicativo	Pretérito anterior, modo indicativo	Pretérito anterior, modo indicativo
12. <i>habré cantado</i>	Future perfect	Futuro perfecto, modo indicativo	Antefuturo, modo indicativo	Futuro perfecto, modo indicativo	Futuro perfecto, modo indicativo
13. <i>habría cantado</i>	Conditional perfect	Perfecto, modo condicional	Antepospretérito, modo indicativo	Potencial compuesto o perfecto, modo indicativo	Condicional perfecto, modo indicativo
14. <i>haya cantado</i>	Present perfect subjunctive	Presente perfecto, modo subjuntivo	Antepresente, modo subjuntivo	Pretérito perfecto, modo subjuntivo	Pretérito perfecto, modo subjuntivo
15. <i>hubiera cantado</i>	Pluperfect subjunctive	Pluscuamperfecto, modo subjuntivo	Antepretérito, modo subjuntivo	Pretérito pluscuamperfecto, modo subjuntivo	Pretérito pluscuamperfecto, modo subjuntivo

16. <i>hubiere cantado</i>	(not included)	(not included)	Antefuturo, modo subjuntivo	Futuro perfecto, modo subjuntivo	Futuro perfecto, modo subjuntivo
17. <i>canta</i>	<i>Tú</i> imperative, affirmative	Imperativo	Futuro, modo imperativo (just labeled imperativo in Zatarain)	Presente, modo imperativo	Presente, modo imperativo
18. <i>cante</i>		Imperativo	Presente, modo subjuntivo (just labeled imperativo-usted in Zatarain)	Presente, modo subjuntivo	Presente, modo subjuntivo
19. <i>cantemos</i>		Imperativo	(Not explicitly listed in Seco or Zatarain)	(Not explicitly listed in Seco 1999)	(Not explicitly listed in Seco 1999)
20. <i>cantad</i>	<i>Vosotros</i> imperative, affirmative	Imperativo	(Not listed in Seco; just listed as imperativo-vosotros-as in Zatarain)	(Not explicitly listed in Seco 1999)	(Not explicitly listed in Seco 1999)
21. <i>canten</i>		Imperativo	(Not listed in Seco; just listed as imperativo-ustedes in Zatarain)	(Not explicitly listed in Seco 1999)	(Not explicitly listed in Seco 1999)
22. <i>cantar</i>	Present infinitive/ Simple infinitive	Infinitivo	Infinitivo	Infinitivo	Infinitivo
23. <i>cantando</i>	Present participle (or Gerund, in Dozier and Iguina)	Participio presente	Gerundio	Gerundio	Gerundio
24. <i>cantado</i>	Past participle	Participio pasado	Participio	Participio	Participio
25. <i>haber cantado</i>	Perfect infinitive	(Spanish term not presented)	Infinitivo compuesto	Infinitivo compuesto	Infinitivo compuesto
26. <i>habiendo cantado</i>	(not included)	(Spanish term not presented)	Gerundio compuesto	Gerundio compuesto	Gerundio compuesto

The official website of the Real Academia Española (<http://dle.rae.es>) gives simple verb forms only, not compound forms such as Present Perfect Indicative *ha hablado*, Pluperfect Indicative *había hablado*, Present Perfect Subjunctive *haya hablado*, or Pluperfect Subjunctive *hubiera hablado*. For the simple forms, the only differences in the online Real Academia (2002) compared with the Academia 1931 and Academia 1973 (according to Seco 1999) are the following:

1. For verb form # 2 (labeled ‘Imperfect Indicative’ in the present study), the Real Academia 2002 uses the term ‘Copretérito’ in addition to ‘Pretérito Imperfecto’ under the ‘Indicativo’ column.
2. For verb form # 3 (labeled ‘Preterite Indicative’ in the present study), the Real Academia uses the term ‘Pretérito Perfecto Simple’ o ‘Pretérito’ under the ‘Subjuntivo’ column instead of just ‘Pretérito Perfecto Simple’.
3. For verb form # 5 (called ‘Conditional’ in the present study), the Real Academia uses the term ‘Condicional Simple o Pospretérito’ under the ‘Indicativo’ column.
4. For verb form # 7 (labeled ‘Imperfect Subjunctive’ in the present study), the Real Academia uses the term ‘Copretérito’ in addition to ‘Pretérito Imperfecto’ under the ‘Subjuntivo’ column.

Bybee (1985: 61) uses the same labels as does Bull (1965: 112-121) and Dozier and Iguina (1999: 147-174), except that Bybee uses the term ‘Past Subjunctive’ instead of ‘Imperfect Subjunctive’. Also, as mentioned above, Bybee organizes the Spanish verb paradigms according to Present (Indicative and Subjunctive), Past (Preterite Indicative, Imperfect Indicative, and Past Subjunctive), and Subsequent (Future and Conditional).

7.7. Conclusion.

The functional significance of the organization and labeling of verb paradigms was discussed in Chapter 4 of the present analysis, while the current chapter has focused on organization and labeling per se. This overview of the traditional ways of organizing and labeling verb paradigms for analytical and pedagogical purposes, along with the comparison of Spanish with several other languages in this regard, serves as background and justification for the grammatical notations used in the present proposed model of Spanish inflectional verbal morphology. Such notations are important because in the model's feature sets these grammatical notations are used as shorthand abbreviations for the many semantic values that such labels might signify in a given context, as discussed in Chapter 4.

The current chapter's discussion also points out the variability, arbitrariness, and language specificity of verb paradigms and their labels and organization. The psychological validity of paradigms is considered, contrasting Bybee's (1985) comments about verb paradigms with the generativist notion of abstract underlying representations for the surface forms of verb paradigms. In addition, the discussion of the role of constants and variables in different kinds of verb paradigms helps to establish a basis for the constants and variables used in the proposed model. The discussion of the psychological validity of paradigms shows the relevance of paradigms to an analogical model of morphology such as Bybee's (1985, 1988) and the possible relevance of different kinds of paradigms used for pedagogical purposes.

CHAPTER 8

BINDING-ACCESS, CONNECTIONS, AND INPUT

8.0. This chapter first looks at why verbs are important in second language acquisition, and then discusses Terrell's (1986, 1991) binding-access model of language acquisition and how it relates to Bybee's (1985, 1988) analogical model of morphology and to the model proposed in the present analysis. Paradigms in language acquisition are then considered, and then the lexicon, lexical access, and matching procedures in the lexicon are discussed. Subsequent sections consider the role of input and interconnections in language acquisition and the importance of frequency of forms as related to regularity and irregularity of forms.

Asher's (1982) observations about the vast differences and sharp contrasts between child and adult language acquisition lead into the discussion of explicit grammar instruction and the role of input and output in the proposed model. This chapter takes another look at the nature of rules and relates the concept of rules as descriptive patterns to language systematicity and variation, showing how the proposed model reconciles Anderson's rule-based model with Bybee's analogy-based model. Finally, this chapter considers the extremely complex, dynamic nature of the lexicon, the human brain, and language acquisition, and sums up the pedagogical implications of the present analysis. The model supports VanPatten's Input Processing model (Lee and VanPatten 1995), Krashen's input hypothesis (1982), Krashen's emphasis on leisure reading for comprehensible input at the intermediate level (1995), as well as Communicative

Language Teaching's emphasis on oral input at all levels. In addition, the model does support some degree of explicit grammar instruction in the classroom.

8.1. The pedagogical significance of lexical organization.

If the model proposed in the present analysis is psychologically plausible, and if it is consistent with Bybee's view of morphology as lexical organization, then it has strong implications for second language acquisition and for the teaching of a second language.

Emmorey and Fromkin (1988) comment that

Psychological models have generally been concerned with how lexical information is accessed or processed and have been less explicit about the representation and structure of the information. The nature of stored representations is important, however, because it may in part determine the nature of the access mechanisms (124).

If inflectional verbal morphology is lexicon-based, that is, if inflected verb forms are composed and stored in the lexicon, based on links between the two parts of the linguistic sign and links among semantically and phonologically similar forms, then it is imperative that the learner have extensive, meaningful exposure to the language. In order for a learner to truly acquire the target language, he or she must be exposed to a significant amount of comprehensible input. This input is necessary for establishing the form-meaning connections that build the feature sets associated with a given verb form, and for establishing and strengthening the analogical connections among forms.

Bybee (1988) refers to the "independently necessary mechanisms of lexical storage" as "the ability to form networks among stored elements of knowledge and the

ability to register the frequency of individual items and patterns” (125). Since, according to Bybee, the frequency of occurrence of a given form determines the strength of that form in a given person’s lexicon, it is essential that a learner be exposed to sufficient tokens, or exemplars, of a given form in meaningful contexts. Whether this input is in the form of reading or listening, or even perhaps in the form of structured output, the feature sets are built, refined, and reinforced through such input. This input, along with the process of building and refining feature sets, strengthens the connections between the *signifié* and the *signifiant* and the connections among forms with partial phonological and semantic connections, thus making those forms more readily available to the speaker for both receptive and productive tasks. Since this connection between the *signifié* and the *signifiant* is essentially arbitrary in any human language, it is especially important for a learner to establish mental connections between a given form and its function in order to make the sign more accessible to the speaker in language processing and production.

8.2. Why are verbs so important anyway?

Terrell (1991) comments that the acquisition of words and morphology is the primary emphasis in beginning levels of foreign language instruction, at least in a typical foreign language classroom the United States. A quick scan of any beginning or intermediate Spanish textbook (or many other language textbooks as well) confirms Terrell’s comment that explicit grammar instruction at beginning levels “involves morphology or morphosyntax: tense, plurality, case, mood, and so forth” (56).

Spanish, unlike its predecessor Latin, does not have noun declension, although Spanish pronouns do have case distinctions. Spanish pronouns vary according to

syntactic function, as well as gender and number, such as *él* (3 Singular) in subject position or as object of a preposition, *lo* (3 Singular) in direct object position, and *le* (3 Singular) in indirect object position. However, although Spanish has no noun inflections (other than pronouns), it does maintain a rich verbal morphology, unlike English.

Therefore, most of the grammar taught to students of Spanish involves verb morphology.

Because of the constant presence of verb conjugations in textbooks and the attention given to verbs in most language courses, verbs seem to be the bane of foreign language students, who groan when an instructor begins to put verb conjugation charts on the board. Yet many complain when grammar is not explicitly taught. Instructors shake their heads in disbelief when students continue to make the same morphological errors over and over again in spite of repeated explanations and practice. Meanwhile, debates rage on about how to teach grammar in the classroom and whether grammar should be explicitly taught. It would naturally follow that any insight into the mental organization and processing of this complex verbal system would be beneficial to teachers of Spanish as a second language and might help them formulate their own approach to teaching this system. It is appropriate, then, to question how lexical and morphological information is organized and accessed by the learner and to examine the psycholinguistic processes used in organizing, processing, and producing linguistic forms.

8.3. Binding-access: comprehension strategies and production strategies

Terrell (1986) defines the acquisition of a form of the target language as “the process which leads to the ability to understand and produce that form correctly in a communicative context” (213), maintaining that “language acquisition consists of at least

two paired components: comprehension strategies and binding and production strategies and access” (1991: 56). This section will look at these two components of language acquisition from Terrell’s viewpoint (1991: 56-58; 1986: 213-227).

Comprehension strategies, as defined by Terrell, are the ways by which a listener makes sense of the input, such as listening for key words and paying attention to context. Terrell uses the term ‘binding’ to refer to “the psycholinguistic linking of meaning to a new form in the target language” (1991: 56) or “the cognitive and affective mental process of linking a meaning to a form” (1986: 214), whether that form is a simple whole word, a complex whole word, or a part of a word, such as a grammatical marker.

Production strategies, on the other hand, are the ways by which a learner or speaker puts linguistic forms together in a linear sequence to constitute output, such as naming a topic and then making a comment about it, and stringing forms together in the appropriate word order for the target language. Terrell uses the term ‘access’ to refer to “the retrieval and production of a form to express some intended meaning in the target language” (1991: 56).

Based on these notions of binding and access, acquisition of a form in the target language, then, is defined as “establishing a connection between concept and form” (56). In Saussure’s terms, this means establishing a connection between the two parts of the linguistic sign: the *signifié* and the *signifiant*. A meaning-form connection means that the speaker/listener can understand the meaning of a given form in the input and that he or she “can access and produce the correct form without undue delay” (56) in language production. For Terrell, then, the acquisition of forms consists of “the positing and storing of pairs of meaning-form relationships and the restrictions on the appropriate

access of these forms to express an intended meaning” and the task of the learner is “to use the input to posit and store correct meaning-form relationships” (56).

This binding task is not too difficult if the parameters in the target language are similar to those of the native language, but when the parameters are different, the task is more difficult. For example, a monomorphemic word such as *boca* ‘mouth’ may not be difficult to acquire if the semantic value of the word in the target language largely parallels the semantic value of the corresponding word in the native language. But to acquire a polymorphemic word, such as all verb forms in Spanish, the task is more difficult, especially to the learner whose native language has different categories and organizes the lexicon in a different way than the target language does, that is, in a way that is unfamiliar to the learner.

For example, in the present author’s experience, many English-speaking students of Spanish have great difficulty acquiring – or even consciously learning – the concept of verb conjugation. According to Terrell’s binding-access model, this difficulty is partially due to the fact that the referents for Spanish verb forms do not exactly correspond to the referents for English verbs. Spanish, for instance, has a split concept for past tense indicative mood that English does not have: the Imperfect Indicative and the Preterite Indicative. The Imperfect Indicative is used for ongoing, habitual actions or states in the past or for actions or states for which the completion is not important or specified. Preterite forms are used for specific events or states seen as completed, or for the beginning or ending of specific events. Also, Spanish has person/number distinctions in all verb forms in all tenses, while English has only the suffix /ed/ (realized as [ed], [d], or [t]) for all persons and numbers in regular Past tense forms. Even the irregular Past tense

forms in English do not have person/number distinctions, except for the Past forms of the verb *to be*. For native English speakers to establish form-meaning connections for forms in Spanish that are not parallel to English forms, they have to bind the forms in different ways to the concepts they represent in order to access the correct form.

Terrell contends that although linguists use metalinguistic terms to describe rules for producing grammatically correct forms, the actual cognitive processes used during speech comprehension and production are not consistent with our popular concept of rule use. In Terrell's binding-access framework, "what must be perceived and stored in the acquisition process are the individual meaning-form pairs, not a grammatical rule" (1991: 57). Thus, "the ability to generalize patterns to new forms and contexts is not due to the learner's having formulated (subconsciously) a rule, but rather is based on a network of meaning-form connections" (57). Native speakers of Spanish are able to generate forms for new, borrowed, or even nonexistent verbs (as in experiments to evoke such forms), such as creating *puchar* out of English *push*; *cachar* from English *catch*; or *formatear* from English *format*, and then conjugating it as any other verb ending in *-ar*. For Terrell, this is not because "they have acquired some morphological 'rule', but, rather, by analogy with the hundreds of other" verb forms already stored (57).

8.4. Terrell's binding-access hypothesis and Bybee's analogical model

Terrell's (1986, 1991) binding-access hypothesis is consistent with Bybee's (1985, 1988) analogical model of morphology. The connections among forms in the binding-access view, such as the connections among similarly inflected verb forms, are parallel to Bybee's analogical connections in the lexicon among semantically and phonologically

similar forms. Terrell, Bybee, and Skousen (1989) all regard language rules as descriptive tools of linguists rather than as generative mechanisms in the perception and production of language. So, in Terrell's framework, learners of Spanish as a second language in a class situation may consciously learn verb traditional conjugations and other 'rules' of the language, but "acquisition depends on [unconsciously] positing meaning-form connections for individual items" (57), just as in Bybee's model of morphology. Learners may consciously know, for instance, that **El niño bebí la leche* violates the rules of Spanish and is not correct because they have correctly learned verb conjugations and know that it should be *El niño bebió la leche* ('The child drank the milk'). If they have genuinely acquired the language, however, they will also recognize an incorrect form "because the words themselves are bound psycholinguistically to meanings with different restrictions on their occurrences" (57). The form *bebí*, for example, is bound to the concept of 'drink', but the occurrence of this particular form is restricted to contexts in which it is *yo* 'I' (1 Singular) drinking the milk, not another person, as in *bebió* (3 Singular). Thus, a native speaker of Spanish or a person who has acquired Spanish as a second language in a natural setting knows intuitively that **El niño bebí la leche* 'doesn't sound right', although he or she may not know exactly why.

This notion of binding-access is also consistent with the model proposed in the present analysis. The rules in the blocks of the proposed model represent the form-meaning relationship that must be established during language acquisition. The top half of each rule, which is the feature set representing the *signifié*, is the 'meaning' part of the form-meaning connection, and the bottom half, the *signifiant*, is the 'form' part of the connection. In Terrell's binding-access framework, just as in Bybee's model and in the

model proposed in the present analysis, each word “is connected to both its meaning and to other words by connections of various sorts: semantic, syntactic, morphological, phonological, and so forth” (Terrell 1991: 57). For example, the verb form *bebió* is connected semantically to *tomó* (which can also mean ‘drank’). *Bebió* shares a partial semantic and phonological connection with *bebí*, since they share initial phonemes and the semantic notion of ‘drink’. *Bebió* is also connected morphologically with *tomó*, since both forms represent a 3 Singular meaning, and they are thus also partially connected phonologically in that they both end with /ó/. In the present model and in Bybee’s model, as discussed earlier, this morphological connection is seen as essentially a semantic connection, since the person/number part of the meaning is really just part of the overall semantic notion represented by the given form, that is, the mental concept of the verb being performed by one person other than the speaker or the person spoken to, which we call ‘third person singular’. The verb form *bebí* is also connected phonologically to the noun *bebé* ‘baby’ by sharing the initial segment /βeβ/, but these two forms share no semantic connections.

According to Terrell’s binding-access model, forms are stored as single units, whole and unanalyzed, in the early stages of acquisition in a natural context. For example, some nouns are acquired first in the singular form, such as *coche* ‘car’, *casa* ‘house’, *boca* ‘mouth’; while some are acquired first in their plural form, such as *ojos* ‘eyes’, *pies* ‘feet’, *llaves* ‘keys’. At least in the early stages of acquisition, no particular part of the word is identified as plural, but, rather, is just part of the overall meaning of the word. As the lexicon grows, the connections also increase, and there are more and more exemplars of singular and plural forms and more and more connections between

and among these forms. These increased connections augment the learner's "ability to understand and generate new forms analogically based on the existing singular-plural and plural-plural connections" (Terrell 1991: 57).

The acquisition of verbal morphology occurs in a similar manner. For example, the learner in a natural context (whether child or adult) hears words such as *habla, bebe, come, escribe, trae, firma*, all 3 Singular forms, and coincidentally, also Familiar Imperative forms. In addition, the adult learner most likely hears Formal Imperative forms as well, such as *hable, firme, escriba*. The learner makes form-meaning connections from context, but, according to Terrell (57), the ending (the morphological marker in traditional terms) is unanalyzed in the early stages of acquisition. The learner gradually hears more and more exemplars of other forms of these and other verbs and continues to make semantic connections based on the overall meaning of the word in context. For example, the learner may hear *hablamos, vamos, comemos* (all 1 Plural forms) with the same basic meaning as *habla, va, come* (respectively), but in a context that includes the speaker and the addressee rather than a third party. As more and more forms are heard, more and more connections are made. For example, connections are made between Verb X [3 Singular] and Verb X [1 Plural], as in *hablamos* and *hablo*, while connections are also being made between Verb X [1 Plural] and Verb Y [1 Plural], as in *hablamos* and *cantamos*. The stronger these connections become through frequency of occurrence in the input, the more able the learner is to comprehend and produce new forms analogically based on the existing lexical connections among forms, such as *habla-hablo-hablas-hablamos* or *hablamos-comemos-bebemos-vamos*.

8.5. How these connections work in the proposed rule-and-feature based model.

In Terrell's and Bybee's models, there are multidimensional, multidirectional connections among forms in the lexicon, such as the above discussed connections between *habla-hablo-hablas-hablamos* or *hablamos-comemos-bebemos-vamos*. The model proposed in the present analysis contains variables, which at the outset are the content-bearing stems of the given verb forms, and constants, which are the phonological material added onto any given variable. One X variable plus the entire model as presented, with all rule blocks, equals the entire paradigm for the given X variable, that is, the entire set of conjugated forms for the given verb, such as in the charts listed in Kendrix' *501 Spanish Verbs* (1990).

Furthermore, each individual rule in the model plus multiple X variables equals a different kind of connection, that is, another direction or dimension, connecting a given affix with all verb forms with that same ending. For example, the rule that specifies /X/ → /X + o/ connects all verb forms that share the ending /o/ (unstressed), such as *hablo, canto, bebo, escribo, vivo, tengo, pongo*. The forms in such a connected network have different semantic values for the stem, but they share the same phonological ending and the same semantic value for person, number, tense, and mood, that is, first person singular, present tense, indicative mood. Similarly, all rules with [+ Pret] in their respective feature sets represent another dimension, or another kind of connection, that is, a semantic connection among all forms of all verbs that represent the various notions associated with the concept of 'preterite', such as *habló, hablaron, canté, cantaste, tuvo, tuvimos, viví, vivieron, vio, viste, fue, fuimos*. All these (and other Preterite forms) share a semantic connection, whether they share phonological material or not. Similarly, all

rules with [+ Subj] in their respective feature sets represent yet another dimension and another paradigm. For various X variables, the connections among forms with [+ Subj] in their respective feature sets are connections among forms that represent the concept ‘subjunctive mood’, regardless of tense, and regardless of the ‘regularity’ of the forms. Of course, there are also semantic connections among Imperfect Subjunctive forms and among Present Subjunctive forms. In addition, there are both semantic and phonological connections among forms in those sets that share phonological material as well as semantic features.

For any given individual learner, the rules in the blocks develop gradually over time, based on the strength of the analogical connections among forms, and this strength of connections is based on frequency of the given forms in the input. In the early stages there are fewer rules in the blocks, because there is not enough input to establish strong form-meaning connections and to establish analogical connections among forms. Thus, the early-stage learner may produce incorrect forms such as **yo bebió* or **ellos habla*. Also, if learners have not heard sufficient exemplars of a given irregular form, they may regularize the form based on analogical connections with similar forms in similar contexts, for example **yo sabo*, instead of *yo sé* [‘know’, 1 Sg Pres Indic], based on an analogical connection with *hablo, bebo, como*. There must be sufficient input, and therefore sufficient connections and sufficient lexical strength, for a ‘more specific’ rule to develop, such as /X/ = /sé/ for the feature set that includes [‘know’, 1 Sg Pres Indic]. After this rule is acquired, it takes precedence over – and blocks the application of – a more general rule, such as /X/ = /X +o/, which would produce the incorrect **sabo*.

Therefore, when student learners of Spanish as a second language produce forms such as **yo sabo* (for *yo sé*) or **él sabió* (for *él supo*), it is because they have not had sufficient input, and consequently they have insufficient lexical connections and lexical strength to establish the more specific rule that produces a suppletive form and blocks application of a later, more general rule. At first, without sufficient analogical connections, the learner is

forced to produce many new forms by a process of analogy to known items. This analogical process of generation is not perfect, since the connections themselves may be both imperfect and tenuous and the network itself too simple. Thus overgeneralizations and various sorts of regularization and ‘irregularization’ may occur (Terrell 1991: 57).

Thus, in the present model, as in Terrell’s binding-access model, the morphological development of a learner is a dynamic, ongoing process, with new forms constantly being “processed from the input and stored along with their connections to other related forms” (57).

According to Terrell (1982, 1986, 1991), in order for a learner to acquire a morphologically complex form, the learner must be able to isolate the form in the input, ascertain the meaning of the given form from the context, and then establish various kinds of associations and connections between the form and the meaning, as well as between the form and other similar forms. This is the process necessary for the establishment of form-meaning connections in the model proposed in the present analysis, as well as in Terrell’s binding-access model, and is consistent with Bybee’s (1985, 1988) model as well.

All three of these models (the proposed model, as well as Bybee's and Terrell's) support Krashen's (1982, 1995; Krashen and Terrell 1983) notion of 'comprehensible input' as an essential factor in language acquisition. According to Krashen's input hypothesis, it is not sufficient for a learner to simply learn and practice 'rules', nor is it sufficient to hear language input that is not at all comprehensible to the learner, such as listening to a radio broadcast in a language that is totally foreign to the listener. There must be some link for the learner to comprehend the input. These links could be immediate visual and demonstrative links, such as when the speaker points to something or performs an action while speaking, or there may be already existing links in the form of vocabulary items in the input that are already familiar to the learner. In addition, according to Krashen, in order for acquisition to occur, the vocabulary and structure of the input must be at a slightly higher level than the current competence level of the learner. Thus, Krashen labels as $i + 1$ the appropriate level of input necessary for acquisition to take place (Krashen 1982; Krashen and Terrell 1983). When there is an appropriate level of comprehensible input, the learner infers the meaning of new vocabulary and structures through contextual clues and thus begins to make form-meaning connections for new lexical items or new structures. And, according to Bybee's model, the more frequently a given form is heard and comprehended in the input, the stronger the lexical connections become.

8.6. Paradigms in language acquisition.

Anderson (1995) says that it is "often argued that words differing only in their inflection group together into paradigms," but he adds that "this is of little help in

identifying ‘inflection’ unless we know when words should be said to belong to the same paradigm” (79). Speakers of a natural human language are not normally aware of the rules, patterns, and paradigms in their native language. According to Bloomfield (1933), psychologists maintain that normal speakers are not capable of the kind of reasoning implied by inflectional paradigms. If this is true, then “the normal speaker, who is not a linguist, does not describe his speech-habits, and, if we are foolish enough to ask him, fails utterly to make a correct formulation” (406). Bloomfield comments that linguists overestimate the ability of speakers to recognize their “speech-habits,” pointing out that linguists “forget that they owe this ability [to describe a language’s rules, patterns, and paradigms] to a sophisticated philosophical tradition” (406). Thus, according to Bloomfield, we must remember that the normal speaker of a language is unable to describe the patterns of his or her own language without specialized training. It would perhaps behoove language teachers, then, to be aware that verb paradigms or conjugations are a construct of linguists and language teachers and are not necessarily obvious or innately accessible to the average language student. As Bloomfield says, such paradigms and other statements about the ‘rules’ of a given language merely describe the action of native speakers of a language and do not “imply that the speaker himself could give a similar description” (406). Bybee (1985) defines “the lexical representation of a paradigm” as “a cluster of words bearing both a semantic and phonological relation to one another, in which one word (and sometimes more than one) is stronger than the others” (124), but this paradigmatic lexical organization does not imply a conscious awareness of paradigms as such.

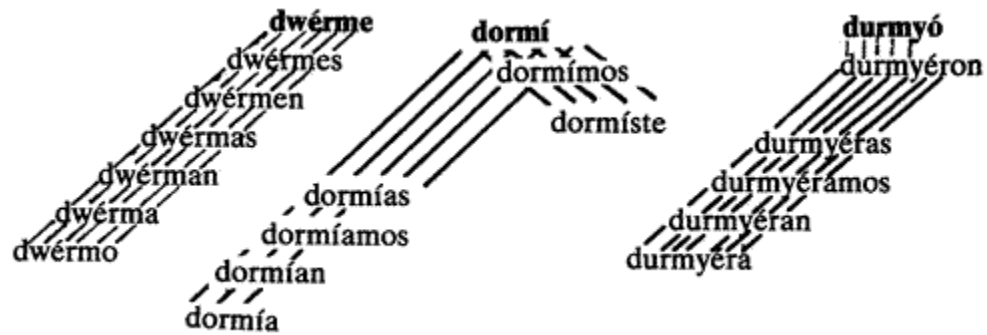
The analogical leveling and regularization processes evident in language change and variation do imply the existence of some type of paradigm in the speaker's consciousness (Aitchison 1991: 147-148; Bybee 1985: 51, 57, 64-65, 126; Wurzel 1984). The conventional verb organizational system based on tense/mood/aspect with six-member person/number-based sets for any given lexeme is pedagogically useful and is a convenient way of organizing verb forms for reference purposes, and is also consistent with Bybee and Pardo's hypothesis of a hierarchical relationship among the surface forms of verb paradigms. Contrary to the generativist notion of an abstract underlying representation for a verb paradigm, in which the surface forms "do not have any official relation to one another" (Bybee and Pardo 1981: 947), Bybee and Pardo found evidence in empirical investigations with speakers that there is psychological validity to such paradigms and that there are indeed relations among surface forms. These researchers found strong evidence that "the person/number forms within one tense or aspect are more closely related to one another than they are to forms outside that tense or aspect" (947). This finding is consistent with Bybee's (1985) contention that aspect and tense are more semantically related to the stem than are person and number (35-37).

However, traditional pedagogical verb paradigms, at best, represent only a small fragment of how verb forms are actually organized, processed, and stored in the mental grammar of a language. As we have seen earlier, there are numerous connections among forms in the lexicon, and these connections are multidimensional and multidirectional. There are connections among forms with the same basic semantic content (the same initial X variable in the present model), such as in all the conjugated forms for a given verb. But there are also connections – other paradigms, perhaps – based on semantic

similarity, such as the connections among all Preterite forms or all Subjunctive forms. Similarly, there are semantic connections among all forms that share the same person/number value, such as *hablas, cantas, escribes, escribas, vives, vivirías, fuiste, eras, supiste, sabías, hablarás, tuviste, cantabas, escribiras, fuisteis* (all 2 Singular forms, but of different tenses, moods, and aspects). In addition to purely semantic connections, of course, there are connections among forms that share semantic and phonological features, such as all 1 Singular Present Indicative forms that end in /o/, all 1 Plural forms that end in /mos/, or all of the above 2 Singular forms that end in /s/. As discussed earlier, this final /s/ connection among all 2 Singular forms, with the exception of 2 Singular Preterite, accounts for the analogical regularization of the 2 Singular Preterite in some dialects, as in **hicistes, *fuistes, *hablastes*.

Therefore, as we saw in Chapter 7 of the present analysis, the traditional six-member person/number-based paradigms grouped by tense, mood, and aspect (as in Figure 1.4, for example, or as in Kendrix' *501 Spanish Verbs*) are not the only possible way of organizing verb forms. Bybee's (1985: 124) graphic in Figure 3.7, for example, shown again in Figure 8.1 below, illustrates the connections among the various forms of the verb *dormir*, with these connections crossing tense and mood categories. Thus, we could organize the verb *dormir* (and other verbs like *dormir*) according to stem alternate instead of according to tense, mood, and aspect. Such a paradigm (as in Figure 3.7 and 8.1) would include /*dwerme/*, /*dwermes/*, /*dwermen/*, /*dwermas/* (all Present Indicative forms), /*dwerman/*, /*dwerma/* (both Present Subjunctive forms), /*dwermo/* (1 Sg Present Indicative). This paradigm includes all the forms that have the diphthongal stem alternant /*dwerm/*, even though these forms belong to different moods.

FIGURE 8.1. Overlapping semantic and phonological connections for some of the forms of the Spanish verb (*dormir* (to sleep) (Bybee 1985: 124)



A related paradigm group is the one containing all the forms that have the monophthongal stem variant, such as /*dormí*/, /*dormimos*/, /*dormiste*/, *dormías*/, /*dormíamos*/, /*dormían*/, /*dormía*/, although these forms belong to different aspects. This group would also contain the Infinitive and Past Participle forms, as well as the Future and Conditional forms. This third group, again crossing mood categories, has all the forms with the /*durm*/ stem alternate: /*durmyó*/, /*durmyeron*/ (both Preterite Indicative forms), /*durmyeras*/, /*durmyéramos*/, /*durmyeran*/, /*durmyera*/ (all Imperfect Subjunctive), and would also include /*durmyendo*/ (Present Participle), although this latter form is not shown in Bybee's graphic. All of the forms in this third group happen to have affixes beginning with *yod*, that is the glide /*y*/, discussed in Chapter 6 as a historical influence on 'raising' the stem vowel in *i*-class verbs.

8.7. Whole words, lexical access, and matching procedures.

According to Chialant and Caramazza (1995), “Recognition of previously encountered words is achieved through a matching procedure between the input stimulus and the lexical representation” (59). If a form in the input is identical to a form in the lexicon, then “the input form can be said to have been identified as a word belonging to our mental lexicon” (59). Chialant and Caramazza maintain that recognition of morphologically complex words is no different from recognition of morphologically simple words. Recognition of *encontrábamos* (‘meet, find’, 3 Plural Imperfect Indicative), for example, is achieved in essentially the same way as the recognition of *casa* ‘house’, that is, through “a matching procedure between the input and the lexical representation” (59). *Casa* may have more lexical strength than *encontrábamos* because of a higher frequency in the input, but the matching process is essentially the same. However, in addition to the whole word matching a lexical entry, the /*âβamos*/ segment of *encontrábamos* would map onto the identical segment in other words, such as *hablábamos*, *cantábamos*. This, in Bybee and Skousen’s terms, means establishing analogical connections among semantically and phonologically similar words or parts of words.

Bybee (1988) comments that the way speakers and linguists “discover that a word consists of more than one morpheme” is that “We find a relation of phonological and semantic identity or similarity between some subpart of the word and a subpart of another word” (127). For Bybee and others who hold to the whole-word hypothesis, even morphologically complex words are stored in the lexicon. When a new word is heard and subsequently learned, whether it is a morphologically complex or morphologically simple

form, “it forms connections with existing lexical material on the basis of its meaning and phonological shape. The word is not physically dismembered, but its parts are nonetheless identified” and are thus connected with parts of other words (127).

Some rules may seem independent of their corresponding representations, for example, when a rule is appropriately applied to a new form that has never been heard before by the speaker. However, Bybee contends that even forms such as these “have to emerge in acquisition from representations (1988: 124). For a learner to acquire rules, he or she “must extract them from the comparison of sets of related forms, which are entered in the mental lexicon” (124). “Rule-like generalizations gradually emerge from stored rote forms, which are initially processed and stored as unanalyzed wholes” (124). Even though a linguist might discover some pattern – and call it a feature-changing ‘rule’ – that alters a base form and thus produces a regular form or even a suppletive form, Bybee argues that such forms are generated by “lexical access rather than by feature-changing rule” (124). She contends that “what are usually thought of as morphological rules do not have a representation that is independent of the lexical items to which they are applicable. Rather, rules are highly reinforced representational patterns or schemas (135).

Regarding rules and rule-like behavior in language, Skousen (1989) comments that “For Chomsky (and virtually all other linguists today) there is no doubt that language is rule governed and that language behavior must be accounted for in terms of explicit rules,” adding that “As a corollary, language acquisition is viewed as learning rules and language change as a change in the rules” (4). According to Skousen, rules per se are more static, while rules as representations, and rules as analogical connections, are more dynamic and therefore are more explanatory of acquisitional development and change.

For Skousen and Bybee, rules as representations and analogical connections are also more explanatory of diachronic change and regional, situational, and individual variation. The important pedagogical implication here is not only that rules as patterns or representations may be more explanatory of acquisitional development, but also that these rules or patterns must be acquired primarily through extensive comprehensible input in meaningful contexts.

8.8. Input, interconnections, and ill-formed tokens.

The importance of input and analogical connections is significant to the present proposed model of Spanish verbal morphology. In first language acquisition, children typically learn some irregular forms before they seem to acquire the rules/patterns for the regular paradigms of their language (Bybee 1988: 136). At some point they begin to regularize the forms in their vocabulary, thus producing ill-formed tokens, or ‘incorrect’ forms, such as English **goed* (for the suppletive form *went*), English **brang* (for *brought*), Spanish **sabo* (for the suppletive form *sé*), or Spanish **andó* (for the allo-class preterite *anduvo*). Before there is sufficient input to make analogical connections and organize patterns in the mind, use of the irregular or suppletive forms seems to be just simple repetition. The child simply parrots what he or she has heard in similar semantic and situational contexts. However, after more exposure to the language around him or her, thus more comprehensible input, the child’s mind begins to develop analogical connections between forms, while at the same time continuing to build the feature sets associated with certain words and contexts. The child may even make analogical

connections and groupings that seem logical and are analogical to similar irregular forms, but that turn out to be erroneous, such as in the above **brang* for *brought*.

In addition to the dynamic nature of the lexicon and of morphological forms in child language acquisition, there is similar variation among individual speakers and among groups of speakers, both synchronically and diachronically. In fact, Bybee proposes a “gradient representation of lexical strength” (1988: 131) that accounts for this variation. Although the lexicon has in the past typically “been conceived of as the mental counterpart of a dictionary, a list of forms set down once and for all,” Bybee proposes

a more dynamic representation in which not all forms have the same status, but rather in which forms are affected by use or disuse. Frequently used forms gain in lexical strength and forms that are not used lose lexical strength. Lexical strength, then, is an index of word frequency (131).

This dynamic view of the lexicon is important and relevant in that it can “account for the psychological, historical, and cross-linguistic effect of frequency on morphology” (131).

Bybee’s observation about the dynamic nature of the lexicon and her hypothesis of “a gradient representation of lexical strength” (131) are relevant to adult language variation as well as the variation and idiosyncrasies of child language. If a speaker of Spanish or English, for instance, has to use a reference book to determine whether the Past Participle of Spanish *bendecir* is *bendecido*, **bendicho*, or *bendito*, or whether the Past Participle form of English *dive* is *dived* or *dove*, then what this means is that the speaker has not had sufficient input to make the analogical connections and groupings necessary for building feature sets and their associated word formation rules. Or it may be that the speaker has had conflicting input, so that he or she is unable to produce the

‘correct’ form with a high level of confidence, as in *sneaked* vs. **snuck*. If a speaker produces an ‘incorrect’ form, what this means, in the framework of the present analysis, is that the rule-and-feature sets in the speaker’s mental lexicon are different from those of the ‘standard’ dialect of his or her language.

Ruiz (1986) comments that even adult native speakers of Spanish may have difficulty knowing whether a given verb has an alternating stem, for instance. He says that native speakers learn such alternations “one by one; that is, diphthongizing stems are lexically marked to undergo the diphthongization rule” (114). This necessity of learning the alternations one by one points out the importance of input for the development of the rule-and-feature sets and semantic-phonological connections in the lexicon. Ruiz (1986) also discusses the instability of the verbal system in Chicano Spanish, pointing out that speakers of this dialect “are more prone to spontaneous creations at the discourse level” (115), resulting in modifications of the lexicon. He further points out that “the lexicon of languages in contact is susceptible to change before other subsystems” (115).

Alvar and Pottier (1983: 273) cite several examples for which analogy accounts for dialectal variation. According to these authors, some regional dialects of Spanish use /oron/ (as in **cantoron*, **echoron*) for 3 Plural Preterite Indicative instead of /aron/, because of analogy with 3 Singular Preterite Indicative /ó/. In addition, according to Alvar and Pottier, other dialects, analogically motivated by the /en/ in many verb forms (*echen*, *hablen*, *canten*, *cantasen*, *hablasen*), have 3 Plural Preterite Indicative ending in /en/, as in **cantoren*, **echoren*.

Albright (2000), discussing lexical gaps, refers to a ‘gradient effect’, similar to Bybee’s gradient representation of lexical strength. He maintains that “gappiness is a

gradient effect, with a whole array of reluctance between complete refusal to answer [when experiments try to elicit missing forms in so-called defective paradigms] (as in *asimos* → ___) and completely automatic responses (as for *cantamos* → *canto*)” (42). For pedagogical purposes, Albright concludes that it is probably useful for students to be made aware “that there are forms missing for a very few common verbs, but it is probably not worth the effort to specifically memorize gaps for many lower frequency verbs. Furthermore, he points out that grammar books and language experts do not always agree on what and where the lexical gaps are in a language.

Regarding words and the lexicon, whether standard or nonstandard, Caron (1992) points out that “A word evokes not only a concept but also the whole set of knowledge associated with that concept” (63). He adds that the meaning of a word is really just a set of implications and associations and is not necessarily limited, and that

the meaning of a word forms an open, undefined set whose content varies between different individuals and different times. This is no longer a ‘dictionary definition’ but an encyclopaedia article (moreover, this encyclopaedia is in a state of permanent revision; every copy is different) (63).

A word also evokes “a set of representations, emotions, attitudes, which are not conceptual in nature,” in other words, connotations, “which can be purely individual but can also be common to a group of individuals, however large or small.” Caron therefore maintains that it may be just an illusion that words have fixed meanings. Since the meaning of a given word is context dependent, Caron questions whether “the meaning of a word is simply what it is used for, the *use* to which it is put in a specific context” (63). This notion explains the metaphoric use of forms that would otherwise be ungrammatical

or semantically nonsensical, such as *El techo se llovió anoche* ('The roof rained last night'). Thus, according to Caron, "The meaning of a word is relative to the 'language-game' in which it is used" (63).

It is wise to keep in mind, then, that we cannot say that people don't say such-and-such because it is 'ungrammatical', but rather that it is ungrammatical because people don't say it. But that observation doesn't mean that someone – and eventually, perhaps, a whole group of people – couldn't start using some new form at any given moment, in any given place, for whatever reason. It's not that people don't use a given form because there's a rule that says they can't say it, but rather that the reason such a form is not a rule or pattern in a given language, dialect, or idiolect is simply that speakers of that language or dialect (or the speaker of that idiolect) do not currently use the given form or expression.

The present author once heard an argument among some linguists about the inherent 'ungrammaticality' of certain constructions. Someone argued, for instance, that it would always – in every circumstance, in any language – be ungrammatical to use a coordinating conjunction without a second item to be coordinated with, as in **I would like some ham and*. Others argued that such a restriction is arbitrary, language-specific, and dependent upon actual use. The person arguing against such universal constraints on grammaticality must have gotten a job as an inventor of product brand names, because there is currently a new advertised product called 'Cookies &'. As Pinker says, "we adults get to say what counts as 'correct', and if we regularize an irregular often enough, we simply declare by fiat that it is not an error!" (1999: 200).

When linguists try to determine the causes of linguistic change, there is a hidden assumption that stability is the norm and that change and variation are the exception. The present author proposes that this assumption may be inherited from Newtonian physics, which assumes clockwork predictability in the universe. In the seventeenth century Sir Isaac Newton changed the way Western civilization viewed the world. Before Newton, humans viewed the universe as chaotic, arbitrary, and incomprehensible. Then Newton devised a few laws that seemed to order the universe, “and suddenly the planets are seen to be moving in simple, predictable orbits! (Waldrop 1992: 327).

The metaphor of the age ... became the clockwork motion of the planets: a simple, regular, predictable Newtonian machine that would run of itself ...

Reductionist science tends to say. ‘Hey, the world out there is complicated and a mess – but look! Two or three laws reduce it all to an incredibly simple system!’ (328).

Linguists have traditionally applied this reductionist science to their own field, incessantly seeking unifying laws that will explain human language and its variation and change. Linguistics is concerned with “a psychologically internalized grammar which generates the utterances of speakers” (Weinrich, Lehmann, and Herzog 1968: 105).

Weinrich, Lehmann, and Herzog maintain, however, that “in a language serving a complex (i.e., real) community, it is *absence* of structured heterogeneity that would be dysfunctional” (101). Milroy (1992) questions why “different dialects remain divergent from ‘mainstream’ norms of language despite the low status usually accorded to them” (ix). He further comments that “Languages are not in reality completely stable or

uniform, and there is absolutely no reason they should be” (4). He assumes that variation and change are the norm and maintenance is the exception.

In the lexicon of the speaker who produces nonstandard forms, whether these are forms that are ‘missing’ in other people’s paradigms, or whether they are socially unacceptable forms, child-language forms, or situational aberrations, these nonstandard forms nevertheless reflect patterns of organization in the lexicon for that particular speaker of the language. That individual person’s patterns of organization may be different from the patterns of the standard form of the language or even from those of other speakers of his or her own dialect. It is highly unlikely that there are two speakers of any language who have exactly the same vocabulary, the same semantic feature sets, and the same collocations sets in their respective lexicons. Given the dynamic and creative nature of human language and the uniqueness of its speakers, it is also unrealistic to think that all speakers of a given language, even of a standard dialect, have identical syntax, phonetics, phonology, and morphology. And, in fact, there is even situational variation for an individual speaker, as Aitchison points out: “Almost all speakers of a language alter their speech to fit the casualness or formality of the occasion, though they are often unaware of doing so” (1991: 35). This variability is what gives each of us our own unique idiolect, and what collectively constitutes dialects.

8.9. Frequency of occurrence, regularity, and irregularity.

In the model of Spanish inflectional verbal morphology proposed in the present analysis, as explained and demonstrated earlier, the most specific rules apply before the more general ones. The more general rules, simply by virtue of being more general, are

likely to have in the input more examples of different ‘types’ – that is, exemplars of a given rule/representation with different X variables, as opposed to ‘tokens’, which are specific instances of a given form in the input. Bybee comments that “A large number of distinct verbs participating in the same pattern [i.e., a high degree of ‘type’ frequency] will serve to strengthen it [the pattern]” (1988: 138). The high frequency of types for a given rule or representation in the input serves to reinforce its connections with other forms, thus reinforcing and strengthening the form-meaning connection between the *signifié* and the *signifiant* of the rule’s variable and the connections among analogous forms. For example, affixes such as English /ed/ [Past] or Spanish stressed /ó/ [3 Sing Pret Indic] have a high degree of type frequency and thus serve to strengthen the given patterns. Thus, a regular rule or representation has more types than an irregular rule/representation, and thus is more amenable to application to newly invented words in a language or to nonce verbs presented to speakers in experimental research. If a speaker hears – or invents – a new verb, he or she can readily make analogical connections with the many existing exemplars in his/her lexicon and promptly produce new forms with the same affix. For example, a speaker could quickly produce **fobbled* from a nonce English verb *fobble* or **fabó* from a nonce Spanish verb **fabar*.

The model proposed in the present analysis is consistent with the observation that the most commonly used forms in any language are the ones most likely to be irregular or suppletive. Bybee maintains that lexical strength accounts for maintenance of irregularity and suppletion in high-frequency forms (1988: 132). Pinker concurs when he comments that “An irregular form that falls below a critical frequency could disappear outright after a few generations” (1999: 201). In Bybee’s view, lexical strength, or the lack thereof,

also accounts for the presence of lexical gaps, for the occasional occurrence of forms that are supposedly missing in certain paradigms, and for the variation in usage of such missing forms.

A highly irregular form that is not in common usage cannot not survive, because it is not present to a sufficient degree in the input of a sufficient number of speakers to allow links to be made between its feature set and the specific rule that represents the given irregular form. A very specific rule (such as a suppletive form) only has enough strength in its connections to survive if it is associated with a commonly recurring form. The very common English verb *be* has the highly irregular forms *am, is, are, was, were, been*, and the common verb *go* has the irregular forms *went* and *gone*. These same common verbs in Spanish also have highly irregular or suppletive forms, such as following forms for *ser* 'be': *eres, es, somos, sois, son* (all Present Indicative), *fui, fuiste, fue, fuimos, fuisteis, fueron* (all Preterite Indicative), *era, eras, era, éramos, erais, eran* (all Imperfect Indicative). The verb *ir* 'go' also has irregular or suppletive forms: *voy, vas, va, vamos, vais, van* (all Present Indicative), *fui, fuiste, fue, fuimos, fuisteis, fueron* (all Preterite Indicative).

The ultimate question in this very complex verbal system is how it is possible for a speaker to generate all of the 'irregular' forms that exist, in addition to all the ones that conform to regular patterns of organization. How is it possible for a speaker to know, much less to remember in every spoken utterance, for instance, that stem changing verbs only diphthongize the root vowel in stressed syllables, but that in certain unstressed syllables the monophthongal mid vowel in these verbs is raised to a high vowel? Or that the past participle of most verbs is formed by taking the regular stem plus the class vowel

and adding /do/, but that certain verbs do not follow this pattern and have entirely different forms, such as *dicho* or *hecho*? How is it possible for a speaker to mentally formulate all these rules in so little time and with so little input and then apply these rules with near perfect precision every time he or she speaks? And then how is he or she to know, in addition, all the minute exceptions to these rules and apply each of these more specific rules in only certain contexts? This is Plato's timeless paradox, and it is still the essential problem of linguistics.

In this rule-and-feature based model of morphology, the associations between the feature sets and the rules that produce the phonological material representing these semantic notions can only be made through consistent and sufficient exposure to the language. It is through this input that the connections are strengthened and thus made automatic, both the connections between the feature sets and the rules, and the analogical connections among similar phonological forms and similar feature sets. Therefore, if an 'incorrect' form is produced, that is, a form that does not conform to the recognized norms of usage of the language, it is because that particular speaker has formed associations among the rules and feature sets that are different from, or perhaps not as specific as, those of other speakers of the same language. But he or she is nevertheless speaking according to the patterns, rules, representations, and feature sets of his/her own dynamic mental model. The models that form in the mind of each individual are dynamic, nonstatic models that change and expand as more input is received and more feature sets are refined, and as the system becomes better organized and more simplified. Ironically, this continuous reorganization and simplification of the overall system take

place at the same time that the speaker's utterances are becoming increasingly more complex.

8.10. The problem: classroom learning vs. child language acquisition

If Krashen's (1982, 1995; Krashen and Terrell 1983) input hypothesis is valid, and if the above hypotheses regarding the strength of lexical connections is true, then there remains a grave problem for adults or adolescents learning a second language in a classroom setting. According to Asher's (1982) calculations, the average child by age six has heard oral input in his or her native language for 17,520 hours and has produced vocalizations perhaps an average of an hour a day during those six years, amounting to 2,190 hours of speaking. "In comparison, the student in the classroom in one year has listened to a foreign language for 320 hours [based on 32 weeks times five hours per week in class plus five hours a week listening to tapes] and had produced vocalizations for 27 hours, assuming the student talks 10 minutes per class meeting" (Asher 1982: 2). Asher then adds that "if we expect the student in the classroom to have the fluency of a six-year-old child, the student should listen to the foreign language for 55 years of college instruction and the student should have the opportunity to vocalize in the foreign language for 81 years of college instruction" (2).

This time comparison is impressive enough without considering that the actual time of a typical foreign language class at the university level at the beginning of the twenty-first century is only two and a half hours a week, only half the time allowed for in the above comparison. Furthermore, it is doubtful that many students listen to tapes for five hours a week outside of class. Therefore, the analogy of fifty-five years of listening

and eighty-one years of speaking is actually a gross underestimate of the comparison between university-level second language learning and a child's first language acquisition. The real mystery, then, as Asher points out, is that "so much is accomplished in a highly condensed period of time" (2).

8.11. Explicit grammar instruction.

Explicit grammar instruction may help adult or adolescent learners to be cognitively aware of verb paradigms, whether in their traditional groupings or in groupings according to their stem variants as above. It may be useful, at least for some adult or adolescent learners, to memorize verb paradigms for conscious use, especially when there has not been sufficient time for comprehensible input to support the building of connections and feature sets in the lexicon. Explicit grammar rules can, in this sense, be viewed as simply a shortcut to communication along the road to true acquisition of the language. Terrell (1986) comments that "knowledge of verb paradigms plays little or no role in beginners' comprehension – which in early stages depends mostly on the recognition of the verb stems," but he also adds that "Learning exercises may in some cases aid the acquisition process for many students" (225).

However, foreign language instructors should be aware of the difference between conscious use of rules or paradigms on the one hand and what Bloomfield (1933) calls unconscious 'speech-habits' on the other hand (406). Terrell (196) points out that "Learning is not acquisition, and it does not *become* acquisition without the necessary comprehensible input" (223). He adds that when a word or form is genuinely acquired, the word begins to "sound like" its meaning (223).

Bloomfield's distinction between conscious, cognitive use of paradigms vs. unconscious 'speech-habits' is analogous to Krashen's learning/acquisition distinction. According to Krashen's acquisition-learning hypothesis, "Language learning is 'knowing about' language, or 'formal knowledge' of a language." It is "explicit' knowledge of rules, being aware of them and being able to talk about them" (Krashen and Terrell 1983: 26-27). Language acquisition, on the other hand, is implicit knowledge and is a subconscious process, like that described above by Bloomfield. Language acquisition involves making and strengthening form-meaning connections, as well as building and refining feature sets, through extensive exposure to comprehensible input. Traditional paradigms and grammar rules may be consciously learned and may well serve the learner in producing forms that have not yet established sufficient lexical strength in their respective connections. However, it does not necessarily follow that these traditional paradigms are the only way verb systems are organized and processed in the mind of the speaker. If such paradigms do exist in the mental grammar, they are only one example of the kinds of lexical connections among verb forms that exist in the mental grammar of a language.

8.12. Input, output, and output as input.

In Bybee's model, "each time a word is heard and produced it leaves a slight trace on the lexicon," and it thus "increases in lexical strength" (1985: 117). In order for these connections to be made and for forms to be mapped onto other forms and larger units (such as similarly inflected forms or collocations in the larger context), there must be sufficient examples of the form ('tokens' or 'exemplars' of the form) in the input. And

for the input to be internalized, to be transformed into ‘intake’, the learner must be attentive to the input, and the input must be comprehensible. Krashen emphasizes the importance of comprehensible input, pointing out that most of the input must be comprehensible. Krashen’s notation of comprehensible input is represented as $i + 1$, “where i is the acquirer’s level of competence” and $i + 1$ is the next immediate stage in a continuum of language comprehension (Krashen and Terrell 1983: 32). Constant aural exposure to a language completely foreign to the listener, with no human contact and no contextual clues, would not constitute comprehensible input, since the listener would have no links to begin with and would have no visual or contextual cues upon which to build feature sets and connections among forms.

Lee and VanPatten (1995) argue that most of traditional grammar instruction and practice “is neither meaningful nor communicative” (93). They propose an input processing model of language acquisition (89-115), with the input structured in such a way as to force the student to focus on form in order to focus on meaning and communication. Ideally, the input is kept at a lexically familiar level so students can focus on form instead of having to focus on form and meaning at the same time. The input is structured in such a way that the student must focus on form in order to comprehend important aspects of the message, such as focusing on different verb endings in order to determine whether an event has already taken place, is currently taking place, or is to take place sometime in the future. Lee and VanPatten contend that with this approach, students’ production, as well as comprehension, is greatly improved, while with the traditional emphasis on grammar explanation and production, students only show gains in production. In some of their studies, the input processing groups actually

performed better on production tasks than did the production-oriented control groups. This emphasis on the processing of input could greatly enhance the establishment of form-meaning connections, and thus facilitate the building of the feature sets necessary for processing verbal morphology according to the model proposed in the present analysis.

Although this input processing model of language acquisition sounds promising for building feature sets and establishing and strengthening connections, it has several strategic problems. If the input is genuinely communicative, rather than simulated or contrived, and if the learner is genuinely motivated to comprehend the input, then the results might be reminiscent of Rousseau's *Emile* (1762). The young boy Emile is motivated to learn to read when he misses a party because he is unable to read the invitation. In the case of Emile, his tutor creates a real-life situation that motivates the boy to learn to read. Similarly, when people learn a second language in real-world settings, they often attend to form out of necessity to understand the message. The motivation to learn, as in the case of Emile or a second-language learner in a genuine communicative setting, is a powerful factor in language learning or acquisition. In a classroom situation, however, it is difficult, if not impossible at times, to create genuine communicative motivation. In addition, even if contrived input forces students to focus on form in order to comprehend the meaning, such input requires considerable effort on the part of the instructor. Nevertheless, the point is well made that if attention to form is necessary to understand the message, the resulting input will probably be more comprehensible and will thus increase and strengthen lexical connections.

In addition to the importance of input, Terrell (1991) also argues that a learner's own output can actually qualify as comprehensible input. One role of explicit grammar instruction in the classroom, then, is to provide a framework through which a learner can structure and produce correct output, which can then be processed as input and thus strengthen lexical connections. Terrell comments that "adults do not automatically use input to develop competence in the way Krashen has suggested" (Terrell 1991: 53). Terrell cites his own personal example of trying to learn Arabic while in Morocco. He says that since even the simplest input was incomprehensible to him, he successfully tried a conscious output strategy. Whenever he needed to say something, he would ask someone in French how to say such-and-such in Arabic. He would then repeat the phrase several times to himself and then try it out repeatedly, eventually acquiring some basic vocabulary and functionally useful phrases such as those necessary for purchasing or ordering food or drink. He says that what he acquired, however, "was my own version of these words and phrases based on my output, not on the input since I rarely heard these same phrases in the input directed to me" (53). Terrell concludes that explicit grammar instruction for adult second language learners can be valuable, especially if it is true that adult learners do not process input the same way children do. For adults, then, conscious knowledge of regular grammatical patterns/rules in the language and traditional verb paradigms might increase the comprehensible input in the form of the learner's own output, as it did for Terrell in Morocco.

The present author's own personal experience is also testimony to Terrell's notion of output as input and to the value of conscious knowledge of grammatical 'rules'. After studying Spanish at the university level many years ago, I did not have occasion to hear

or speak Spanish for about twenty years. However, during those twenty years with virtually no exposure to Spanish, I had a repertoire of Spanish songs that I sang to my children. I also frequently talked to myself or to my young children in simple, contrived phrases constructed from the vocabulary and grammar that I had learned in high school and college Spanish, and I frequently used the grammatical structures in the songs as models for constructing new phrases. Then, after many years of no exposure to Spanish and no use of the language other than this self-talk and memorized repetition of song lyrics, I was amazed to discover that I could still understand spoken Spanish and was quickly able to acquire (or re-acquire) spoken fluency.

One important factor in the continued reinforcement of Spanish structures and the strengthening of lexical connections for me, I believe, was that the songs that I had memorized and regularly sang to my children contained examples of nearly every important grammar point taught in beginning and intermediate Spanish courses. In addition, since I had studied all those grammar points exemplified in the songs, conscious knowledge of at least some of the rules added another dimension to the strengthening of lexical connections. I was aware of the meaning and make-up of those forms and was able to mentally analyze them and use them as reference models in creating my own self-talk. The simple repetition of forms through the medium of song lyrics, along with the recombination of the forms into new phrases in self-talk, strengthened the form-meaning connections and the analogical connections among similar forms through frequency of tokens. So, according to Bybee's (1985, 1988) hypothesis regarding lexical strength related to lexical frequency, and according to Terrell's (1991) notion of output as input, I was repeatedly exposed to exemplars of numerous morphologically complex forms, even

though I was not in a Spanish-speaking environment. One might suspect, then, that although neither rote memorization nor explicit grammar instruction per se may lead to true acquisition of a second language, some conscious knowledge of rules and some rote memorization may be very useful to the learner. The ability to analyze structures and to analogically construct new combinations of forms and structures may encourage an output-as-input strategy, thereby strengthening lexical connections and increasing the speed and efficiency of processing and producing verb forms and overall acquisition of the language.

8.13. Rules, systematicity, and variation.

Pinker (1999) comments that “the mind, like any complex device, is a system of mechanisms optimized for different jobs” (146). He adds that

Any theory that has one mechanism doing all the work is proposing a kind of crippleware that the human brain is bound to outperform.... If the mechanism is a set of rules, it loses the advantage of cacheing the results of frequently performed computations so that it can look them up quickly rather than recomputing them every time. If the mechanism is a set of associations, it loses the advantage of variables and the rules that combine them” (146).

The present analysis recognizes the unfathomable complexity of the human mind – and of the language that it produces – and brings “order and coherence” to part of that domain. The proposed model reconciles Anderson’s rule-based model (1986, 1988a, 1988b, 1995) with Bybee’s (1985, 1988) analogy-based model. The model allows for “cacheing the results of frequently performed computations,” although it sees those

‘computations’ as gradient analogical connections in a dynamic, ever-changing lexicon, rather than as part of a fixed “modularized set of rule-symbol systems” (Chandler 1993: 594) acting outside the lexicon. Like a true rule-based model, the proposed rule-and-feature based analogical model does recognize “variables and the rules that combine them.” The rules in the proposed model, however, are descriptions of the lexical forms. The rules are the representations themselves, rather than independent mechanisms that produce the representations. In the proposed model, “The usage is the description,” rather than usage being “a function of the description,” as in an independent-rules approach (Derwing and Skousen, 1989a: 63).

The present analysis and proposed model allow for the observable fact that one of the rules of language “is that the rules can be broken, which indicates a fundamental difference between the rules of grammar and the laws of science” (Bonnycastle 1997: 91). As Bonnycastle points out, “People break the rules of grammar all the time.... It simply means that the speaker is not a member of the social group that habitually uses correct English [for example]. Members of that group can, if they want, break the rules to achieve a desired effect, without changing the rules” (90). However, diachronic language change and synchronic language variation clearly demonstrate that if enough people in the right social class break the rules enough times, the rules can indeed change, and often do. In fact, one can justifiably argue, as Derwing and Skousen (1989a) do, that in human language “speaker variability is the rule rather than the exception” (64). The present analysis takes that variability into account, recognizing the dynamic nature of the lexicon and of language itself and rejecting the strict systematicity of a true rule-based approach. Language in general, and morphology and the lexicon in particular, are “a

flexible, open and ill-defined thing, with analogy... seen as the primary creative mechanism” (Derwing and Skousen 1989a: 63).

8.14. Conclusions: Coherence, connections, and complexity

Anderson regards (1988b) “a ‘theory’ of a given domain [a given component of the mental grammar of a language] as a framework of assumptions and propositions within which it is possible to bring order and coherence to particular facts within that domain” (189). The present analysis of the Spanish verbal system, based in form on Anderson’s Extended Word-and-Paradigm model of morphology, but in theory based primarily on Bybee’s analogical model of morphology, does indeed bring descriptive order and coherence to the Spanish verb system and thus is capable of contributing to a better understanding of the mental organization and processing of such a complex system.

The theories and morphology models of Anderson, Bybee, and Marslen- Wilson all provide a formal and theoretical basis for the proposed model of Spanish inflectional verbal morphology, although some aspects of these theories may seem contradictory. Although Anderson’s model provides the formal basis for the model, and his notion of an a-morphous morphology is essentially consistent with the proposed model, the proposed model, unlike Anderson’s, is not a generativist model. The proposed model also parts with Anderson in that the present model considers inflectional morphology to be based in the lexicon, regarding the lexicon as a dynamic, multidimensional, interconnected component of the grammar, as does Bybee.

The proposed model appears to be rule-based, since it is indeed a collection of Word-Formation Rules as described by Anderson, but we have seen that the proposed model views rules differently than does a generativist model. The rules in the model, as in Bybee's model, are descriptions of patterns of lexical organization and connections, and these lexical rules are seen as indistinct from lexical representations. Hence, representations and rules, or items and processes, are just opposite ends of a lexical continuum. On this continuum, lexical strength is inversely proportionate to regularity of form, since the highly lexicalized suppletive forms at one end of the continuum are dependent upon high frequency of use for their survival. The more regular representations/rules at the other end of the continuum are less dependent upon frequency, since their regularity itself makes their forms more predictable through analogical connections with similar forms. The rules in the proposed model that are most specific, those at the 'suppletive' end of the continuum, are the rules that must apply first for any given form, before application of the more general rules at the other end of the continuum. The application of these 'most specific' rules, or even their very existence for a given speaker, is dependent upon the lexical strength of those rules, and that lexical strength is dependent upon sufficient input to establish strong connections between the form and its function.

The model proposed in the present analysis has strong implications for how languages are learned and how they might be effectively taught. Simply memorizing grammatical rules and traditional verb paradigms may not be the best way to establish form-meaning connections, or to bind meaning to form and thus make the forms more accessible in comprehension and production. In order for form-meaning connections to

be made and for feature sets to be built, and for analogical connections among forms to be established and strengthened, students need to be exposed to a rather significant amount of comprehensible input through a combination of listening and reading. However, we have also seen how conscious knowledge of rules and patterns can be useful to a limited degree in building the lexical connections necessary for language acquisition and, possibly, in helping to produce output as input. The proposed model and its implicit emphasis on a large amount of meaningful input support VanPatten's Input Processing model (Lee and VanPatten 1995), Krashen's input hypothesis (1982), Krashen's emphasis on leisure reading for comprehensible input at the intermediate level (1995), as well as Communicative Language Teaching's emphasis on oral input at all levels. In addition, the model does support some degree of explicit grammar instruction as a shortcut in establishing form-meaning connections through increased awareness of form and as an aid in consciously constructing output that can then become input.

The Spanish verb model proposed in the present analysis is a product of, and hybrid of, several models and theories. The proposed model's Word-Formation Rules, although formally based on Anderson's generativist morphology model, are really just a different conceptual illustration of Bybee's analogical model of morphology. In addition, the proposed model supports Marslen-Wilson's cohort model of lexical processing as well.

We can conclude by extending Pinker's comment about language organization and processing to the realm of second language acquisition and teaching: "Any theory that has one mechanism doing all the work is proposing a kind of crippleware that the human brain is bound to outperform" (1999: 146). The proposed model strongly supports

an instructional approach that emphasizes providing a significant amount of comprehensible input, along with some degree of explicit grammar instruction. Given the unfathomable capacity and complexity of the human brain, the present analysis also supports an “informed eclecticism” (Richards and Rogers 1986: 158) in instructional philosophies and strategies, rejecting strict adherence to any one technique or teaching style.

APPENDIX
VERB STEM VARIANTS
AND FEATURE SETS

This appendix lists the lexical stem sets ('allomorphs' in traditional terminology) for a sampling of Spanish verbs, along with their respective feature sets. Special attention is given to exceptional/regular and idiosyncratic/suppletive verbs. It should be noted that since the proposed model is presented in an 'a-morphous' framework (as in Anderson 1995), these 'lexical stems' are not really 'allomorphs', nor are they considered true morphemes and may not in all cases correspond to what are traditionally considered 'stems'.

For ease of recognition by the reader, conventional orthography is used in this verb listing rather than phonemic transcription. However, since the model deals primarily with the spoken language, phonemic transcriptions are used in the model itself. For this reason, orthographic variation is not listed in the stem variants, since it is assumed that regular orthographic conventions will apply in the written form of the verbs.

The following spelling changes apply as indicated in all written forms of Spanish verbs, primarily for the purpose keeping the sound of the consonant consistent, plus an arbitrary spelling rule regarding <z> and <c>.:

1. <c> → <qu> / __ e (as in *indica* → *indique*)
2. <g> → <gu> / __ e (as in *llega* → *lleque*)
3. <gu> → <g> / __ o (as in *sigue* → *sig*)
4. <g> → <j> / __ a (as in *dirige* → *dirija*)
/ __ o (as in *escoge* → *escojo*)
5. <c> → <z> / __ o (as in *hice* → *hizo*)
6. Regular written accent marks do not apply to one-syllable verb forms, as in *vi*, *vio*, *di*, *dio* (Preterite forms of *ver* and *dar*).
7. For some one-syllable verb forms, written accent marks are used to distinguish homographic forms, as in *dé* ('give', 1 Sg Pres Subj) vs. *de* (Preposition).
8. <i> / [- Stress] → <y> / V__V (between two vowels), as in **leió* → *leyó*

Note: Although the verbs below are grouped together to show the common types of stem variation, any given verb may actually exhibit more than one type of irregular behavior, such as in *venir*, which has diphthong/mid-vowel alternation, mid-vowel raising, velar insertion in the stem, and, in addition, is in the class designated here as *allo*-class based on regularities within the irregularity, hence part of the exceptional/regular group.

Categorical/regular (most verbs)

INFINITIVE/VOWEL CLASS	DEFAULT STEM (INFIN/3 SG PRES INDIC STEM)
------------------------	--

Beber /e-class	beb
Cantar /a-class	cant
Hablar /a-class	habl
Comer /e-class	com
Vivir /i-class	viv

Phonologically variant: Diphthong/mid vowel alternation only

INFINITIVE/ VOWEL CLASS	STEM VARIANTS/ FEATURE SETS
----------------------------	--------------------------------

	SYLLABLE [+ STRESS].	SYLLABLE [- STRESS].
Cerrar /a	cierr	cerr
Contar /a	cuent	cont
Despertar /a	despiert	despert
Doler /e	duel	dol
Encender /e	enciend	encend
Encontrar /a	encuentr	encontr
Fregar /a	frieg	freg
Jugar /a	jueg	jug
Llover /e	lluev	llov
Oler /e	huel	ol
Pensar /a	piens	pens
Perder /e	pierd	perd
Probar /a	prueb	prob
Volver /e	vuelv	volv

Phonologically variant: Mid-vowel raising only

INFINITIVE/ VOWEL CLASS	STEM VARIANTS/ FEATURE SETS
----------------------------	--------------------------------

	{ SYLLABLE [+ STRESS] } { /y/ IN FOLLOWING SYLLABLE }	DEFAULT/'ELSEWHERE' STEM
Elegir /i	elig	eleg
Gemir /i	gim	gem
Medir /i	mid	med
Pedir /i	pid	ped
Repetir /i	repit	repet

Exceptional/regular: Allo-class

INFINITIVE/ VOWEL CLASS	STEM VARIANTS / FEATURE SETS					
Andar / a	anduv / {[Pret Indic] [Imperf Subj]}	and / elsewhere				
Estar / a	estuv / {[PretIndic] [ImperfSubj]}	est / elsewhere	estoy / [1 Sg Pres Indic]	Ending-stress / [Pres]		
Poner / e	pus / {[Pret Indic] [ImperfSubj]}	pon / elsewhere				pondr / {[Fut][Cond]}
Hacer / e	hic / {[PretIndic] [ImperfSubj]}	hac / elsewhere				har / {[Fut][Cond]}
Haber / e	hub / {[PretIndic] [ImperfSubj]}	hab / elsewhere	he / [1 Pres Indic]	ha / [{2Sg,3Sg,3Pl} Pres Indic]	hay / [Pres Subj]	habr / {[Fut][Cond]}
Saber / e	sup / {[PretIndic] [ImperfSubj]}	sab / elsewhere	sé / [1 Sg Pres Indic]			sabr / {[Fut][Cond]}
Venir / i	vin / {[PretIndic] [ImperfSubj]}	ven / elsewhere	veng / {[1 Sg Pres Indic] [Pres Subj]}	[Pres /Syllable [+ stress]]		vendr {[Fut] [Cond]}

Exceptional/regular: j-allo-class

INFINITIVE/ VOWEL CLASS	STEM VARIANTS / FEATURE SETS				
Decir / i	dij {[Pret] [Imperf Subj]}	dec / elsewhere	dig / {[1 Sg Pres Indic][Pres Subj]}	dic / [Pres / syllable + stress]	dir / {[Fut] [Cond]}
Conducir/i	conduj {[Pret] [Imperf Subj]}	conduc elsewhere	conduzc {[1 Sg Pres Indic] [Pres Subj]}		
Producir /i	produj {[Pret] [Imperf Subj]}	produc elsewhere	produzc {[1 Sg Pres Indic] [Pres Subj]}		
Traer / e	traj {[Pret] [Imperf Subj]}	tra elsewhere	traig {[1 Sg Pres Indic] [Pres Subj]}		

Phonologically variant: Velar insertion

INFINITIVE/ VOWEL CLASS	STEM VARIANTS / FEATURE SETS				
Caer / e	caig / {[1 Sg Pres Indic] [Pres Subj]}	ca / elsewhere			
Hacer / e	hag / {[1 Sg Pres Indic] [Pres Subj]}	hac / elsewhere	hic / {[Pret Indic] [Imperf Subj]}	har / {[Fut] [Cond]}	
Poner / e	pong / {[1 Sg Pres Indic] [Pres Subj]}	pon / elsewhere			
Salir / i	salg / {[1 Sg Pres Indic] [Pres Subj]}	sal / elsewhere			
Traer / e	traig / {[1 Sg Pres Indic] [Pres Subj]}	tra / elsewhere			
Venir / i	veng / {[1 Sg Pres Indic] [Pres Subj]}	ven / elsewhere	vin / {[PretIndic] [ImperfSubj]}	vendr {[Fut] [Cond]}	vien / [Pres /Syllable [+ stress]]

Idiosyncratic/suppletive

INFINITIVE/ VOWEL CLASS	STEM VARIANTS / FEATURE SETS									
Ir / I- class; a-class / [Imperf Subj]	voy / [1Sg PresIndic]	va / [Pres Indic]	fue / {[3Pret Indic] [Imperf Subj]}	vay / [Pres Subj]	fu / [Pret Indic]					i / elsewhere
Ser / e	soy / [1 Sg Pres Indic]	se / [Pres Subj]	ere / [2 Sg Pres Indic]	es / [3Sg Pres Indic]	fu / [Pret Indic]	so / [Pres Indic]	era / [Imperf Indic]	fue / {[3Pret Indic] [Imperf Subj]}	s / elsewhere	
Haber/e	hub / {[PretIndic] [ImperfSubj]}		he / [1 Pres Indic]	ha / [{2Sg,3Sg,3Pl} Pres Indic]	hay / [Pres Subj]	habr / {[Fut] [Cond]}				hab / elsewhere
Dar / a- class; e- class /[Pret] [Imperf Subj]	doy / [1 Sg Pres Indic]									d / elsewhere

REFERENCES

- Aitchison, Jean. 1991. *Language change: progress or decay?* Cambridge: Cambridge University Press.
- Albright, Adam. 2000. The acquisition of basehood in paradigms: An overview.
Available: <http://www.humnet.ucla.edu/humnet/linguistics/people/grads/aalbrigh/papers.html> [2000, Dec. 6].
- Alvar, Manuel, and Bernard Pottier. 1983. *Morfología histórica del español*. Madrid: Biblioteca Románica Hispánica, Editorial Gredos.
- Anderson, Stephen R. 1985. Inflectional morphology. [In: Shopen, T. (ed.). *Language typology and syntactic description*, v III, *Grammatical categories and the lexicon*. Cambridge: University Press. 150-201.]
- Anderson, Stephen R. 1986. Disjunctive ordering in inflectional morphology. *Natural Language and Linguistic Theory* 4. 1-32.
- Anderson, Stephen R. 1988a. Inflection. [In: Hammond, Michael, and Michael Noonan (eds.). *Theoretical Morphology*. San Diego: Academic Press. 23-43.]
- Anderson, Stephen R. 1988b. "Morphological theory." [In: Newmeyer (ed.). *Linguistics: The Cambridge Survey*, v I, *Linguistic Theory: Foundations*. Cambridge: Cambridge University Press. 146-191.]
- Anderson, Stephen R. 1995. *A-Morphous Morphology*. Cambridge: University Press.
- Anshen, Frank, and Mark Aronoff. 1988. Producing morphologically complex words. *Linguistics* 26: 641-655.
- Aronoff, Mark. 1976. *Word Formation in Generative Grammar*. Cambridge, MA: MIT Press.

- Aronoff, Mark. 1992. "Stems in Latin verbal morphology." [In: Aronoff, Mark (ed.). *Morphology Now*. Albany: SUNY Press. 5-32.]
- Asher, James J. 2000. *Learning Another Language Through Actions: The Complete Teacher's Guide*. 6th ed. Los Gatos, CA: Sky Oak Productions.
- Azevedo, Milton M. 1992. *Introducción a la lingüística española*. Englewood Cliffs, NJ: Prentice Hall.
- Beard, Robert. 1995. *Lexeme-Morpheme Base Morphology: A general theory of inflection and word formation*. Albany: SUNY Press.
- Beyer, Thomas R, Jr. 1998. *501 English Verbs*. Hauppauge, NY: Barron's.
- Bloomfield, Leonard. 1933. *Language*. New York: Henry Holt and Company.
- Brown, H. Douglas. 1994. *Principles of Language Learning and Teaching*. New York: Prentice Hall.
- Bull, William. 1965. *Spanish for Teachers*. New York: Ronald Press Co.
- Butt, J. 1997. *Spanish Verbs*. Oxford: Oxford University Press.
- Bybee, Joan. 1985. *Morphology: a Study of the Relation between Meaning and Form*. Amsterdam: Benjamins.
- Bybee, Joan. 1988. Morphology as lexical organization. [In: Hammond, Michael, and Michael Noonan (eds.). *Theoretical Morphology*. San Diego: Academic Press. 119-142.]
- Bybee, Joan, and Elly Pardo. 1981. On lexical and morphological conditioning of alternations: a nonce-probe experiment with Spanish verbs. *Linguistics* 19: 937-968.

- Cadierno, Teresa. 1995. Formal instruction from a processing perspective: An investigation into the Spanish past tense. *The Modern Language Journal*, 79, ii: 179- 193.
- Caron, Jean. 1992. *An Introduction to Psycholinguistics*. Toronto: University of Toronto Press.
- Carstairs-McCarthy, Andrew. 1987. *Allomorphy in Inflexion*. Beckenham: Croom Helm.
- Celce-Murcia, Marianne. 1992. Formal grammar instruction: An educator comments. *TESOL Quarterly* 26.2: 406-409.
- Chandler, Steve. 1993. Are rules and modules really necessary for explaining language? *Journal of Psycholinguistic Research* 22: 593-606.
- Chandler, Steve. 1994. An exemplar-based approach to language acquisition. Paper presented at the Workshop on Cognitive Models of Language Acquisition, University of Tilburg, The Netherlands, 21-23 April 1994.
- Chialant, Doriana, and Alfonso Caramazza. 1995. Where is morphology and how is it processed? The case of written word recognition. [In: Feldman, Laurie Beth (ed.). *Morphological Aspects of Language Processing*. Hillsdale, NJ: Lawrence Erlbaum Associates: 55-75.]
- Chomsky, Noam. 1972. *Language and Mind*. New York: Harcourt Brace Jovanovich.
- Chomsky, Noam. 1980. *Rules and Representations*. New York: Columbia University Press.
- Chomsky, Noam. 1986. *Knowledge of Language*. New York: Praeger.
- Chomsky, Noam. 1993. *Language and Thought*. Wakefield, RI: Moyer Bell.

- Cirre, José F., and Manuela M. Cirre. 1981. *España y los españoles*. New York: Holt, Rinehart and Winston.
- Cook, Vivian. 1993. *Linguistics and Second Language Acquisition*. New York: St. Martin's.
- Cook, Vivian, and Mark Newson. 1996. *Chomsky's Universal Grammar: An Introduction*. Oxford, UK: Blackwell.
- Damasio, Antonio R. 2002. Remembering when. *Scientific American*. September 2002: 66-73.
- Da Silva, Zenia Sacks, and Gabriel H. Lovett. 1965. *A concept approach to Spanish*. New York: Harper & Row.
- Davies, Paul. 2002. That mysterious flow. *Scientific American*. September 2002: 40-47.
- Derwing, Bruce L., and Royal Skousen. 1989a. Morphology in the mental lexicon: A new look at analogy. *Morphology Yearbook 2*: 55-71.
- Derwing, Bruce L., and Royal Skousen. 1989b. Real-time morphology: Symbolic rules or analogical networks? [In: Hall, Kira, Michael Meacham, and Richard Shapiro (eds.). *Proceedings of the Fifteenth Annual Meeting of the Berkeley Linguistics Society*. Berkeley: Berkeley Linguistics Society. 48-62.]
- Dorwick, Thalia, and Anna Maria Girones. 1995. *¿Qué tal? An Introductory Course*. Boston: McGraw Hill.
- Dozier, Eleanor, and Zulma Iguina. 1999. *Manual de Gramática*. Boston: Heinle and Heinle.
- Eddington, David. 1995. The psychological relevance of phonological generalizations in Spanish: An experiment. *Hispania* 78.4: 875-884.

- Eddington, David. 1996a. The psychological status of phonological analyses. *Linguistica* XXXVI: 17-36.
- Eddington, David. 1996b. The role of experiments in phonological investigation. *Journal of Quantitative Linguistics* 1999, 6: 14-28:
- Eddington, David. 2000. Analogy and the dual-route model of morphology. *Lingua* 2000, 110: 281-298.
- Eddington, David. 2000. Spanish gender assignment in an analogical framework. *Journal of Quantitative Linguistics* 2000, 9: 949-975.
- Emmorey, Karen and Victoria Fromkin. 1988. The mental lexicon. [In: Newmeyer, Frederick (ed.). *Linguistics: The Cambridge Survey*, Vol. III: *Language: Psychological and Biological Aspects*. Cambridge: Cambridge University Press. 124-149.]
- Ezzell, Carol. Clocking cultures. *Scientific American*. September 2002: 74-75.
- Faulkner, William. 1948. *Intruder in the Dust*. New York: Vintage Books/Random House: 190.
- Fayer, Mischa H. 1959. *Basic Russian*. New York: Pitman Publishing Corporation.
- Garrett, Merrill. 1988. Processes in language production. [In: Newmeyer, Frederick (ed.). *Linguistics: The Cambridge Survey*, Vol. III: *Language: Psychological and Biological Aspects*. Cambridge: Cambridge University Press. 69-96.]
- Hammond, Michael, and Michael Noonan (eds). 1988. *Theoretical Morphology*. San Diego: Academic Press.

- Harris, John. Aspects of Spanish verb morphology. [In: Hagiwara, Michio Peter (ed.).
Studies in Romance Linguistics. Rowley, MA: Newbury. 44-60.]
- Hashemipour, Peggy, Ricardo Maldonado, and Margaret van Naerssen (eds.). 1995.
*Studies in Language Learning and Spanish Linguistics in honor of Tracy D.
Terrell*. New York: McGraw-Hill.
- Hayakawa, S. I. 1949. *Language in Thought and Action*. New York: Harcourt, Brace and
Company.
- Hockett, Charles. 1954. Two models of grammatical description. *Word* 10: 210-234.
- Hooper, Joan. 1976. *An Introduction to Natural Generative Phonology*. New York:
Academic Press.
- Horwitz, Elaine K. 1986. Some language acquisition principles and their implications for
second language teaching. *Hispania* 69, 3: 684- 689.
- Jackendoff, Ray S. 1972. *Semantic Interpretation in Generative Grammar*. Cambridge,
MA: MIT Press.
- Jannach, Hubert. 1980. *German for Reading Knowledge*. Boston: Heinle &
Heinle.
- Joos, Martin. (ed.). 1958. *Readings in Linguistics* (2nd ed). Chicago: University of
Chicago Press.
- Kendris, Christopher. 1990. *501 Spanish Verbs*. Hauppauge, NY: Barron's.
- Kagan, Olga and Frank J. Miller. 1996. *V Puti: Russian grammar in context*. Upper
Saddle River, NJ: Prentice Hall.

- Kiparsky, Paul. 1973. 'Elsewhere' in phonology. [In: Anderson, Stephen R., and Paul Kiparsky (eds.). *A Festschrift for Morris Halle*. New York: Holt, Rinehart and Winston.
- Krashen, Stephen. 1982. *Principles and Practice in Second Language Acquisition*. New York: Pergamon Press.
- Krashen, Stephen. 1992. Formal grammar instruction: Another educator comments. *TESOL Quarterly* 26.2: 409-411.
- Krashen, Stephen. 1995. What is intermediate natural approach? [In: Hashemipour, Peggy, Ricardo Maldonado, and Margaret van Naerson (eds). *Studies in Language Learning and Spanish Linguistics in honor of Tracy D. Terrell*. New York: McGraw-Hill. 92-105.]
- Krashen, Stephen, and Tracy Terrell. 1983. *The Natural Approach*. Oxford: Pergamon Press.
- Kretschmar, William A., Jr. 1996. Dialectology and sociolinguistics: Same coin, different currency. *Language Sciences*, Vol. 17, No. 3: 271-282.
- Lathrop, Thomas A. 1980. *The Evolution of Spanish: An Introductory Historical Grammar*. Newark, DE: Juan de la Cuesta.
- Lee, James F., and Bill VanPatten. 1995. *Making Communicative Language Teaching Happen*. New York: McGraw Hill.
- LEXESP. 2002. [Online.] Available: <http://www.uniovi.es/UniOvi/Apartados/Departamento/Psicologia/metodos/soft/corpus/>

- Lightbown, Patsy M., and Nina Spada. 1990. Focus on form and corrective feedback in communicative language teaching: Effects on second language learning. *Studies in Second Language Acquisition* 12.4: 429-48.
- Lightbown, Patsy, and Nina Spada. *How Languages Are Learned*. Oxford: Oxford University Press.
- Lipski, John. 1985. /s/ in Central American Spanish. *Hispania* 68, March 1985: 143-149.
- Lucy, John Arthur. 1996. *Grammatical categories and cognition: A case study of the linguistic relativity hypothesis*. Cambridge: Cambridge University Press.
- Mackenzie, Ian. 2001. *A Linguistic Introduction to Spanish*. Muenchen: Lincom Europa.
- Manley, Joan H., Stuart N. Smith, John T. McMinn, and Marc Prévost. 1994. *Qu'est-ce qu'on dit?* Boston: Heinle and Heinle.
- Marslen-Wilson, William. 1987. Functional parallelism in spoken word-recognition. *Cognition*, 25: 71-102.
- Marslen-Wilson, William. 1989. Access and integration: Projecting sound onto meaning. [In: Marslen-Wilson, William (ed.). *Lexical Representation and Process*. Cambridge, MA: MIT Press. 3-24.]
- Marslen-Wilson, William. 1999. Abstractness and combination: The morphemic lexicon. [In: Garrod, Simon and Martin J Pickering (eds.). *Language Processing*. East Sussex, UK: Psychology Press, Ltd: 101-119.]
- Mateo, Francis and Antonio J. Rojo Sastre. 1999. *Langenscheidt's Pocket Bescherelle Spanish Verbs*. New York: Langenscheidt.
- McCarthy, John J. 1979. *Formal Problems in Semitic Phonology and Morphology*. PhD dissertation, MIT. [Distributed by IULC; published by Garland, NY, 1982.]

- McClelland, James L., and David E. Rumelhart (eds). 1986. *Parallel distributed processing: Explorations in the microstructure of cognition*. Cambridge MA, MIT Press/Bradford Books.
- McLaughlin, Barry. 1987. *Theories of Second Language Learning*. London: Edward Arnold.
- Merriam Co., G & C. 1965. *Webster's Seventh New Collegiate Dictionary*. Springfield, MA.
- Milroy, James. 1992. *Linguistic Variation and Change: On the Historical Sociolinguistics of English*. Oxford: Blackwell.
- Mitchell, Rosamond, and Florence Myles. 1998. *Second Language Acquisition Theories*. London: Arnold.
- Moreland, Floyd, and Rita M. Fleischer. 1977. *Latin: An Intensive Course*. Berkeley: University of California Press.
- Musser, George. A hole at the heart of physics. *Scientific American*. September 2002; 48-49.
- Newmeyer, Frederick J. 1983. *Grammatical Theory*. Chicago: University of Chicago Press.
- Newmeyer, Frederick J. 1986. *The Politics of Linguistics*. Chicago: University of Chicago Press.
- Newmeyer, Frederick J. 1998. *Language Form and Language Function*. Cambridge, MA: MIT Press.
- Newmeyer, Frederick J. (ed). 1988. *Linguistics: The Cambridge Survey, v I, Linguistic Theory: Foundations*. Cambridge: Cambridge University Press.

- Newmeyer, Frederick J. (ed). 1988. *Linguistics: The Cambridge Survey*, v III, *Language: Psychological and Biological Aspects*. Cambridge: Cambridge University Press.
- Pinker, Stephen. 1989. Language acquisition. *Foundations of Cognitive Science*, Cambridge, MA: MIT Press: 359-366.
- Pinker, Stephen. 1997. *How the Mind Work*. New York: W. W. Norton.
- Pinker, Stephen. 1999. *Words and Rules*. New York: Basic Books.
- Pinker, Stephen, and Alan Prince. *On language and connectionism: Analysis of a parallel distributed processing model of language acquisition*. Cambridge, MA: The Center for Cognitive Science, MIT.
- Rouse, W.H.D. (trans.). 1956. *Great Dialogues of Plato*. New York: Mentor Books, New American Library.
- Pyles, Thomas, and John Algeo. 1993. *The Origins and Development of the English Language*. Fort Worth: Harcourt Brace Jovanovich College Publishers.
- Ranson, Diana. 1995. Morphological complexity as a criterion for evaluation of language difficulty: the French and Spanish Verb Systems. [Paper delivered at ACTFL, Anaheim, CA.]
- Real Academia Español. 2002. Formas verbales. [Online.] Available: <http://dlc.rae.es> [accessed August 4, 2002].
- Richards, Jack C. and Theodore S. Rogers. 1986. *Approaches and Methods in Language Teaching: A Description and Analysis*. Cambridge: Cambridge University Press.

- Rousseau, Jean Jacques. 1762. *Emile or On Education*. Translated by Grace Roosevelt.
Available: <http://www.ilt.columbia.edu/Projects/emile/emile.html> [2002, Nov. 19].
- Ruiz-M., Hildebrando. 1986. Vocalic alternations in the speech of Spanish bilinguals.
Revista Canadiense de Estudios Hispánicos, vol. XI, no. 1: 99-118.
- Ruiz-M., Hildebrando. 2000, Nov. 14. Personal communication by e-mail.
(hruiz@arches.uga.edu).
- Rumelhart, David E. 1989. The architecture of mind: A connectionist approach. [In: Posner, Michael I. (ed). *Foundations of Cognitive Science*. Cambridge, MA: MIT Press. 133-159.]
- Saussure, Ferdinand de. 1916. *Course in General Linguistics*. [Bally, Charles, and Albert Sechehaye (eds.). Baskin, Wade (trans). New York and Toronto: McGraw-Hill. 1959.]
- Scalese, Sergio. 1988. Inflection and derivation. *Linguistics* 26: 561-581.
- Schulz, Renate A. 1991. Second language acquisition theories and teaching practice: How do they fit? *The Modern Language Journal*, 75, i: 17-26.
- Seco Reymundo, Manuel. 1999. Gramática esencial de la lengua española. Madrid: Espasa Calpe, S.A.
- Simon, Herbert A., and Craig A. Kaplan. 1989. Foundations of cognitive science. [In: Michael I. Posner (ed.). *Foundations of Cognitive Science*, Cambridge, MA: MIT Press. 1-20.]
- Skousen, Royal. 1989. *Analogical Modeling of Language*. Dordrecht, The Netherlands: Kluwer.

- Skousen, Royal. 1995. Analogy: A non-rule alternative to neural networks. *Rivista di Linguistica* 7: 213-232.
- Spencer, Andrew. 1988a. Arguments for morpholexical rules. *Journal of Linguistics* 24: 1-30.
- Spencer, Andrew. 1988b. Lexical rules and lexical representation. *Linguistics* 26: 619-640.
- Spencer, Andrew. 1997. *Morphological Theory*. Oxford: Blackwell.
- Stemberger, Joseph Paul. 1994. Rule-less morphology at the phonology-lexicon interface. [In: Lima, Susan D., Roberta L. Corrigan, and Gregory K. Iverson (eds.). *The Reality of Linguistic Rules*. Amsterdam: Benjamins. 147-169.]
- Stemberger, Joseph Paul and Brian MacWhinney. 1988. Are inflected forms stored in the lexicon? [In: Hammond, Michael, and Michael Noonan (eds). *Theoretical Morphology*. San Diego: Academic Press. 101-116.]
- Stix, Gary. *Scientific American*. September 2002: 36-39.
- Swain, Merrill. 1985. Communicative competence: Some roles of comprehensible input and comprehensible output in its development. [In: Gass, Susan M., and Carolyn G. Madden (eds). 1985. *Input in Second Language Acquisition*. Rowley, MA: Newbury House. 235-53.]
- Taft, Marcus. 1988. A morphological-decomposition model of lexical representation. *Linguistics* 26: 657-667.
- Terrell, Tracy D. 1977. Constraints on the Aspiration and Deletion of Final /s/ in Cuban and Puerto Rican Spanish. [In: Hashemipour, Peggy, Ricardo Maldonado, and Margaret van Naerssen (eds.) 1995. *Studies in Language Learning and*

- Spanish Linguistics in honor of Tracy D. Terrell*. New York: McGraw-Hill. 261-280.]
- Terrell, Tracy D. 1982. The natural approach to language teaching: An update. [In: Hashemipour, Peggy, Ricardo Maldonado, and Margaret van Naerssen (eds). 1995. *Studies in Language Learning and Spanish Linguistics in honor of Tracy D. Terrell*. New York: McGraw-Hill. 18-34.]
- Terrell, Tracy D. 1986. Acquisition in the Natural Approach: The Binding/Access Framework. *Modern Language Journal* 70iii: 213-27.
- Terrell, Tracy D. 1991. The role of grammar instruction in a communicative approach. *The Modern Language Journal* 75i: 52-63.
- VanPatten, Bill. 1987. How juries get hung: problems with the evidence for a focus on form. *Language Learning* 39: 243-260.
- Waldrop, M. Mitchell. 1992. *Complexity: The Emerging Science at the Edge of Order and Chaos*. New York: Simon and Schuster.
- Warren, Paul, and William Marslen-Wilson. 1987. Continuous uptake of acoustic cues in spoken word recognition. *Perception and Psychophysics* 4: 262-275.
- Weinrich, Harald. 1964. *Estructura y función de los tiempos en el lenguaje*. [Versión española de Federico Latorre. 1968. Madrid: Editorial Gredos.] (*Tempus, Besprochene und Erzählte Welt*. Stuttgart: W. Kohlhammer Verlag.)
- Weinrich, Uriel, William Labov, and Marvin I. Herzog. Empirical foundations for a theory of language change. [In: Lehmann, Winfred Philip, and Yakov Malkiel (Eds). 1968. *Directions for Historical Linguistics: A Symposium*. Austin: University of Texas Press. 95-195.]

- Whorf, Benjamin L. 1956. *Language, Thought and Reality*. [Carroll, John B. (ed.).]
Cambridge: MIT Press.
- Wurzel, Wolfgang. 1984. *Flexionsmorphologie und Natürlichkeit*. Berlin: Akademie
Verlag. [Schentke, Manfred (trans.). 1989. *Inflectional Morphology and
Naturalness*. Dordrecht, The Netherlands: Kluwer.]
- Zwicky, Aenold. 1985. How to describe inflection. *Berkeley Linguistics Society* 11: 371-
86.