STUDIES IN FUNCTIONAL AND STRUCTURAL LINGUISTICS 53

Spanish Phonology and Morphology

David Eddington

Spanish Phonology and Morphology

Studies in Functional and Structural Linguistics (SFSL)

Taking the broadest and most general definitions of the terms functional and structural, this series aims to present linguistic and interdisciplinary research that relates language structure — at any level of analysis from phonology to discourse — to broader functional considerations, whether cognitive, communicative, pragmatic or sociocultural. Preference will be given to studies that focus on data from actual discourse, whether speech, writing or other nonvocal medium.

The series was formerly known as *Linguistic & Literary Studies in Eastern Europe (LLSEE)*.

Volume 53

Spanish Phonology and Morphology:
Experimental and quantitative perspectives
by David Eddington

Spanish Phonology and Morphology

Experimental and quantitative perspectives

David Eddington
Brigham Young University

John Benjamins Publishing Company
Amsterdam/Philadelphia

The paper used in this publication meets the minimum requirements of American National Standard for Information Sciences – Permanence of Paper for Printed Library Materials, ANSI Z39.48-1984.

Library of Congress Cataloging-in-Publication Data

David Eddington
Spanish Phonology and Morphology : Experimental and quantitative perspectives / David Eddington.
p. cm. (Studies in Functional and Structural Linguistics, ISSN 0165–7712 ; v. 53)
Includes bibliographical references and index.
1. Spanish language--Phonology. 2. Spanish language--Morphology.
I. Title. II. Series.

PC4131.E33 2004
465--dc22 2004062251
ISBN 90 272 1562 6 (Eur.) / 1 58811 612 3 (US) (Hb; alk. paper)

John Benjamins Publishing Co. · P.O. Box 36224 · 1020 ME Amsterdam · The Netherlands
John Benjamins North America · P.O. Box 27519 · Philadelphia PA 19118-0519 · USA

To my wife and inspiration, Humi

Table of contents

Acknowledgments

I owe the particular view of language I espouse in this book to the writings and insights of three scholars: Joan Bybee, Bruce Derwing, and Royal Skousen. Although their ideas and mine do not always mesh completely, their calls for substantive evidence have been heeded and their influence is evident throughout the book.

I would like to thank several professors of mine who guided and inspired me in the early years of my career, who instilled me with a love of Spanish and an enthusiasm for linguistics: Halvor Clegg, Willis Fails, Orlando Kelm, and Carlos Solé. I have also been blessed with the friendship of numerous colleagues at Brigham Young University and the University of New Mexico as I worked to prepare the manuscript for this book.

The anonymous reviewers provided invaluable input about the style and content of the manuscript, as did Joan Bybee, José Ignacio Hualde, Tim Face, and Alan Meredith. In many places they saved me from having to extract my foot from my mouth post-publication. Converting the initial drafts into a publishable book may be attributed to the editing expertise of Mel Thorne and the Humanities Publication Center at Brigham Young University, as well as to Stephen Mouritsen.

However, the single most supportive and encouraging influence behind this book, and for life in general, is my wife Humi.

Introduction

I had originally intended to subtitle this book *A View from Left Field*. The difficulty with that is twofold. First, the subtitle would not have given much insight into the actual approach taken. Second, the term 'left field' has already been used in a linguistic monograph (Zwicky et al. 1971) whose subject matter and tone I do not intend to mimic. Nevertheless, the baseball metaphor does have its merits. What exactly is meant by 'left field'? In baseball, the majority of the players are stationed in the infield where most of the action in the game takes place. A player positioned in left field is much less likely handle the ball, (especially in some early baseball fields that were irregularly shaped), and plays at some distance from the rest of the members of the team. The term 'left field' may also be due to the fact that Babe Ruth played right field in Yankee Stadium. Fans with seats in left field were not in a good position to observe this stellar player (Ammer 1993). Hence, the left field belongs to heterodox, unconventional, nontraditional ideas located far from the mainstream infield and distant from the more publicized players. The left field also tends to have fewer adherents than the more traditionally accepted views.

In one sense, distinguishing left fielders from infielders in linguistics is easily accomplished. One need only listen to the jargon. Infield players employ terms such as *feature percolation, constraint ranking, C-command, optimality, and universal grammar*. Left fielders, on the other hand, have their own parlance that includes *corpus, token frequency, statistical significance, usage,* and *experiment*.

Where do left fielders come from? Well, some have apparently been in left field throughout their careers. Others had coaches from the left field. A few 'Michael Jordaned' their way into the game from a different sport. Nevertheless, there are those who started playing in the infield. These players witnessed the radical changes in theoretical game plans that were implemented every decade or so. Each new game plan promised to rectify the problems of past approaches and revolutionize the sport. However, these players became disenchanted with mainstream linguistics or mainstream methodology and transferred to the left field. For them, a change of game plan meant taking the

plays that have been used for decades (passivization, nasal assimilation, plural formation, etc.) and merely fitting them into a new theoretical mold. The proponents of the theoretical game plans seemed to be most interested in the formalisms and mechanisms that the plans entailed, but these players felt that few new insights into the actual mechanisms involved in playing the sport were brought to light with the institution of each new plan.

A left-field approach to the game, because of its nontraditional and often iconoclastic nature, is sure to be viewed as controversial by most infielders. However, the left field itself is not a homogenous group of like-minded individuals, so even some left fielders may find certain approaches to be a bit too radical for their taste. One difficulty with the approach I take in this book is that I run the risk of misclassifying players. Players I may feel are left fielders may contest that assignment insisting that they are really shortstops or first basemen. Secondly, the left field has recently become much more populated and it is impossible to include every player on the roster of left-field Spanish linguistics. As is often the case, once a viewpoint that what was once considered avant-guard and unconventional gains a significant number of adherents it becomes the new orthodoxy. Left field linguistics has certainly grown substantially in the past few decades. The ever-increasing number of sociolinguists alone attests to this fact. However, rather than taking the place of the old orthodoxy, the various subfields of left field linguistics have simply established their own.

One thing I have failed to do in this grossly overextended baseball metaphor is give a more recognizable name to what I consider the left field. In my view, it comprises all approaches that are quantitative and experimental in nature. This would include researchers who seek quantitative and experimental evidence for generative constructs (see 2.4.1). Among others, much but not all of what is done in the fields of corpus linguistics (Biber 1998), sociolinguistics (Labov 1994), psycholinguistics (Gernsbacher 1994), language acquisition (Doughty & Long 2003; Lust & Foley 2004), prototype theory (Taylor 1995), laboratory phonology (Pierrehumbert, Beckman, Ladd 2000), exemplar theory (Chandler 2002), phonetics (Lass 1996), connectionism (McClelland 1988), and usage-based linguistics (Barlow & Kemmer 2000) would fit into my theoretical left field.

Most quantitative left field researchers would agree that their studies are not attempts to organize linguistic information in the most concise or elegant way, or within the most popular formal framework of the day, nor are the results of their studies designed to reflect the grammar of a fictitious ideal speaker-hearer. Instead, those who employ these approaches share the sense

that their studies give some insight into how actual Spanish speakers process language in real time.

Before delving into specific quantitative approaches to Spanish linguistics, it is important to discuss the psychological status of linguistic analyses since it is not an uncontroversial issue – this is the topic of Chapter One. In Chapter Two, I argue that quantitative and experimental data are an essential part of understanding how people process language. More specifically, I present an apologia for the experimental approach to linguistics. The remainder of the book focuses on a range of issues in Spanish linguistics that have been analyzed from experimental and quantitative approaches. Chapter Three discusses processes that have been proposed for Spanish without firm evidential grounding. Chapters Four and Five discuss the important and often overlooked roles that frequency and analogy play in cognition. Empirical analyses of diphthongs, syllables, and stress assignment are the focus of Chapter Six, and Chapter Seven covers studies of morphological processing in word recognition.

No monograph is able to cover all facets of a particular topic, and this one is no exception. The title itself narrows the focus to Spanish phonology and morphology which excludes discussion of studies of semantics and syntax for example. However, valuable studies of Spanish morphology and phonology have been produced by connectionists, sociolinguists, corpus linguists, and researchers in language acquisition. Each of these fields has large numbers of active investigators, professional conferences, and area specific journals. For this reason, I have chosen not to include such studies in the present work. Readers who are interested are referred to the vast amount of already available literature on those topics.

Chapter 1

The psychological status of linguistic analyses[1]

1. Introduction

The psychological status of linguistic analyses was a much debated topic among what could be termed 'practicing linguists' starting in the 1970s (Baker 1979; Botha 1971, 1973; Chomsky 1972, 1976; Cutler 1979; Derwing 1973, 1979; Goyvaerts 1978; Itkonen 1976, 1978a, 1978b; Lass 1976b; Rischel 1978; Steinberg 1975). In the following two and a half decades since then the subject has been continued mainly by philosophers of language (e.g., Itkonen 1983; Katz 1985; Katz & Postal 1991; Yngve 1986, 1996, 2000) and linguists with experimental tendencies (e.g., Pierrehumbert, Beckman, & Ladd 2000; Stemberger 1994). Has the issue of psychological reality been resolved? No, instead of resolving the debate many researchers seem to have put it on the back burner while they turned their attention to the newly emerging theories of the day (e.g., minimalism, autosegmental phonology, government and binding, optimality theory, etc.). The purpose of this chapter is to examine the status of linguistic analyses as psychologically pertinent analyses.

1.1 The psychological status of formal mechanisms

Prior to the advent of generative linguistics, linguistics in the United States was decidedly behavioristic. The behaviorists, led by Bloomfield, felt that the psyche was an unobservable entity, and as such, any arguments that were psychological in nature were unscientific, and therefore, rejected. It is Chomsky who is principally responsible for resurrecting and propagating the idea that linguistic analyses are not mere descriptions of linguistic data. He believes that they represent the speaker's actual underlying knowledge of the language, which he terms competence. According to Chomsky, linguistics is a "branch of cognitive psychology" (1972: 1). More specifically, he asserts that linguistic rules and principles are psychologically real (1980a: 48).

Many contemporary linguistic analyses follow Chomsky in asserting psychological relevance. For example, it is extremely common for an analysis to begin by claiming that its goal is to provide a descriptively adequate account of some linguistic phenomenon. By definition, a descriptively adequate analysis describes the "linguistic intuition of the native speaker" (Chomsky 1964: 28–29). Therefore, in one sense, an analysis that claims descriptive adequacy also purports to be psychologically valid, and not to merely describe the language data.

The issue of the psychological status of linguistic theories centers on formal mechanisms such as rules, parameters, and constraints. Once a generalization has been formalized into a rule or system of constraints, what exactly is meant when it is claimed that the rule or constraint is "psychologically real"? There are two senses in which grammars may be considered psychologically real. Cutler (1979: 79) defines them in this way:

> In the strong sense, the claim that a particular level of linguistic analysis X, or postulated process Y, is psychologically real implies that the ultimately correct psychological model of human language processing will include stages corresponding to X or mental operations corresponding to Y. The weak sense of the term implies only that language users can draw on knowledge of their language which is accurately captured by the linguistic generalization in question. (see also Steinberg 1975: 218–20)

The strong sense implies a close relationship between the way a theoretical analysis works out on paper and the internalized representations and mental processes speakers possess. On the other hand, the weak sense involves little correspondence between theoretical constructs and psychological mechanisms. Accordingly, if an analysis produces the same output as do the speakers of the language, then in a limited way it may claim to have achieved psychological validity (Rischel 1978: 442).

Formal linguistic mechanisms are often spoken of as if they were algorithmic operations or mental processes. If they were then they would be candidates for attaining reality in the strong sense. However, though they are often spoken of in this way they are usually defined in terms of speakers' intuitions, tacit knowledge, or underlying representations, not in terms of psychological mechanisms. Since they are defined in these terms they cannot be considered psychologically valid in the strong sense. As a result, one must speak of the psychological reality of a formal mechanism in Cutler's weak sense. This means that what is potentially real in an analysis is not the formal notation with its transformations, derivations, parameters, rules, and constraint rankings.

What is real is "the function that these constructs serve to specify" (Matthews 1991: 197), or the contents of the rules (Rischel 1978: 442).

1.2 Reasons for doubting the psychological reality of linguistic analyses in the weak sense

In principle, an analysis of a given linguistic phenomenon may achieve psychological significance in the weak sense of the term. However, there are even reasons for doubting the psychological reality of formal analyses in the weak sense. These fall into four categories: (1) the 'truth equals reality' issue, (2) the empirical status of linguistic analyses, (3) the question of autonomous versus nonautonomous methodology, and (4) the narrow base of evidence. Each of these will be discussed below.

1.2.1 *Truth equals reality*

One of the most common criticisms of linguistic analyses is that there is insufficient justification that they are more than mere descriptions, but actually have some kind of bearing on human language faculties. The objection is that a detailed, rigorous, or sophisticated description of a linguistic phenomenon does not necessarily indicate that the phenomenon has any relevance to linguistic cognition. This sentiment has been expressed by a number of linguists (e.g., Botha 1971; Derwing, Prideaux, & Baker 1980; Goyvaerts 1978; Lass 1976b; Morin 1988; Skousen 1989).

In response to such charges, Chomsky has attempted to explain the justification for assigning mentalistic status to grammatical constructs. As far as he is concerned, linguistics achieves this in the same way empirical sciences such as physics do, the only difference being that researchers in physics deal with physical entities while linguists "are keeping to abstract conditions that unknown mechanisms must meet" (Chomsky 1976: 9).

Linguistic theories typically include entities such as constraints, parameters, transformations, and rules, none of which are directly observable. This is analogous to the way quantum physics includes entities such as quarks, or psychology includes concepts such as the superego or motivation. Since these entities are not directly observable, their existence is motivated entirely on indirect evidence. Mohanan's (1986: 185) statement illustrates argumentation based on indirect evidence:

> It does not make sense to ask, 'does the unconscious mind really exist?' When faced with this question, what the psychologist does is to produce a range of

> facts which the notion of the unconscious mind can account for. Therefore, the question ought to be, 'What kind of EVIDENCE do we have for ASSUMING the theoretical entity 'unconscious mind'? Similarly, it does not make sense to ask if syllables and segments really exist in the human mind; instead, we should ask what kind of evidence we have to assume that they are part of the mental representations of language users.

According to this line of reasoning, if a theoretical entity aids in describing the structures or distributional patterns of a language, then there is ample evidence that the entity exists.

Chomsky protests that the reality of theoretical linguistic entities is often questioned, while the entities of other sciences are not. In linguistics, as well as in other sciences, theoretical entities are postulated because the data and evidence that have been gathered to date support their existence. In any science, once all the evidence has been gathered and the best possible theory based on that evidence has been put forth, the theory is not only deemed true within its theoretical domain but it is also considered to reflect reality as well (Chomsky 1980a: 190–191; see also Chomsky 1986: 252–257).

If one assumes that there is no distinction between the best possible analysis of some linguistic phenomenon and the psychological relevance the analysis may have for language speakers, it would necessarily follow that all aspects of a 'correct' analysis are also aspects of a psychologically significant one. Harmon (1980: 21–22) explains why this is not necessarily the case:

> Given any theory we take to be true, we can always ask what aspects of the theory correspond to reality and what aspects are mere artifacts of our notation. Geography contains true statements locating mountains and rivers in terms of longitude and latitude without implying that the equator has the sort of physical reality the Mississippi River does. (see also Matthews 1991: 196)

The stance taken by Chomsky leaves no room for some elements of the theory to be real and for others to be simply artifacts of the formal notation used within a given theoretical framework. This is surely an untenable position, one which Chomsky later abandoned, admitting that "there is a question of physical (or psychological) reality apart from truth in a certain domain" (1980b: 45). In other words, the search for linguistic realities and psychological realities belongs to separate domains (see Yngve 1986, 1996, 2000).

The question that now arises is this: If constructs belonging to the best analysis of a linguistic phenomenon are not necessarily real, why are constructs belonging to other sciences (e.g., atoms, quarks, black holes) considered real by the same token? The answer is simply that generative linguistics is a funda-

mentally different type of science than is physics or psychology, for example (Katz 1981, 1985; Lass 1984: 104). In the next section, I will argue that the theories, constructs, and methodologies of sciences such as physics and psychology are empirical, while much of what is done in contemporary linguistics is non-empirical.

1.2.2 *Empiricality*

Chomsky charges that linguistics is confronted with issues that other sciences are not forced to deal with. More specifically, linguists are often asked to produce "more convincing evidence" that their theoretical constructs are real, yet the evidence that scientists working in the hard sciences produce in favor of a theoretical entity is often accepted at face value (Chomsky 1980a: 22). This difference stems from the greater confidence that people have in the physical reality of a theoretical construct which belongs to an empirical science than they have in one pertaining to a nonempirical science. A great deal of what is done in contemporary formal linguistics falls into the latter category.

Quite a large body of literature exists which demonstrates that much of what is done in linguistics is nonempirical in nature, and should therefore be grouped together with other nonempirical sciences such as formal logic, pure mathematics, and philosophy (Baker & Hacker 1984; Botha 1971, 1973; Derwing 1973; Hacker 1990; Hall 1987; Itkonen 1976a, 1978a, 1978b, 1983; Katz 1981, 1985; Katz & Postal 1991; Lass 1976b; Ringen 1975; Sampson 2001; Steinberg 1975; Yngve 1986, 1996, 2000). There are several reasons cited for making this classification. One of the most common has to do with falsifiability.

1.2.2.1 *Falsifiability.* Falsifiability is the criterion that distinguishes empirical theories from nonempirical theories (Popper 1968: 27–48). In order for a theory to be empirical it must be potentially falsifiable. Theories that can be potentially falsified possess a sense of concreteness and reality which is lacking in theories for which there is no counterevidence or other possible refutation. One theoretical notion that is not subject to potential falsification is that of the ideal speaker-hearer. Many analyses claim to describe the linguistic system of an ideal speaker-hearer. It should be clear that a concept such as the ideal speaker-hearer is useful in constructing descriptive grammars. In the beginning stages of theory building, a model such as the ideal speaker-hearer is often proposed that excludes some factors for simplicity's sake. The model of the ideal speaker-hearer does not incorporate certain factors such as individual differences, slips of the tongue, or ambient noise, in an attempt to devise a

simple yet general model of language. This, of course, is no different than what a physicist does when excluding friction as a factor in a model of what happens to objects as they move through space. The crucial difference is that at some point the physicist factors the effect of friction back into the theory.

What ultimately makes the ideal speaker-hearer unfalsifiable is that the factors that are originally omitted in order to simplify the model are never reincorporated. In most formal linguistic analyses, the model language of the ideal speaker-hearer, void of all real-life variables, is in itself the object of study; therefore, it is not subject to falsification (Hall 1987:38; Prideaux 1980:247; Wheeler 1980:77–78). No evidence of any kind could possibly falsify what the linguistic system of an ideal speaker-hearer may or may not contain. This is true because the ideal speaker-hearer is a convenient fiction whose mental grammar is not subject to inspection.

What constitutes the grammar of an idealized speaker-hearer is a perfectly legitimate realm of inquiry, as Carr (1990, 2000), Katz (1981, 1985), and Katz and Postal (1991) have demonstrated. Pedagogical grammars, for example, describe an ideal state of affairs supposedly shared by all speakers of a given language. However, it is the supposition that all elements of such a grammar necessarily have correlations to the grammars of actual speakers that leads to skepticism (Derwing 1980:173). This is not to deny the possibility that some aspects of the ideal speaker-hearer's grammar may very well relate to aspects of the mental grammar of actual speakers. However, exactly which aspects are and are not psychologically real is indeed an empirical question, the answer to which should not be determined on the basis of a priori assumptions nor nonempirical evidence. Consider the concept of the phoneme, for example. In some ways, it is fairly abstract; unlike allophones, one cannot measure the fundamental frequency or duration of a phoneme. Initially, then, it appears that phonemes have no spatiotemporal manifestations apart from their respective allophones and should therefore escape possible spatiotemporal falsification. However, the reality of the phoneme has been demonstrated in studies of perception (e.g., Jaeger 1980), and perception is a process that occurs in time.

The distinction that is often made between competence and performance has been accused of immunizing analyses from falsification. Competence is defined as a speaker's knowledge of language, while performance is the use of that knowledge (Chomsky 1980a:205). Competence is an idealized concept that comprises the system of rules or other formal mechanisms that are thought to underlie a speaker's ability to produce and understand language. Performance is made up of specific utterances, and unlike competence, it includes errors.

The insulating effect the competence/performance distinction has on theoretical analyses is clearly seen in this example: Suppose that two subjects react differently from each other to a question in a psycholinguistic experiment. One could say that the subjects have the same underlying knowledge of the language (competence), but that the experimental design was responsible for the differences. That is, something in the experiment kept the subjects' reactions from reflecting their competence. Remember that competence is idealized in that it is free from errors and extralinguistic influences, while performance includes errors. Therefore, it could be argued that the experiment must have measured performance not competence. An analysis that claims to describe the linguistic competence of the subjects would be unaffected by the differing responses given by each subject since the results of the experiment would be indicative only of the subjects performance (see Derwing 1983: 66; Wheeler 1980: 78–90).

The question of where competence stops and performance begins is also difficult if not impossible to determine. Any time an utterance is made, whether in the laboratory or in spontaneous speech, it necessarily involves both competence and performance. That is, there is no such thing as "pure competence," unfettered by performance factors (Hacker 1990; Schütze 2003; Stemberger 1994; Wheeler 1980: 67; Zimmer 1969: 320). In fact, everything known about language is actually based on performance (Pierrehumbert, Beckman, & Ladd 2000: 290). In short, since competence is not subject to inspection it is does not lend itself to possible refutation.

It must be stressed that the fact that few linguistic analyses meet the requirement of potential falsifiability has no bearing on whether they are good NONEMPIRICAL analyses, only that they are not empirical. As such, they may be validated or refuted in the same way a philosophical argument is. According to Itkonen (1976a: 15–16), the nonempirical linguist and the logician have a similar goal:

> to generate all and only intuitively valid formulae: insofar as they fail to do this, their systems are (non-empirically) falsified … not by reference to some specific spatiotemporal occurrences, but showing that it does not capture the concept which it tries to capture.
>
> (See also Carr 1990: 66; Kac 1992: 39; Linell 1976: 84–85)

In a sense, nonempirical theories are falsified by one type of data, and empirical ones by another. Clearly, formal theories of language are not usually judged on the basis of spatiotemporal data, but on the degree to which they correctly capture the linguistic system of the language in question.

1.2.2.2 *Spatiotemporality.* In addition to being falsifiable, empirical theories must deal with events, activities, or processes that take place in time and/or across space (Itkonen1978a: 80; Popper 1968: 27–48, 102–103). This means that the falsification of such theories must be done on the basis of events or actions that take place across time and/or space. Nonempirical theories, on the other hand, either do not deal with spatiotemporal events, or deal with them but are formulated in such a way that they can be falsified, but not by recourse to spatiotemporal events (Itkonen 1978a: 155).

In contrast to the definition supported above, Carr (1990, 2000) defends formal linguistics as a truly empirical science. In his view, "linguistic realities are quite reasonably taken to be speaker-external, and thus not psychological in nature" (1990: 113). He suggests that formal hypotheses about grammatical structure and organization are subject to falsification, which makes them empirical. The crucial difference is that the type of falsification Carr refers to is not based on spatiotemporal events but on whether an analysis captures the concept it was initially designed to capture. In fact, besides grammaticality judgments, Carr eschews all other spatiotemporal data as irrelevant to proving or disproving linguistic theory as conceived of by Chomsky, although such spatiotemporal evidence may be relevant to other approaches to linguistics (2000).[2] In other words, Carr maintains that formal linguistics is indeed empirical but that it differs from other empirical enterprises in that its tenets are subject only to possible nonspatiotemporal falsification. As Love (1992) points out, this is an extremely controversial position because it would make linguistics the only empirical science not based on spatiotemporal events.

There is a relationship between a theory's spatiotemporal falsifiability and its reality. A theory that can be proved or disproved on the basis of spatiotemporal events possesses a sense of tangibility and concreteness because events that are considered real (as opposed to abstract) take place in time and space. This same sense of tangibility is missing in a theory for which there are no spatiotemporal manifestations or that eludes possible substantive falsification.

Remember that the constructs and operations of many linguistic theories are not thought to model the actual steps or processes utilized by actual speakers in producing and comprehending language. Instead, they represent the idealized language of an ideal speaker-hearer. Idealizations are a fact of life in all sciences. However, any idealized state of affairs is expected to reflect reality. Unfortunately, the same cannot be said of idealizations in many aspects of linguistic theory (Bresnan & Kaplan 1982: xxiii). This is simply because the theories do not profess to relate to actual spatiotemporal events, which prompts the question: "What kind of reality can be ascribed to the notion of rule whose

mental existence is not open to introspection and whose operations are ordered in non-real time?" (Derwing & Skousen 1989: 54). The answer, of course, is that such operations are purely notational or formulaic (Mohanan 1997; Steinberg 1975: 246–247). Therefore, they fall outside the realm of empirical science and into the realm of nonempirical science. A good case may be made for the reality of theoretical entities that are demonstrable by reference to spatiotemporal events. By the same token, the reality of theoretical entities that are not based on spatiotemporal evidence is highly speculative.

Of what value are formal theories of language that do not relate to actual events or processes then? The best answer is given by Mohanan (1997):

> Proposals for context free and context dependent PS rules, structure building and structure changing rules, transformational and non-transformational grammars, constraints, percolation, and so on, are formal frameworks in this sense: they give us a domain specific formal language and a calculating system to deduce the consequences of the laws and representation of the organization of human language, but they do not tell us anything about the content or substance of these laws and representations.

Mohanan goes on to state that even though formal frameworks do not "make any claims about the world," exactly which framework is the best formal characterization of human languages is a question that can be answered.

In the above discussion, I have given several reasons for classifying many linguistic theories as nonempirical. Some have reacted unfavorably to any suggestion that linguistics is less than empirical, which is most probably due to the fact that "the word *empirical* has become so prestigious that it has blinded linguists to the respectability of non-empirical theories" (Lass 1976a: 217). The frequent, almost hackneyed, use of the word *empirical* in linguistic literature suggests a belief that a science becomes empirical or that evidence becomes empirical by mere denomination not by conforming to empirical criteria.

1.2.3 *Methodology: Autonomous versus nonautonomous*

The nonempirical status of many analyses is one reason that people question claims of psychological relevance. Another reason is that most analyses are carried out in almost complete isolation from the speakers of the language themselves. A number of linguists have suggested that there are two distinct types of linguistics each with a different methodology. These approaches will be referred to as autonomous and nonautonomous approaches.[3] In this section, I discuss the reasons why there is so much doubt about psychological claims that are arrived at by autonomous methods.

Kac (1974:42) defines autonomous linguistics as a field of study that attempts to produce psychologically significant theories without performing psychological experiments. It is the study of language structure in isolation from the pragmatic environment in which speech occurs, and without recourse (or at most with only minimal recourse) to actual speakers. It involves the study of the patterns, alternations, and structures that are found to exist in linguistic data. It is a metaphysical or philosophical realm of inquiry that deals with axiomatizations about linguistic structure which "make it possible to deduce all true statements about the system from a small set of prior assumptions about its nature" (Kac 1974:44).

Nonautonomous linguistics, on the other hand, is the study of language as an entity that is inseparable from the speakers of the language. It examines language with the methodological tools of the experimental cognitive psychologist and attempts to determine what speakers know about their language and how they actually process and store linguistic information. The objection that is commonly raised against autonomous methodology is this (Derwing 1980): How can autonomous analyses profess to be pertinent to the speaker's knowledge or manipulation of linguistic elements if they are arrived at with little or no recourse to actual speakers as if human language were an entity separate from humans? Of course, an autonomous analysis may be considered potentially real in that it "reflects a kind of abstract complexity with which somehow the human brain must cope" (Goyvaerts 1978:12), but it does not necessarily spell out how the brain copes with it. Autonomous linguistics studies the structures that exist in languages. Therefore, it is a field of inquiry that is psychologically relevant in that the structure of a given language is what renders it "capable of being learned and employed by speakers" (Kac 1974:42). It does not, however, express what speakers actually know about the structures of their language, nor how they utilize them. Only nonautonomous analyses are justified in making the claim that they characterize speakers' actual, not simply their potential linguistic competence (Kac 1974:42).

Although theories arrived at by autonomous methods may be useful and potentially real, they do have serious limitations. Researchers who fail to fully incorporate the speakers of a language into a theory about the linguistic system of those speakers run the risk of treating language as an entity completely separate from humans; as a result, it becomes easy to confuse psychological reality with descriptive validity. As Black and Chiat note (1981:42), it is a common practice to mislabel a "linguistically valid analysis" as "psychologically valid" when no psychological substance is intended. In Smith's (1999:94) words "to claim that the grammar you have hypothesized in order to explain a range of

different facts is psychologically real is simply to claim that it's correct... .[It is not] a claim about how the language is neurophysiologically implemented in the brain." This sort of terminological confusion continues in linguistic thought to the present day.

Autonomous linguistics has had tremendous success in discovering linguistic structures, patterns, and generalizations that are to be found in language data. Because the data have been produced by humans, it is possible that humans have knowledge of, or utilize those structures, patterns, and generalizations, however, their existence is not proof that speakers do use them, nor that they have knowledge of them. It only demonstrates that those structures and patterns are available to be potentially known or used. In order to determine what is actually known or utilized by the speakers the focus of the research must turn back to the speakers themselves.

It is highly possible that some patterns, structures, and generalizations exist in a given language that speakers do not use or have knowledge of. A phonological structure may have arisen by chance or may be the result of a diachronic process that has long since died. Certain alternations or phonotactic patterns may be due to purely articulatory or aerodynamic influences (see Diver 1979), and in that case, would neither be knowable nor psychological. The task of nonautonomous linguistics is to differentiate between those structures that possess potential psychological relevance and those that have actual psychological significance.

Of course, it should be noted that some analyses do include native speakers' intuitions about the correctness or incorrectness of linguistic structures. Certainly this is a step toward the development of a mentalistic theory of linguistics. However, the methodology used in gathering intuitions is frequently poor. Often, one's intuitions about one's own speech are in direct conflict with what one actually produces (Sampson 2001). Even when the methodology is good, it is still just as prone to problems of validity as any other type of experimentally acquired data (Bard, Robertson, & Sorace 1996). For these reasons, there is no valid motivation for giving speaker intuitions a privileged status over other types of evidence.

One difficulty with the use of intuitions is that many analyses are founded solely on the intuitions of the very linguists who perform them. A theory of psychology based on data psychologists are able to glean from their own psyches would surely be considered seriously methodologically flawed.[4] The task of any scientist is to gather evidence in the most objective way possible in order to be able to assert with the utmost confidence that the evidence is valid beyond the laboratory. An analysis that centers on the linguist's personal intro-

spections about some phenomenon is suspect because if a scientist is able to invent the facts he uses to test his or her hypotheses, "the door is wide open to self-fulfilling predictions" (Sampson 2001: 124). Ohala (1990: 163) suggests that linguists' personal introspections are a poor source of evidence because

> The knowledge the linguist has about spoken forms, including historical derivational relationships between words, the morphemic structure of complex words, and the inductively-based knowledge of common cross-language sound patterns, is projected onto the mental grammar of linguistically-naive native speakers.

Any characterization of what constitutes the knowledge of the speakers of a language should ultimately come from linguistically naive speakers of the language.

Unfortunately, even studies that incorporate the intuitions of naive speakers are frequently methodologically weak. It is common for analyses to be based on the responses of only one or two subjects, which is far from a representative sample of a population. Very rarely do analyses founded on intuitions attempt to control for factors that may affect the internal or external validity of outcome, nor do they observe the procedures established and followed by other empirical sciences for determining the significance of their results statistically. The lack of sound psychological methodology is what prompts Derwing (1979: 117) to conclude that

> The so-called 'Chomskyan Revolution' may well have entailed a TERMINOLOGICAL re-orientation in the direction of the psychologization of linguistic jargon and the supposed domain of its interest and claims, [but] no corresponding METHODOLOGICAL revolution accompanied these changes.
>
> (see also Geeraerts 1989)

In summary, the psychological status of many analyses is suspect simply because its methodology is not trusted. This does not imply that the autonomous approach should be eliminated, only that one should be skeptical of psychological conclusions arrived at by autonomous methods. It is necessary to draw a sharp line between nonautonomous and autonomous approaches, as well as between psychological and nonpsychological conclusions (Carr 1990: 34–38; Itkonen 1976b: 219; Stemberger 1996).

1.2.4 *Evidence base*

One reason for questioning the psychological validity of many analyses is that the evidence on which they are founded is almost exclusively internal. Critics call for more external evidence, and charge that the external evidence that goes

against a particular theory is often overlooked or rejected as irrelevant (Mohanan 1986: 185). An actual definition of what sorts of evidence fit into each category is essential.

In phonology, internal evidence is primarily drawn from the data gleaned from a corpus of utterances. It involves generalizations that are based on the surface regularities and alternations found in the language data which spell out the distribution of linguistic structures, as well as what structures or elements are not found in a particular context. The internal evidence about the phonological system of a particular language consists of phonetic and phonemic alternations along with the restrictions that exist in the distribution of the phonological elements (Zwicky 1975: 154).

The major difference between internal and external evidence is that internal evidence comes from language used in unexceptional ways, such as the printed language and careful, monitored speech. External evidence, on the other hand, is evidence gathered from the actual use of language, especially its use in unusual and exceptional ways and situations (Campbell 1986: 171). Among other places, external evidence comes from language games, speech errors, intralanguage borrowing, aphasia, spelling errors, historical change, stylistic variation, informant judgments,[5] and of course, psycholinguistic experiments (Zwicky 1975: 154–5).

Some have thought that the two types of linguistic evidence correspond to two fields of linguistic inquiry. According to this view, internal evidence is relevant to an approach that has as its goal to describe the structure of a language without asserting that the resulting grammar has mentalistic import. On the other hand, a grammar that claims to be psychologically significant must not be founded on internal evidence alone, but must crucially include external evidence as well (Carr 1990: 34–38; Lass 1984: 214–215; Ohala 1990: 159–160; Wheeler 1980: 54). Itkonen's statement is representative of this sentiment (1978a: 220–221):

> It ought to be self-evident that this psycholinguistic hypothesis must be tested on the basis of new, independent [read *external*] evidence provided, above all, by psycholinguistic experimentation, and not on the basis of those very same grammatical descriptions [read *internal evidence*] which, in the first place, gave rise to the psycholinguistic hypothesis in question.

The major point of contention is that the language internal evidence, which forms the basis for assuming the existence of a theoretical entity, may not at the same time constitute proof of the existence of the entity in the psyches of the speakers of the language. Failure to make this distinction results in equating

the phenomenon explained with the explanation of the phenomenon (Higginbotham 1991:555; Ohala 1990:159; Sampson 2001:124). In other words, it is a case of elevating a linguistic description to the status of a psychological explanation (Black & Chiat 1981:48).

Of course, this opinion is not shared by all linguists. The assumption that certain kinds of evidence count as evidence for a good theory while others count as evidence for the psychological validity of the theory is considered absurd by Chomsky (1980a:107–108; 1995:33; see also Schütze 2003). What disturbs him is that linguistics is often asked to provide a new kind of evidence in order to prove its constructs are real, which is something not asked of researchers in other sciences. Chomsky gives the following analogy (1980a:189–191).

When an astronomer hypothesizes that certain thermonuclear reactions occur in the sun's interior, he or she presents all the available evidence and concludes that those reactions are physically real. Once all the available evidence has been gathered and interpreted, there is no evidence, short of physically exploring the sun's interior, that would entitle the astronomer to claim "a higher order of physical reality" than before. However, linguists are often asked for a different or better kind of evidence before the reality of their theories is accepted.

If the analogy between the astronomer and the linguist is extended, the reason more evidence is called for becomes clear. Imagine that the astronomer almost exclusively admits only that evidence that can be gathered through an optical telescope. He or she is convinced that observations through the telescope provide the best evidence for proving hypotheses about thermonuclear reactions in the sun. Furthermore, there are other means besides the telescope available for gathering information about the sun. The astronomer unhesitatingly states that these other methods may be very useful and condones their use, but makes no effort to use them in order to gain further insight. Moreover, when other astronomers present evidence that these other methods have provided, the astronomer either deems them irrelevant, or accepts only the evidence which corroborates the evidence obtained through the telescope. Would it be unusual or unwarranted under these circumstances for scientists to question the reality of the astronomer's theory and ask for more evidence?

The fact that a large body of evidence is commonly overlooked in the formation of linguistic theories has prompted some to group the overlooked types of evidence together (external) and to contrast them with the more commonly used types of evidence (internal). Chomsky is correct when he asserts that once all the evidence has been gathered and it supports a given theoretical construct,

one may safely conclude that the construct is real. However, many theories cannot presume reality because all the evidence is not considered.

When linguists are asked for more evidence, they are in effect being asked to include the evidence obtained by other means along with the evidence gathered through the 'telescope' (i.e., internal reconstruction). A true incorporation of all the available evidence does not, of course, mean that evidence of one type, or from one source, is accepted if it corroborates the theory and ignored if it does not. Unfortunately, this is precisely the manner in which external evidence is often treated (see Bertinetto 1992 for an example). Once both the internal and external evidence support a given theoretical entity a stronger case may be made for the reality of the entity.

There is a further difficulty with the analogy between the thermonuclear reactions in the sun and the linguistic knowledge of language speakers. The analogy exemplifies the widespread view that the only evidence, or at least the best evidence, is obtained through observation. The thermonuclear reactions in the sun may be studied only by observation. They may not be manipulated, nor tested, nor experimented with under controlled circumstances. The same is not true of human language capabilities, which lend themselves to methods other than passive observation–experimentation. Perhaps the reason that some make such a sharp distinction between internal and external evidence is that external evidence seems to be more telling of how language is actually manipulated, as well as what kinds of knowledge are drawn upon in order to produce and comprehend it.

Chomsky's position on the value of external evidence is somewhat inconsistent. On a number of occasions he has asked that more and varied kinds of evidence, including experimental evidence, be admitted into the pool of linguistic evidence (1981: 9; 1986: 36–37). Yet on another occasions, he questions the utility of certain experimental results (1982: 33; 1995: 33). His low regard for external evidence is clearly demonstrated below:

> As an objection of a narrower sort, one can take it seriously as an argument that the evidential base is too narrow to carry conviction; one who believes this might ask what other kinds of evidence would strengthen or undermine the theories we are led to construct on the basis of the (not inconsiderable) evidence that we can now readily obtain. In practice, what has been produced along these lines *has not been very informative*, but certainly any improvement in this regard will be welcome. (1986: 260, emphasis is mine)

In theory, Chomsky invites all kinds of evidence, but in practice he finds only a restricted kind of evidence truly compelling.[6]

In Section 1.2.1, it was seen that Chomsky originally held that there was no difference between the best theory and the reality of the constructs proposed by the best theory. If this is correct, then it follows that there cannot be two distinct kinds of evidence; one that is better for proving the psychological relevance of a theory and another that serves only to prove that the theory is true within its theoretical realm. This position is perfectly rational assuming no difference between the best theory and psychological reality. However, Chomsky later conceded that one must sort out what elements of the best (i.e., true) theory are real and which are artifacts of the theoretical notation. This concession would in turn logically entail another concession: one kind of evidence may be more pertinent to the truth of a theory, while another may have more significance for the psychological reality of the theory. Unfortunately, Chomsky has not yet made this concession (Botha 1989: 184–185).

Others maintain that the dichotomy exists and that external evidence offers more insight into the mind. According to this point of view, external evidence is vital to discovering which of the potentially internalizable linguistic structures are actually captured and productively used by the speakers of a language (Campbell 1979: 77; Skousen 1975: 20–21). In this regard, Mohanan (1986: 58–59) states:

> In the absence of clear evidence, we are forced to make guesses about which of the patterns have been internalised by a language user and which of the patterns are simply accidental correlations in the corpus. As soon as clear evidence from psycholinguistic experimentation on the storage, recognition, and production of linguistic forms becomes available, we must be willing to revise our initial guesses on the basis of new evidence.

In their introductory text on classical generative phonology, Kenstowicz and Kisseberth (1979: 154, 232) lament that the analyses of the languages they present are based almost entirely on internal evidence. They acknowledge that the lack of external evidence raises serious questions about the psychological validity of the analyses presented. In short, there are those who affirm that external evidence, and not internal evidence, will ultimately reveal what theoretical constructs are psychologically real (Derwing, Prideaux, & Baker 1980: 6; Itkonen 1978a: 85; Wheeler 1980: 65).

Perhaps the best stance to take on this issue is that all evidence is good and useful (Schütze 2003), but that some types are simply more telling of cognitive processing. Internal evidence is ambiguous to interpret in mentalistic terms; experimental evidence is less so in that it can be refined with experimental controls (Ohala cited by Fromkin 1980: 210–211). Nevertheless, internal evi-

dence is relevant to the search for psychological reality in that it determines what structures exist in a language and are thus available to be potentially internalized by the speakers of the language.

In this section, I have presented several arguments that demonstrate why there is doubt concerning the reality of many theories of language, while at the same time people are apt to believe that the theoretical constructs of other sciences are real. One may not understand the difference between empirical and nonempirical sciences, or between internal and external evidence, or understand what autonomous linguistics is. Nevertheless, it is clear that there are large differences between facts and evidence in chemistry or astronomy and facts and evidence in much of linguistics. One does not need to look closely to see that there is a wide discrepancy between the way linguistic theories are proven and the way theories in the empirical sciences are proven.

Linguistic theorizing is carried out and proven real in a very different manner than other sciences that claim to have mentalistic import. Therefore, when linguists are asked for convincing evidence for the psychological significance of their theories, they are in essence being asked for evidence that is recognizably psychological in nature, and is gathered by empirical means. Once these theories are determined, tested, and proven in such a way, there will be fewer calls for better evidence and fewer charges that the theories are not necessarily psychologically relevant, or in other words, that they are true outside their theoretical realm.

1.3 The relationship between formal and empirical analyses

Most contemporary analyses of language are formal. For the purposes of this book, formal analyses may defined as those that are primarily based on internal evidence, utilize autonomous methodologies, and are nonempirical in nature. Formal linguistics is a self-contained field that focuses on linguistic constructions and patterns themselves with little regard for the cognitive, social, and communicative functions of language. The structure of a language is considered of utmost importance rather that the use of language as a communicative tool. Explanations of linguistic phenomena are embodied in the formal notations of the model. Empirical approaches contrast with formal approaches in that they make use of both internal and external evidence, are empirical by definition, and are carried out with nonautonomous methods. Linguistic phenomena are explained by reference to language behavior in real time.

Many of those who argue that there is a difference between formal and empirical linguistics do so in order to illustrate how an empirical analysis is

more justified in making mentalistic claims than a formal analysis. Aside from that, few elevate one approach over another. In fact, many who make the distinction state that both types of linguistic research are useful and worthwhile and that there are important questions to be answered in both fields (Carr 1990, 2000; Hutchinson 1974: 73; Itkonen 1976: 6; Kac 1992: 54–57; Katz 1981, 1985; Lass 1976a: 220). However, the domains of both approaches are not kept separate. Researchers often attribute the same characteristics they assume an ideal speaker-hearer has to actual speaker-hearers; many also erroneously use performance-related data to justify models of competence and vice versa (Hale and Reiss 2000; Stemberger 1996).

The principal cause of the blurring of boundaries is the desire to provide a unified account of both the structure of a language and what knowledge speakers have of the language. This dual position is exemplified in Chomsky's writings. On the one hand, he asserts that generative linguistics is the study of abstract linguistic entities. At the same time, he proposes that it is the empirical study of human cognitive abilities (Katz & Postal 1991: 541–547; Olshewsky 1985). Katz (1985: 193) illustrates why this position is untenable:

> No one confuses psychological theories of how people make inferences with the logical theories of implication, or psychological theories of how people perform arithmetical calculations with mathematical theories of numbers. Yet, in the exact parallel case of linguistics, conceptualists do not make the distinction, conflating a psychological theory of how people speak and understand speech with a theory of the language itself.

Conflation of the two domains is responsible for the practice of carrying out an analysis in one approach, and making claims that correspond to the domain of the other. An example of how a phonological analysis is accomplished will be helpful in illustrating this point.

Suppose that a phonologist goes about analyzing a pattern found in a language within a formal approach. The phonologist, unlike the native speaker, may have knowledge of the history of the language, as well as what goes on in the phonologies of related languages. He or she is also aware of what kinds of patterns are common and uncommon in languages of the world, as well as what phones constitute natural classes. The task is undertaken in accordance with various formal principles; the phonologist seeks to provide an elegant and simple analysis that accounts for the greatest number of lexical items. An analysis that includes principles of universal grammar and has independent motivation is more highly valued than one that does not. An analysis of this type claims to account for the language of an idealized speaker-hearer.

Up to this point, the phonologists method is perfectly valid, carried out within a formal approach, and there has been no intermingling of formal and empirical domains. This analysis may legitimately claim to capture the phonological pattern or structure it intends to. As a result, it is a valid analysis of the language phenomenon. However, empirical claims are often ascribed to formal analyses. Any of the following claims, if used in conjunction with an analysis such as the one described above, would constitute an improper mixture of formal and empirical approaches: (1) the analysis is empirical; (2) the analysis is descriptively adequate, that is, it is not simply valid as a description of the data on which it was based, but it represents the knowledge that actual speakers have of their language.

The fact that an alternation, constraint, or structure can be found and described using formal methodology is not grounds for supposing that it is somehow represented in native speakers' minds, nor that they have any sort of conscious or tacit knowledge of it. Researchers have arrived at many creative and elegant analyses of linguistic phenomena, but in many cases it is still to be seen which of these are captured by native speakers, which are the leftovers of diachronic changes, and which have synchronic psychological relevance.

If formal and empirical approaches are both valid ways of approaching linguistic questions, why is it necessary to distinguish between the two? The answer is that invalid conclusions are reached when the two are confused (Yngve 1986, 1996). The first step towards resolving this problem is recognizing that there are indeed two approaches (Prideaux 1971: 346). One must make a choice between determining what would be the ideal system of an ideal speaker-hearer and determining what real speakers actually know about their language, as well as how they actually produce and comprehend it.[7]

Accordingly, a good analysis recognizes the domain to which it belongs, the methods and criteria that are valid in that domain, and does not make claims outside of its domain. For example, the results of a psycholinguistic probe into how native speakers utilize a certain phonological pattern should not have any bearing on what would be the most rigorous, concise, or elegant way to account for that pattern in a given formal framework. Conversely, the most intuitively valid formulae for describing a pattern are to be ascribed only to an ideal speaker-hearer, not necessarily to actual speakers of the language. Hale and Reiss (2000) for example, are among those who recognize and maintain a strict respect for the boundary between domains. In their approach to phonology, they explicitly reject any tangible substantive data such as phonetic and experimental evidence, since these have no relevance to their autonomous theory of phonology.

The emphasis on keeping formal and empirical approaches separate does not mean that there is not, or should not be, an interface between them. Theory that never goes beyond the theoretical stage is just as uninsightful as data gathering and experimentation carried out with complete disregard to theoretical underpinnings. Theory that is clearly stated should naturally lead to and precede empirical research. In the case of linguistics, formal theories may serve as the basis for empirical investigation (Baker 1979: 141; Black & Chiat 1981: 51–54; Carr 2000; Derwing 1979: 125; Kac 1980: 243; Pierrehumbert, Beckman, Ladd 2000).

1.4 Conclusions

Many linguistic theories claim to be relevant to linguistic cognition. However, the constraints, representations, derivations, modules, and parameters are not thought to mirror the actual algorithms that are used in the course of speech comprehension and production. Instead, they are considered to be abstract representations of speakers' underlying knowledge of their language.

There are two senses in which an analysis may be considered psychologically real. In the strong sense, it is real if the steps in the derivation, for example, have counterparts in actual language production or comprehension. This is not what most researchers interpret the steps in an analysis to represent. Therefore, an analysis may be psychologically real only in the weak sense of the word. A grammar is real in the weak sense if the outcome of the grammar corresponds to the output and intuitions of the speakers of the language. This means that orderings, constraints, and parameters are not real, only the structures and generalizations that they are designed to describe.

Although linguistic entities have the potential of being real in the weak sense, their psychological validity has been questioned on a number of different grounds. The first has to do with the truth versus reality issue, which questions the supposition that the best formal analysis of a phenomenon is necessarily a psychologically significant analysis. If the best formal theory is taken to be necessarily real, then there is no way to distinguish between theoretical constructs that may have psychological validity and those that are merely notational artifacts. An analysis may reveal many linguistic patterns, but their existence does not necessarily imply that they are significant for the speakers of the language.

A second reason why the reality of many analysis is doubted is that most contemporary approaches to linguistics are nonempirical. That is, they are not stated in such a way as to make them subject to potential falsification based on events that take place across space and in real time. The reality of entities

belonging to the nonempirical sciences is questionable in comparison to the reality of entities of the empirical sciences. This is true because the empirical sciences deal with entities that are proven and disproven with recourse to spatiotemporal events. Theoretical entities that escape possible falsification in space and time are not likely to exist in space or time either. Of course, the existence of a given linguistic phenomenon is falsifiable on the basis of a corpus of utterances. However, its psychological status must be determined separately.

In order to determine whether a linguistic phenomenon is psychologically valid it needs to be verified by psychological means. Therefore, the third reason for doubting the psychological validity of phonological theories is that many are established with little or no recourse to the speakers of the language via experimental psychology.

The fourth cause for skepticism is that most linguistic analyses are founded on an extremely narrow base of evidence. Internal evidence such as internal reconstruction, data gleaned from a corpus of utterances, and carefully monitored speech are the principal sources of evidence upon which contemporary analyses are typically based. External evidence, as found in speech errors, language games, and psycholinguistic experimentation, is often overlooked or at least not commonly sought. However, external evidence appears to be more telling of what speakers know about their language than internal evidence since it involves actual language use and manipulation. Therefore, a claim regarding the psychological significance of an analysis is strengthened if it is supported with external as well as internal evidence. Such evidence also avoids circular argumentation, which occurs when the observations that form the basis for assuming the existence of a theoretical entity are, at the same time, used as proof of the existence of the entity.

In summary, many linguistic analyses are assumed to be psychologically significant, but there are many grounds for challenging this assumption. This is not to say that some aspects of the theory may well be psychologically real. Linguistic analyses attempt to codify linguistic systems that exist in the minds of speakers and have somehow been codified by them. Therefore, it is highly possible that there is some correspondence between an analysis and the actual knowledge that speakers have about their language. The charge is simply that most are based on evidence that is not empirical and that is obtained by means which are not recognizably psychological. Once theories of language are established in this manner their psychological validity will be challenged less often.

Chapter 2

The role of experiments in linguistics[8]

2. Introduction

The goal of linguistic investigation has always been to discover and systematize the patterns and generalizations found in the systems of natural languages. This could be called the search for linguistic realities. Since the advent of generative linguistics the search for linguistic realities has been extended and the search for psychological realities rekindled following the behaviorist era.

In the previous chapter, I argued that linguistic analyses may adequately describe linguistic realities. However, they are not necessarily adequate descriptions of psychological realities as well. In this chapter, I discuss the role of experiments in the search for linguistic analyses that are relevant to linguistic cognition. I claim that a stronger case may be made for the psychological relevance of an analysis that includes experimental evidence. I will present and contest several objections to the experimental approach that have been voiced over the years. I will also review some examples of experimentally acquired evidence that bear directly on the issue of the psychological significance of certain linguistic analyses.

2.1 The role of experiments in the search for psychological realities

As discussed in Chapter One, there are several reasons to be skeptical of the psychological significance of many phonological analyses. However, there are three ways in which experimentally adduced evidence can aid in the search for psychological realities: (1) experiments provide empirical evidence; (2) experiments involve attempts to gain insight into the cognitive organization of language users by more direct means; (3) experiments help determine which linguistic realities are psychologically pertinent and which are not.

The goal of an experiment is to produce data that helps support or refute a hypothesis. Besides experiments there are other means of finding relevant data. Sociolinguistic interviews and corpus studies are among these. Data from these methods are just as useful in hypothesis testing as experimentally acquired data

are. On the one hand, language observed under more natural conditions is less likely to be tainted with the influences of the experimental procedure (see Section 2.3.4). On the other hand, natural evidence has a higher chance of being affected by factors that cannot be controlled for when compared with experimentally acquired evidence. In any event, my use of the term 'experiment' in this chapter can be extended to include evidence based on careful observations that are submitted to statistical analysis.

The linguistic realities (patterns, generalizations, constraints, parameters, etc.) that researchers have observed are based on human language; therefore, it is possible that they also have some relevance to the way people process language. However, the fact that an analysis yields a certain structure or pattern only demonstrates that it is available to be potentially known or used. In order to determine what is actually known or utilized by the speakers, the focus of the research must turn to the speakers themselves. Language is a human behavior and needs to be investigated as such (Tobin 1997).

It is highly possible that a linguistic structure may prove useful in systematizing a given language but may have no correlation in the minds of language speakers. A structure may be the result of a defunct historical change. Certain alternations may exist in the language, but may not be recognized or utilized by the speakers of the language, which would suggest that they play no role in language processing. I hold that evidence from psycholinguistic experimentation helps to differentiate between linguistic realities that are psychologically relevant, and those that are not.

The second advantage of experimental evidence is that it deals with events that take place through space and in time. A hypothesis stated in such a way that it is subject to experimental refutation is empirical. As mentioned before, there is a willingness to accept the reality of hypotheses that are supported by spatiotemporal evidence. By the same token, there is skepticism about the reality of hypotheses for which there is no spatiotemporal evidence. It is for this reason that well-designed experiments are thought to offer insight into human linguistic capacities. If one's goal is to study the variety of structures and patterns in languages, psychological experimentation is not essential. However, if one's goal is to study how humans produce and comprehend language, direct access to the speakers of the language is imperative.

2.2 Strong and weak reality

I would like to return to the notion of strong and weak reality that was defined in Section 1.1. The strong sense of reality implies a close relationship between

the formal mechanisms that the linguist proposes and the internalized representations and mental processes speakers use in comprehension and production. The weak sense involves little correspondence between formal constructs and psychological mechanisms. In many respects, human language processing can be considered a virtual black box. Since the box cannot easily be opened for inspection, the major clues to its contents come from its output. The weak reality approach to the black box involves determining *in principle* what the box might contain. If an analysis produces the same output as the box, then in a limited way it may claim to have achieved psychological validity (Rischel 1978: 442). The strong claim to reality, on the other hand, involves establishing *in fact* what the box contains.[9]

It is often uncertain what level of psychological reality phonological analyses strive to attain. At times, they are spoken of as if the formal apparatuses are step-by-step formulae for assembling or producing forms. Bromberg and Halle (2000: 35) certainly take this realist stance: "Do speakers *really* retrieve morphemes from their memory, invoke rules, go through all these labours when speaking? We think they do." If this is the case, the rules would be candidates for attaining reality in the strong sense. However, this is not the most widely accepted opinion. Most linguists do not generally define formal mechanisms in these terms. Chomsky and Halle (1968: 117) state that

> Although we may describe the grammar G as a system of processes and rules that apply in a certain order to relate sound and meaning, we are not entitled to take this as a description of the successive acts of a performance model.

This sentiment is evident in more contemporary linguistic thought as well. For instance Kager (1999: 26) observes that

> explaining the actual processing of linguistic knowledge by the human mind is not the goal of the formal theory of grammar … a grammatical model should not be equated with its computational implementation.
>
> (see also Bradley 1980: 38)

If formal rules do not reflect the mental algorithms that are used in speech and comprehension, the natural question is in what sense they are relevant to linguistic cognition. The current stance is that formal mechanisms such as rules somehow represent a speaker's tacit knowledge of the language – knowledge in the sense of being able to speak and comprehend the language. For example, in response to charges that phonological rules have no psychological validity, Kiparsky (1975: 198) states

> To deny that grammatical rules are utilized in speech behavior is not necessarily to deny their psychological reality … In phonology, the system of rules and underlying forms might be a representation of the speaker's *knowledge* of the systematic relationships among words in the language; not in any sense a mechanism which is applied whenever words are spoken and heard.

If formal linguistic mechanisms are defined in these terms, they cannot ever attain psychological reality in the strong sense because the strong sense implies that they correspond to mental operations and mechanisms. If linguistic analyses are merely abstract representations of a speaker's ability to speak and understand the language, instead of mirroring actual processes, they may only attain reality in the weak sense. This means that what is potentially real in an analysis is not the parameter settings, rules, constraints, intermediate derivations, or orderings, but merely "the function that these constructs serve to specify" (Matthews 1991: 197). Any analysis is real in the weak sense so long as it produces the same outcome as is produced by the speakers of the language, regardless of the way it goes about producing the output.

Misunderstandings about what level of reality linguistic analyses represent sometimes lead to confusion. For example, Derwing (1979: 114) charges that "a grammar that describes utterance forms can no more 'explain' them than can a description of a painting tell how the painting came about – or a 'grammar of a cake' tell how to make a cake." In other words, a description of some phenomenon is not at the same time an explanation. Derwing feels that linguistic analyses and the formalisms they employ should explain language production and comprehension. Fromkin (1980: 200) charges that this assumption is responsible for Derwing's misunderstanding of what rules and the like represent; they are not thought to explain language behavior, but are merely an abstract representation of the ability to use language. Most linguistic analyses do not claim to explain the actual processes involved in language perception and production, only to describe linguistic structure assuming that the structure is somehow relevant to actual processing.

In regards to Derwing's analogy between a linguistic analysis and painting a picture, it could be countered that a description of the structure, composition, and use of color and light in a painting does not explain how it was painted by the artist. However, it does say something about what factors the artist had to have in mind while producing the painting. It is in this abstract way that analyses correspond to speakers' linguistic abilities. Of course, some linguists (e.g., Bresnan & Kaplan 1982: xxii) perceive the search for the actual cognitive processes that underlie language use as a much worthier goal.

In summary, many researchers are interested in analyses of language that possess some relevance to the human psyche. Confusion often arises because formal mechanisms are often spoken of as if they represented actual steps in language processing. However, they are actually defined in other terms (Carr 2000). They are defined, not as linguistic algorithms, but as a more generic type of knowledge that underlies speakers' ability to speak and comprehend their language. An analysis that posits algorithms may justifiably profess to achieve psychological reality in the strong sense of the word. However, an analysis that deals in abstract representations may only hope to be proven real in the weaker sense of the word. I will return to the notion of strong and weak reality as it pertains to experimental evidence in a later section.

2.3 Criticisms of psycholinguistic experiments

What I have suggested is that experimental evidence has advantages over the more commonly utilized types of evidence in resolving issues of psychological reality. The experimental approach, however, has not been free from criticism. I review several of these criticisms in this section.

2.3.1 *Lack of adequate knowledge*

One misconception about experiments is the idea that they are valuable only in a field that is well developed, and about which much is known. For example, Matthews (1991: 190–191) is of the opinion that

> we know very little about the computational machinery involved in language processing. We are therefore not in a position to use experimental evidence regarding language processing.

Kac expresses similar sentiments. He concludes that since an adequate theory of linguistic structure has not been achieved, mingling psychology and linguistics is a wasteful and unfruitful endeavor (1974: 45–46, 1980: 243).

It is difficult to accept this line of reasoning. It is tantamount to refusing to perform experiments relating to subatomic particles on the grounds that so little is known about them. In the same way, it would be absurd for a psychologist to denounce the utility of psychological experiments simply because he or she felt that too little is known about perception, learning, or cognition. In actuality, much of what is known about physics and psychology is a direct result of theory-based experimentation and could not have been established in any other way.

In the empirical sciences, theory building and experimentation are inseparably connected. Extensive knowledge, well-developed and completely adequate theories are not prerequisites for experimental research. The only prerequisite is a hypothesis that is consistent with the bulk of the existing scientific knowledge and that is stated in such a way as to specify what outcome would support the hypothesis and what outcome would disprove it (Bunge 1980: 33). Once a hypothesis has been confirmed or refuted the theory is then modified, which in turn leads to better hypotheses and more experimentation. Analyses that are purely theoretical are just as meaningless as experimentation carried without a theoretical base. Theory development and empirical research must go hand in hand. The role of descriptive analysis is to provide insight into the structure of language upon which theories are built. In Kac's (1978: 155) words

> The question how a language is organized (which is the same as the question of what its structure is) is a different one from that of how a speaker comes to be able to use it – though in answering the first we contribute to some extent to the answer to the second since it is precisely the fact that languages have structure that renders them knowable and learnable in the first place.

In linguistics, descriptive analysis must precede empirical investigation (Baker 1979: 141; Derwing 1979: 125; Kac 1980: 243).

2.3.2 *Experiments and competence*

A common objection to experimental evidence is that it is not pertinent to the generative domain of enquiry. Kiparsky (1968: 174) phrases it in these terms:

> The fact that grammars are not performance models presumably means that the answer to the question of whether they are correct competence models is not likely to be forthcoming by any currently known experimental techniques until the contributions of competence can be separated out from the facts about performance.

In one respect, Kiparsky is absolutely correct. Competence is an idealized concept that comprises the system of formal mechanisms that are thought to underlie a speaker's ability to produce and understand language; performance is the actual realization of the speaker's linguistic ability (Chomsky 1980a: 205). As long as this view is maintained competence is effectively shielded from experimental probes and possible refutation. This is so because, according to this dichotomy, all spatiotemporal manifestations of language fall into the domain of performance, and as a result, can never be directly relevant to the study of abstract representations of linguistic ability that competence is thought to embody.

Suppose, for example, that two subjects react in a different manner to a question in a linguistic experiment. One could say that the subjects have the same underlying linguistic ability (competence), but that the experimental design was responsible for the differences. That is, something in the experiment kept the subjects' reactions from reflecting their competence (Wheeler 1980: 78–90). Therefore, it could be argued that the experiment must have measured performance instead. Since generative theories are theories of competence, and by definition the results of any experiment are indicative only of performance, experimental results are irrelevant to a theory of competence.

Derwing (1983: 66) demonstrates how a similar argument can effectively insulate a theory from any sort of counter evidence:

> Suppose we find some child who is quite adept at basic arithmetic. One possible hypothesis about the 'competence' thought to underlie this skill might be to attribute the child, not with something so mundane as a learned, laborious, step-by-step procedure for carrying out simple arithmetic operations, but rather with knowledge of number theory. And what if experimental results are found that seem to fly in the face of this hypothesis? Just chalk them up as 'performance errors' and the well-formed theory remains inviolate.

This scenario played itself out in early history of Transformational Syntax. According to this framework, certain sentences, such as positive statements, were seen as less complex than sentences such as passives that had to undergo transformations. A series of psycholinguistic experiments found no consistent correspondence between measured processing times and transformational complexity (Greene 1973). Rather than accept this as evidence against the notion of transformational complexity, it was dismissed as irrelevant. Testing formal theories by empirical means was thought to reflect a failure to distinguish language from thought and competence from performance (Smith 1999: 108). Experimental data measures performance, and Transformational Syntax is a theory of competence not performance.

Until competence is defined in such a way that it is subject to possible falsification, any empirically testable hypothesis will ultimately be a hypothesis about an aspect of performance. Therefore, researchers whose principle concern is psychological reality should be content to relegate competence to the domain of nonempirical science along with logic, virtue, number theory, and ethics, and focus on the reality of entities that emerge through spatiotemporally observable performance.

2.3.3 *Negative experimental results*

Kiparsky (1975) voices another objection to the use of experiments. In this case, it has to do with the validity of negative experimental results. He accepts the results of production and perception tests, but only insofar as they produce evidence in favor of the psychological significance of an analysis. According to him, negative evidence is always inconclusive. His justification for this reasoning is this: If an archaeologist finds bones at an archaeological dig, the existence of those bones provides positive evidence that a certain animal inhabited that area. However, if no bones are found, that is not evidence that the animal in question never lived there.

This illustration is objectionable on two counts. In the case of bones found at an archaeological dig, Kiparsky's logic is sound. However, this does not imply that negative evidence is irrelevant to the testing of a hypothesis in all fields of research. For example, the claim that cold nuclear fusion had been produced in the laboratory was corroborated by some experiments and refuted by others. In the end, the claim was rejected on the basis of the negative evidence presented.

Second, Kiparsky's argument cannot be logically extended to physics nor to phonology because archaeologists look for evidence of things that *used to exist* in a given place. Physicists and mentalistic linguists seek evidence about things that theoretically exist *in the present*. Fortunately, not all linguists are willing to dismiss negative evidence so quickly. For example, Mohanan (1986: 58–59) suggests that if there is abundant evidence that a rule does not play a part in storage, recognition, or production, it should not enter into a description that professes psychological significance.

2.3.4 *Experiments and external validity*

Another criticism of psycholinguistic experiments is what Kiparsky and Menn term the "strangeness effect" (1977: 63–64). This suggests that the unusual circumstances that are involved in obtaining experimental evidence influence the subjects to give unusual responses. In other words, the experimental situation causes the subjects to answer in ways they would not under normal circumstances. The strangeness effect is known in science as threats to the external validity of an experiment. External validity involves the extent to which experimental results can be considered valid outside of the experimental setting.

An experiment in any science must try to control for external validity. It is a factor in any field that employs experimental paradigms. One way to control for external validity is to compare the results of experiments performed in the laboratory with results obtained from sociolinguistic studies in which emphasis is put on data gathering in a more naturalistic setting. External validity may be

checked in another way also. Ideally, a hypothesis about psycholinguistic functions should be tested with a number of distinct experimental paradigms. In this way, influences that are attributable to the experimental situation may be factored out. If several different experiments yield similar results, the results are more likely to be due to actual mental processes, not to something inherent in one particular experimental design (Derwing 1979; Pierrehumbert, Beckman, Ladd 2000; Schütze 1996).

2.3.5 *Conflicting experimental results*

The fact that experiments often produce conflicting results could be adduced by a skeptic to argue that experiments are ineffective in deciding questions of psychological reality. This position is clearly untenable. Conflicting experimental evidence is a reality for all fields of science that incorporate experiments. The mere existence of conflicting evidence is not grounds for abandoning the experimental approach (Schütze 2003). On the contrary, it should compel researchers to refine their experimental methods as well as to explore others. It should lead to closer inspection of the phenomenon under investigation as well as the experimental means used to probe it.

2.3.6 *Experiments and strong reality*

Ultimately, the claim that an analysis is real in the strong sense means that it relates to actual mental mechanisms. Since little is currently known about the workings of the brain at this level, it may seem impossible for any analysis to achieve strong reality. However, this is not the case. Consider the work on the processing of regular and irregular past tense verbs in English.

Two theories exist. One suggests that different mechanisms are responsible for regular and irregular inflection (Marcus et al. 1993; Pinker 1991; Pinker & Prince 1988, 1994; Prasada & Pinker 1993). Others contend that regular and irregular past tense forms are produced by the same mechanism (Bybee 1985, 1988, 1995; Daugherty & Seidenberg 1992, 1994; Eddington 2000a; Seidenberg 1992; Stemberger 1994). Both of these hypotheses lend themselves to possible empirical refutation which makes them viable candidates for strong reality. If one area of the brain is consistently activated when processing regular forms and another when processing irregular forms, that would constitute evidence that regular and irregular forms use different mental mechanisms. This is actually the goal of several recent studies (e.g., Jaeger et al. 1996).[10]

2.3.6.1 *Experiments, strong reality, and competing analyses.* As already discussed, most analyses do not purport to represent mental mechanisms, which

prevents them from attaining strong reality. If one were to choose a given linguistic phenomenon it would be a simple task to find a number of analyses carried out within different frameworks that claim to account for the phenomenon in question. In many cases, each of the analyses would produce the same outcome, but they each assume a differing mechanism in order to yield the output. Often, each successive analysis is regarded as superior to its predecessors. The question is whether experimentation can help decide between such competing analyses.

Most psycholinguistic evidence involves this assumption: A theory postulates certain mental mechanisms, therefore, if the results of an experiment correspond in a statistically significant way to the hypothesized mechanisms, that constitutes evidence in favor of their reality. Botha (1971:128–30) challenges the validity of this assumption:

> The fact that a theory correctly *predicts* some events does not necessarily imply that it also correctly *describes* or *represents* the structure of the mechanism from the operation of which the predicted events result. The predictive and representative or descriptive functions of a theory are distinct, and achievement of success in one of them does not necessarily imply that success has been achieved in the other as well. [Emphasis is original]

In other words, if the human mind is viewed as a black box, any number of different mechanisms may be responsible for producing the output of the box. That experimental evidence supports one hypothetical mechanism is no guarantee that a distinct mechanism is not actually responsible for the output (Botha 1971:131–135). Therefore, if two distinct mechanisms are postulated that would produce the same output, psycholinguistic evidence would be incapable of determining which mechanism is real (Fromkin 1975:56).

Botha and Fromkin are correct in this regard. As already discussed, a linguistic analysis that does not claim to mirror actual mental processes has the potential of relating to psychological mechanisms only in the weak sense. In this sort of analysis, only its substance may be proven to have some significance for speakers. Consequently, any and all analyses that make the same predictions are significant in the weak sense if those predictions are borne out.

For many linguists, it is frustrating that psycholinguistic experimentation is unable to decide between competing analyses of the same phenomenon. Much of the contemporary work in linguistics centers on demonstrating the superiority of one analysis of a given process over another. Therefore, it appears that if experimentation is incapable of settling the issue of which analysis is most correct it is of little value. An example from Spanish will clarify this point.

2.3.6.2 *Spanish diphthongization and strong reality.* The alternation between the unstressed mid-vowels [e, o] and the stressed diphthongs [jé, wé], has been accounted for in several different analyses. A crucial part of these analyses is that they distinguish between unstressed mid-vowels that alternate with stressed diphthongs (e.g., *c[o]ntámos* 'we count,' *c[wé]ntan* 'they count') and those that do not (e.g., *t[o]sémos* 'we cough,' *t[ó]sen* 'they cough').

A number of ways of distinguishing diphthongizing from non-diphthongizing mid-vowels have been proposed. Harris (1969:74–75; 1977) claims that the mid-vowels that do not undergo diphthongization are represented underlyingly as /o/ and /e/. Those that are transformed into diphthongs appear in the deep structure with the diacritic feature [D]. St. Clair (1971:421) utilizes a tense-lax distinction in his analysis. The lax vowels, which he transcribes as /O, E/, become diphthongs, while the tense vowels, /o/ and /e/, do not. Hooper's analysis (1976:157–160) is different in that it does not transform vowels into diphthongs. Instead, in stem morphemes that demonstrate the alternation, both the mid-vowel and the diphthong are listed disjunctively. For example the verb *contar* 'to count' is represented as in (1):

(1) /k$\left\{\begin{smallmatrix} o \\ we \end{smallmatrix}\right\}$nt-/

As a result, only mid-vowels that are disjunctively listed with a diphthong alternate. In a later analysis, Harris (1985a:31) suggests that the diacritic that marks mid-vowels that undergo diphthongization is part of the syllable structure.[11] According to this view, only mid-vowels that appear adjacent to an empty prosodic slot trigger the formation of diphthongs.

The four formal representations of the stems *cont-* 'to count' and *tos-* 'to cough' are listed in (2) for comparison:

(2) a. Harris 1969, 1977 /k$\left[\begin{smallmatrix} o \\ +D \end{smallmatrix}\right]$nt-/ /t$\left[\begin{smallmatrix} o \\ -D \end{smallmatrix}\right]$s-/

b. St. Clair /kOnt-/ /tos-/

c. Hooper /k$\left\{\begin{smallmatrix} o \\ we \end{smallmatrix}\right\}$nt-/ /tos-/

d. Harris 1985a

/k o n t-/	/t o s-/
| | | | |	| | |
x x x x x	xx x

Each of these analyses express the generalization that some mid-vowels, but not all, alternate with diphthongs, and each effectively distinguishes diphthongizing from non-diphthongizing mid-vowels. In other words, they each embody the substance of the alternation. At this level, the claim that diphthongization

is psychologically valid in the weak sense means that it is in some way relevant to the language faculties of Spanish speakers.

Proving one of these analyses is more psychologically valid than the others would involve more than confirming the reality of the substance of the alternation. It would include confirming the reality of the actual form of the representation used in the analysis. Although the *substance* of each analysis is the same, (i.e., some mid-vowels alternate with diphthongs), the *form* of each one is distinct. For instance, in Harris's earlier analyses the diacritic [D] is crucial, while in St. Clair's the abstract feature [lax] is. Any strong claim to the reality of an analysis at this level is necessarily a claim that the actual form of the representation has a correlate in speakers' minds. Therefore, an attempt to prove that one analysis is true and another false would necessarily entail demonstrating the psychological reality of notational elements such as empty prosodic slots or features such as [D] or [lax].

2.3.6.3 *Concepts and strong and weak reality.* The notational elements of linguistic investigation are conceptual elements on a par with the notational elements of logic and mathematics. Concepts are created fictions that are defined and granted existence by those who use them; they may be useful within the domain in which they have been created, but unlike physical elements they do not exist outside of the conceptual domain (Bunge 1980).

A number of linguists have related linguistic representations to maps (Harmon 1980: 21–22; Matthews 1991: 196). This is a fruitful analogy. Suppose that in addition to latitude and longitude, other notational devices have been invented that also locate topical features on the Earth. It could be argued that one device does the job better or is more elegant or precise than another, but, it would be absurd to assert that one is real while the others are not. The islands, oceans, and rivers themselves are real. The devices used to locate them are not real, but are simply extremely useful tools. So it is with the theoretical notations of linguistic analyses – they are convenient fictions.

The fact that notational elements belong to the conceptual realm means that they are not subject to empirical testing. In other words, it would be impossible to discover empirical evidence to support the psychological reality of one of the analyses of Spanish diphthongization in (2) over another. What spatiotemporal event would argue in favor of the feature [D] over the disjunctive ordering of vowels and diphthongs? Would it be possible to formulate a hypothesis about the existence of empty prosodic slots in such a way that it would be clear what evidence would disprove or support them? (e.g., if we find X, we

know that diphthongizing vowels appear next to an empty prosodic slot.) Any attempt along these lines would be ultimately futile.

Botha and Fromkin are correct in regarding experiments as an unsatisfactory means of distinguishing between competing analyses that are notational variants. Empirical data is relevant only to empirically confirmable entities. Since the notational elements of these analyses belong to the conceptual realm they simply cannot be proven by empirical means such as experimentation. Therefore, the charges that theoretical problems cannot be solved via experimentation are essentially complaints that experiments cannot resolve nonempirical questions.

In formal linguistics, the majority of the evidence presented to demonstrate the superiority of one analysis over another is nonempirical (e.g., appeals to universal grammar, simplicity, elegance, generality, etc). One analysis may indeed be proven superior to another in the conceptual realm (see Section 1.2.2), but it does not follow that the superior analysis more closely models the mental processing of language that exists in the empirical/physical world. Unfortunately, linguists often make the mistake of confusing the empirical and conceptual domains (see Section 1.3).

Can experimental evidence decide between competing analyses of the same phenomenon? Can it be used to answer questions about mental processing? Yes, but the analysis must be based on an empirically testable theory of actual language processing, not merely a conceptual theory of language structure. The reality of an analysis in the weak sense may be verified or refuted experimentally. This entails determining whether the substance the analysis captures is psychologically significant. In this case, if experimental results yield an output similar to those predicted by a hypothesized mechanism then the existence of some mechanism is supported. The results do not indicate what the form of the mechanism is, only that speakers appear to have or not to have internalized some mechanism that generates the output (Cutler 1979: 79; Steinberg 1975: 218–220). In other words, experiments are effective in determining which analyses have mental significance and which do not. Although it is more difficult, it is also possible to verify the strong reality of an analysis. The studies on the processing of the English past tense cited in Section 2.3.6 demonstrate this possibility.

2.4 Examples of experimentally acquired evidence

2.4.1 *English ng and lexical phonology*

Nonce word experiments have the ability to answer some questions about the psychological validity of grammars. For example, Schlenck (1988) carried out a nonce word experiment in English to test a hypothesis of lexical phonology. Lexical phonology hypothesizes that certain phonological processes are related to certain groups of affixes. In English, the combination of letters *ng* is pronounced either [ŋg] or [ŋ]. With a few exceptions, [ŋg] appears before affixes such as *-ation, -ize,* and *-er*. These are called Class I affixes. On the other hand, before Class II affixes such as *-ing, -ly,* and *-less, ng* is pronounced [ŋ], as seen in (3):

(3)	a.	Class I aff.	*-ation*	*prolongation*	[ŋg]
			-er	*stronger*	[ŋg]
	b.	Class II aff.	*-ing*	*prolonging*	[ŋ]
			-ly	*strongly*	[ŋ]

The empirical question is whether this analysis merely describes a linguistic reality in English or whether it represents a psychological reality as well.

In Schlenck's study, subjects were recorded as they read a fairy tale that included a number of nonce words containing *ng* followed by different affixes. The nonce words were then transcribed phonetically. The predictions inherent in lexical phonology were borne out. Before Class I affixes, [ŋg] appeared significantly more often than [ŋ]. Before Class II affixes, [ŋ] appeared significantly more often than [ŋg]. The subjects' ability to use [ŋ] or [ŋg] in the appropriate context implies that they have some sort of knowledge of the distributional pattern of these phones in English and put that knowledge to use.

2.4.2 *The English vowel shift*

Perhaps one of the most experimentally tested rules in English is the English vowel shift. The vowel shift rule is designed to account for the following vocalic alternations (Chomsky & Halle 1968: 50–55):

(4)	a.	[aj]∼[I]	*divine-divinity*
	b.	[ij]∼[ɛ]	*serene-serenity*
	c.	[ej]∼[æ]	*sane-sanity*
	d.	[æw]∼[ʌ]	*profound-profundity*
	e.	[uw]∼[ɑ/ɔ]	*lose-lost*
	f.	[ow]∼[ɑ]	*verbose-verbosity*

In 1977, Halle reformulated the rule so that it incorporated another vocalic alternation (614–618):

g. [uw]~[ʌ] *reduce-reduction*

However, this reformulation of the rule excluded the alternation between [uw] and [ɑ/ɔ]. In 1985, Halle and Mohanan added another alternation to the vowel shift rule (72–9):

f. [ɔj]~[ʌ] *destroy-destruction*

Each pair of vowels is thought to be derived from a single underlying long vowel that never surfaces. For example, the alternation between [ij] and [ɛ] stems from an underlying /e:/. Although these alternations arose from historical processes, they are asserted to constitute part of the internalized mental grammar of Modern English speakers.

A great deal of experimentation has been done in order to investigate the psychological status of the vowel shift rule (see Jaeger 1986; Wang & Derwing 1986 for summaries). The results of these experiments have led to a greater understanding of the vowel shift as well as to a better understanding of how to tap speakers' linguistic knowledge experimentally. Experimentation into the vowel shift has helped to determine what kinds of experiments provide insight into the mind. One of the earliest experimental techniques applied to the vowel shift question was morpheme combination. In this type of experiment, the subjects are given a word and a suffix and then asked to combine them to construct a new word (Myerson 1976; Ohala 1974; Steinberg & Krohn 1975). For instance, subjects are asked to combine *maze* and the suffix *-ic* to form the word *mazic*. The outcome that would be predicted by the vowel shift rule is [mejz] > [mæzIk], but the most common way subjects handled these questions was to leave the vowel unchanged (i.e., [mejz] > [mejzIk]). These negative results were at first regarded as evidence against the psychological reality of the vowel shift rule.

Although morpheme combination experiments do not support the vowel shift rule, other experimental methods have provided some positive evidence (e.g., Eddington 2001b). Further evidence in favor of certain aspects of the vowel shift rule has been found through a variety of experimental procedures. Limited positive evidence was found by means of preference experiments, learning experiments, memory experiments, and concept formation experiments (Jaeger 1986: 88–90).

In other words, morpheme combination experiments produced negative results while many other methods yielded positive ones. This finding reveals

less about the vowel shift than it does about the validity of morpheme combination experiments. It seems that morpheme combination experiments are incapable of tapping into psycholinguistic abilities while other methods are. The realization that morpheme combination experiments are poor experiments, at least in the case of the English vowel shift, came about as a direct result of conducting further psycholinguistic research into the vowel shift question, even when the initial experiments did not prove fruitful. This demonstrates that experimental inquiries into a subject not only provide a better understanding of the subject itself, but they also provide a better understanding of experimental techniques. It is in this way that the experimental approach is able to hone and refine its methods.

Experiments into the vowel shift have also indicated the degree to which the rule is psychologically valid. The most significant finding is that not all of the eight proposed vowel shift alternations are supported. It appears that only five of the eight alternations have psychological significance for linguistically naive English speakers. The bulk of the data supports the validity of the alternations [aj]∼ [I], [ij]∼ [ɛ], [ej]∼ [æ], [ow]∼ [ɑ], and [uw]∼ [ʌ], but not that of [æw]∼ [ʌ], [uw]∼ [ɑ/ɔ], and [ɔj]∼ [ʌ] (Jaeger 1984, 1986; Wang & Derwing 1986, 1994). This means that none of the proposed vowel shift rules in (4) correctly groups the psychologically significant alternations together, and at the same time excludes the insignificant ones. Therefore, the psychological validity of any of the proposed vowels shift rules, as they are currently formulated, is dubious.

As is often the case, the answer to one question provokes the formulation of another one. In the case of the vowel shift, the question that arises is what the five psychologically valid alternations have in common. Jaeger (1984, 1986), Moskowitz (1973), and Wang and Derwing (1986, 1994) suggest that the five significant alternations correspond to spelling rules. When children are taught to read English they are told that each of the five vowels has a short and a long sound. The five significant vowel shift alternations correspond exactly to the short and long varieties of the five written vowels:

(5) *i* [aj]∼[I]
e [ij]∼[ɛ]
a [ej]∼[æ]
o [ow]∼[ɑ]
u [uw]∼[ʌ]

Of course, this claim constitutes a hypothesis that has yet to be tested. It would be possible to replicate the tests using illiterate English speakers as subjects.

Negative results to the experiment would corroborate the spelling rule hypothesis. A positive outcome, on the other hand, would suggest that the vowel shift alternations are not rooted in spelling rules.

2.5 Conclusions

The purpose for citing these experiments has been to demonstrate the utility of experiments in shedding light on theoretical issues. As a result of these experiments, much more is known about the psychological role certain alternations have for linguistically naive speakers. It is essential to note that this knowledge could not have resulted from a nonempirical study of language internal data.

Of course, experiments have their difficulties. First of all, they are more difficult to carry out than paper-and-pencil formal approaches, which makes it appear that the field of empirical linguistics progresses slowly (Ohala 2003). Secondly, experiments may not be used to decide between competing conceptual analyses that are essentially notational variants of each other; however, they may provide clues about the mentalistic import of conceptual analyses in the weak sense. Proving the reality of an analysis in the strong sense is also possible but much more difficult because it also requires a hypothesis that lends itself to confirmation on the basis of spatiotemporal evidence.

One advantage of experimentation is that it is not linked to a formal paradigm that gets replaced every few years. This means that the results of a quantitative study may be relevant to the field for many years (Ohala 2003). Of course, the results of psycholinguistic experiments should always be evaluated carefully. Definitive conclusions about the mental capabilities of language speakers should not depend solely on the outcome of one experiment. The results of any one study are always tentative and inconclusive. Experiments must be replicated, and this will often yield conflicting results.[12] For this reason, judgment should be suspended until a large body of evidence from varying sources has been accumulated.

Chapter 3

Testing untested notions

3. Introduction

In the linguistic literature, it is fairly common to find analyses of linguistic phenomena that are based on only a handful of examples. Evidence for syntactic theories, for instance, is often given as short lists of sentences that are marked as grammatical and others marked ungrammatical with the ominous asterisk. Often, the researcher him or herself decides what to accept as grammatical and ungrammatical. This along with the paucity of examples in many analyses has led to skepticism about the validity of the phenomena studied. For this reason, some investigators have centered their efforts on considering a wider range of data and testing the intuition of a larger number of speakers. The aim of the present chapter is to examine several notions that have gone untested, or were at best poorly tested in the beginning.

3.1 Vowel opening in the wake of s-deletion

One of the most studied phonetic processes in Spanish is what is known as *s*-aspiration[13] and deletion. The phoneme /s/ is subject to being pronounced as the aspirate [h] or is deleted altogether. This is most common in word-final and syllable-final position. Hence, *esto que es* 'what is this' may be rendered [ehto ke eh] or [eto ke e]. The degree to which aspiration and deletion occur varies geographically and is governed by a number of sociolinguistic factors.

In word-final position, the functional yield of /s/ is great. It distinguishes between singular and plural (*mesa*∼*mesas* 'table∼tables') as well as between second person singular and third person singular verbs (*vienes*∼*viene* 'you come∼s/he comes'). If /s/ is realized as [h] in these contexts, lexical distinction is maintained, but if it is deleted it could lead to confusion. A number of scholars suggest that if the preceding vowels are /a e o/, when /s/ is deleted they are realized as the more open vowels [ɑ ɛ ɔ] (Alonso, Zamora, & Canellada 1950; Cassano 1972; Honsa 1965; Hooper 1978; Navarro Tomás 1966; Salvador 1977;

Zamora Vicente 1974). Accordingly, the distinction between *tú pones*∼*él pone* 'you put∼he puts' is maintained in terms of vowel closure [tu ponɛ∼el pone].

A close reading of the primary literature on the subject shows that the open/closed vowel contrast described is not based on spectrographic analysis but apparently on the authors' own ear. Hammond (1978) devised a study to more accurately test the quality of vowels preceding a deleted /s/. The participants in his study were speakers of the Miami-Cuban dialect. Hammond recorded four speakers reading a series of sentences such as *Ayer pintaron las casas* 'yesterday they painted the houses,' and *Ayer pintaron la casa* 'yesterday they painted the house.' From these recordings certain words were chosen for comparison. For example,

Written word	Phonetic realization	Gloss
casa	[kasa]	house
casas	[kasaø]	houses
casas	[kasas]	houses
pescado	[peøkaðo]	fish
pecado	[pekaðo]	sin

The vowel qualities of words such as [kasaø], [kasa], and [kasas] were compared. Hammond found that in some instances the second and third formants indicated that a deleted word-final /s/ did co-occur with a slight opening of the preceding vowel. However, this phenomenon was not consistent. In some instances, vowels preceding a deleted /s/ were also found to be slightly longer in duration, but again it was inconsistent.

A different pattern emerged when the deleted /s/ in question appeared word-internally as in word pairs such as [peøkaðo] and [pekaðo]. In these pairs, there were no consistent differences in vowel quality. However, vowels preceding deleted /s/ (e.g., the /e/ of [peøkaðo]) were all longer than those that did not precede an /s/ (e.g., [pekaðo]). This supports the idea of compensatory lengthening; when an element is deleted its length remains but is transferred onto an adjacent segment.

What is odd about this finding is that word-final /s/ plays a functionally important role in Spanish morphology in distinguishing plurals and certain verbal forms. Word-internal /s/, on the other hand, serves to differentiate only a relatively small number of lexical items from each other. Nevertheless, the word-internal distinction is maintained via compensatory lengthening while the seemingly more important word-final distinction is not. In a study of Puerto Rican Spanish, López Morales (1989) also found that the morphological status of word-final/s/ has no bearing on its propensity of being deleted.

In addition to making measurements of vowels before deleted /s/, Hammond also performed a perception experiment with the same materials he examined spectrographically, that is, the sentences produced by native speakers. Other speakers of Miami-Cuban Spanish listened to these recordings and were asked to choose which sentence or word they had heard. For example they heard sentences such as:

Es importante que [salɣa] 'It's important for him/her/you to leave'
Es importante que [salɣas]
Es importante que [salɣaø]

Their task was to choose *salga* or *salgas* given to them in written form. They also heard individual words such as [peøkaðo] and [pekaðo] and chose either *pecado* or *pescado*. The subjects were able to correctly differentiate between cases involving word-internal /s/ at a rate of about 92%. However, only about 58% of the word-final cases were correctly chosen, which is only slightly better than 50% random selection. Of course, we know that in these data an /s/ that has been deleted word-internally lengthens the adjacent vowel and thereby gives a crucial clue as to the identity of the word. The same does not occur consistently with word-final /s/. The word-internal vowel length clued the speakers in to the correct identity of the words.

Hammond originally set out to test the reality of vowel opening in conjunction with /s/ deletion. While he found no consistent evidence for that process, he did unwittingly discover something else, namely, that in certain contexts speakers can differentiate between words such as *pescado* and *pecado* even when no /s/ is present. However, the distinction is not one of vowel quality but of vowel duration. One must be careful not to draw unwarranted conclusions from this study. The fact that vowel opening in the wake of /s/ deletion was not found in these particular speakers of Miami-Cuban Spanish cannot be construed to mean that opening may not occur in other dialects.

Figueroa (2000) expanded the evidence base on vowel opening by performing an experiment using Hammond's methodology. In the perception study, she found that Puerto Ricans were not able to distinguish between words such as *ves* [beø] 'you see,' and *ve* [ve] 's/he sees.' However, word-internally, they were able to differentiate between words such as *pastillas* [paøtiy̆as] 'pill' and patillas [patiy̆as] 'sideburns.' Her acoustic study revealed that vowels preceding deleted /s/ were consistently longer than vowels not preceding deleted /s/, but only word-medially, not word-finally. No consistent evidence for changes in vowel quality arose in any position.

The issue of vowel opening was further explored in the Eastern Andalusian dialect of southern Spain by Martínez Melgar (1986). Her subjects were asked to name objects in a picture book in the context of a phrase. For instance, they had to insert the words *patata* and *patatas* 'potato/es' in the sentence *Veo una/dos* _____ *en la foto* 'I see one/two _____ in the picture.' Judging by the spectrograms given, all of the plural forms she inspected lacked either final [s] o [h] and therefore ended in open syllables. She compared the final vowels of the plural and singular forms and found that, with the exception of /a/, the final vowels of the plural forms were longer in duration. The vowel quality differed somewhat but not in any systematic way.

While most studies failed to find the hypothesized vowel opening, three more recent studies do provide some limited evidence for its existence, at least in Eastern Andalusian. In separate experiments (Llisterri & Poch 1986; Martínez Melgar 1994; Sanders 1998), vowel formats of singular and plural nouns were contrasted. Speakers of that region were asked to pronounce singular and plural nouns in contexts such as: *digo* _____ *por ti* 'I say ___ for you,' and *uno/dos* ___ *pequeño/s* 'one/two small ___.' Llisteri and Poch observed a systematic opening of stressed and word-final /o/ and /e/ in the plural forms. Sanders recorded opening of /a/, /o/, and /e/ in his three subjects. Martínez Melgar's results are based on the speech of 91 subjects. She also found systematic opening of the mid-vowels in plural forms when contrasted with singulars. In some phonetic contexts /i/ and /u/ were observed to be more open in plurals, but to much smaller degree than the mid-vowels.

Given the present state of research into vowel opening in the wake of the deletion of /s/, two conclusions may be reached. First, this process does not occur in every Spanish dialect that exhibits aspiration and deletion. It appears to be limited to Eastern Andalusia. If one accepts the theory that American Spanish has its roots in Western Andalusia it could explain why this opening process has not been documented in America.[14] Second, even in dialects where it does occur it does not affect all of the vowels equally; it seems to be most prevalent among the mid-vowels. The results of these experiments tell us that researchers should be wary of making sweeping generalizations based on their own intuitions or informal observations. The postulation of a process should always be accompanied by valid supporting evidence.

3.2 Secondary stress

Given a multisyllabic word in Spanish, there is generally one syllable that is more prominent than the rest. This prominence is referred to as primary stress.

Acoustically, the syllable that receives primary stress in Spanish tends to have a rise in pitch in comparison to other syllables in the word, tends to be longer in duration, and to a lesser extent, tends to be produced with more volume (Enríquez, Casado, & Santos 1989; Quilis 1971). While the existence of primary stress is generally accepted in Spanish, whether some syllables receive a secondary stress is debated. The existence of secondary stress sets up a tripartite stress system in which syllables can be unstressed, have primary stress, or have secondary stress.

A number of positions have emerged concerning secondary stress. Quilis (1981), for instance, denies that secondary stress exists. Stockwell, Bowen, and Fuenzalida (1956) claim that the only place secondary stress exists is on word-initial syllables, hence *cìvilización, òbligacionísta*. Harris (1983, 1991b), Navarro-Tomás (1957), and Roca (1986) discuss a more complex system, the rhythmic hypothesis, in which stressless syllables alternate with stressed syllables: *òbligàcionísta*, *cìvilìzación*. Although Roca, Navarro-Tomás, and Harris' systems are more involved than these examples demonstrate, they differ from the other two positions in that more than one secondary stress may occur per word.

Prieto and van Santen (1996) set out to test the three positions described above. They recorded a Mexican speaker pronouncing triads such as *número/numéro/numeró* 'number, I number, s/he numbered,' *míntola*[15] */tintéro/tintoréra* '?, inkwell, species of shark,' and *horóscopo/escapó/escápo,* 'horoscope, s/he escaped, I escape.' All of the words were embedded in carrier phrasees. In this way, they could measure differences in the syllables [nu], [tin], and [po] in each triad. They measured the duration of syllables with alleged secondary stress and compared them to those that are unstressed or have primary stress. They found no differences in length between stressless syllables and syllables with secondary stress in either word-initial (e.g., ***tìn**toréra*), word-final (e.g., *horósco**pò***), or word-medial position (e.g., *dèsem**pòl**vorízo*). However, syllables receiving primary stress were longer in duration.

Prieto and van Santen then measured the fundamental frequency and amplitude peak of the syllables and found that syllables with primary stress were characterized by higher peak amplitude and a rising pitch. Secondary stress on word initial syllables was marked by higher amplitude than all other syllables except the syllable with primary stress. It was also marked by a downward sloping pitch. In other words, evidence was found for secondary stress in the first syllable of words as asserted by Stockwell, Bowen, and Fuenzalida (1956). That is, the first syllable of a word such as *dèsempòlvorízo*, 'I remove the gunpow-

der,' carries a secondary stress, but the third syllable does not, as the rhythmic hypothesis predicts (Harris 1983, 1991b; Navarro-Tomás 1957; Roca 1986).

Díaz-Campos (2000) was also interested in the existence of secondary stress. In his study, he focused on the syllables [de] and [po] in the triad *depósitò/dèposíto/depòsitó* 'deposit, I deposit, s/he deposited.' The syllables [es] and [ti] were compared in the test words *estímulò/èstimúlo/estìmuló* 'stimulus, I stimulate, s/he stimulated.' This method allowed comparison of primary stress, secondary stress, and lack of stress in the same syllable (i.e., [tí] [tì] [ti]). The test words were presented to six female speakers of Peninsular Spanish in carrier sentences such as *necesitas estímulo grande* 'you need a big stimulus.'

A comparison of syllables with primary stress, secondary stress, and no stress revealed that there were no consistent differences in amplitude or fundamental frequency among the three. The differences were all related to duration. Primary stressed syllables are longer than both stressless and secondary stressed syllables. What is more, there were no differences between the length of secondary and stressless syllables, which puts doubt on the existence of secondary stress.

In summary, Díaz-Campos finds no evidence for secondary stress, while Prieto and van Santen find it only in word-initial syllables. A number of reasons for this discrepancy come to mind. Secondary stress may be a salient factor in Mexican Spanish but not in Peninsular Spanish. It may play a part in the speech of some individuals but not in others regardless of their region of origin. In any event, no broad generalizations about the existence of secondary stress in Spanish are warranted by the data presented to date. This should drive researchers to explore this intriguing area of Spanish phonetics in more depth.

3.3 Coronal and velar softening

Formal approaches to linguistics tend to be extremely thorough. Formalists are very successful at hunting down every conceivable generalization in a given language and exploiting every possible relationship between words and phrases. The question that arises when one moves to the realm of performance is whether all of the generalizations linguists account for are also relevant to how native speakers process language. One such generalization is embodied in coronal and velar softening. The question posed in this section is whether coronal and velar softening play a role in synchronic processing.

Coronal softening is evident in a number of morphologically related words. It involves an alternation between the coronal consonants /t/ and /d/[16] and the fricatives /s/ and /θ/.[17]

/t/ ∼/θ/	*inyec/t/ar*	'to inject'	*inyec/θ/ión*	'injection'
	Mar/t/e	'Mars'	*mar/θ/iano*	'Martian'
/t/ ∼/s/	*emi/t/ir*	'to emit'	*emi/s/or*	'emitter'
	perver/t/ir	'to pervert'	*perver/s/o*	'perverted'
/d/ ∼/s/	*alu/d/ir*	'to allude'	*alu/s/ión*	'allusion'
	exten/d/er	'to extend'	*exten/s/ivo*	'extensive'
/d/ ∼/θ/	*aba/d/*	'abbot'	*aba/θ/ial*	'abbatial'

Velar softening is an alternation between the velar consonants /g/ and /k/, and the fricatives /θ/ and /x/.

/g/ ∼/θ/	*distin/g/uir*	'to distinguish'	*distin/θ/ión*	'distinction'
	grie/g/o	'Greek'	*gre/θ/iano*	'Grecian'
/g/ ∼/x/[18]	*ma/g/o*	'magician'	*ma/x/ia*	'magic'
	conyu/g/al	'marital'	*conyu/x/e*	'spouse'
/k/ ∼/θ/	*Costa Ri/k/a*	'Costa Rica'	*costarri/θ/ense*	'Costa Rican'
	católi/k/o	'Catholic'	*catoli/θ/ismo*	'Catholicism'

Within the generative tradition, these alternations have been accounted for by means of rule systems (Harris 1969; Nuñez-Cedeño 1993).

Morin (2002) argues that softening is lexicalized and has no place in the mental grammar of contemporary Spanish speakers. She provides a good deal of data to support this position. First, the phonetic context of softening is very slippery; it generally occurs preceding a front vowel or glide. However, there are exceptions since it also occurs before back vowels (e.g., *emi/s/or, perver/s/o*). On the other hand, it fails to apply in many instances even when followed by a front vowel (e.g., *Puerto Ri/k/o* > *puertorri[k]eño*, **puertorri/θ/eño* 'Puerto Rican' *vago* 'lazy' > *va/g/edad,* **va/x/edad* 'laziness').

Another possibility discussed in the literature is that softening is conditioned by certain suffixes or morpheme boundaries and not by others. The difficulty here is that there are too many exceptions. Morin notes that the suffixes *-e, -ía,* and *-ismo* appear to trigger softening: *api/k/al* 'apical' > *ápi/θ/e* 'apex'; *aboga/d/o* 'lawyer' > *aboga/θ/ía* 'practice of law'; *católi/k/o* 'catholic' > *catoli/θ/ismo* 'catholicism'. However, in other instances these same suffixes do not serve as the conditioning factor: *arran/k/ar* 'to start' > *arran/k/e*, **arran/θ/e* 'start'; *aba/d/* > *aba/d/ía,* **aba/θ/ía* 'abbey'; *taba/k/o* 'tobacco' > *taba/k/ismo,* **taba/θ/ismo* 'nicotine addiction'.

Morin also points out that /t/ > /s/ and /d/ > /s/ are the result of common assimilatory processes in many languages. One formal explanation of these changes is that the continuant feature spreads from the following vowel or glide onto the consonants /t/ and /d/. However, the other alleged softening

alternations (i.e., /t/ ∼/θ/, /k/ ∼/θ/, /g/ ∼/θ/, /d/ ∼/θ/) do not lend themselves to a straightforward assimilatory-based explanation. Morin demonstrates that from both a feature geometry framework and an Optimality Theory standpoint these alternations can only be accounted for with a great deal of ad hoc manipulation of the formal apparatus. In other words, they cannot not be viewed as natural types of consonantal assimilation.

One of the major arguments that Morin provides to support the thesis that softening is lexicalized is that these alternations never arose from a diachronic derivational process in the first place. That is, a word such as *aten/θ/ión* 'attention' did not come from the stem of *atender* 'to tend to' when the suffix *-ión* was added to it. Instead, *atención* was borrowed from Latin ATTENTIONIS at a later date. Historical phonetic evolution transformed Vulgar Latin /tj/ into modern /θ/. If softening was never the result of a historical process related to derivation, it is difficult to assume that the process responsible for it has survived into the contemporary language.

> Without ever having been productive morphophonological alternations in themselves, the apparent [t∼θ] and [k∼θ] alternations of Modern Spanish reflect historical developments … There is no historical evidence of systematic changes that resulted in productive [t∼s], [ð∼s] or [γ∼x] alternations. Words with these apparent alternations were either integral borrowings from Latin, and/or reflect the spelling pronunciations of Spanish at the time they entered the language as learned words. (Morin 2002: 157)

One could argue that the morphophonological relationships may not have their origin in historical developments but that the relationships embodied in the alternations arise synchronically as language learners glean information from the linguistic input they receive and make generalizations from those data.

If softening truly reflects a synchronic process there should be signs that it is productive. Morin tested the productivity of softening alternations by asking subjects to combine nonce words ending in /t, k, d, g/ such as *semedo* and *semoca*, with the suffixes *-ente, -ino, idad, -ico, -ense, -ismo, -ista, -ía*, and *-iano.* The combination experiment was done in the context of a paragraph such as:

En el lenguaje secreto de los niños, un semedo es un animal de cuento de hadas que se parece al dragón blanco de La historia interminable. En la imaginación de los niños existen muchos animales fantásticos que se parecen en algo al semedo. ¿Cómo se llama un animal que se parece al semedo? Es un animal ____*iano.*

'In secret child language, a *semedo* is an animal from a fairy tale that looks like the white dragon in *The Never Ending Story*. In the children's imaginations,

there are many fantastic animals that resemble the *semedo* in one way or another. What adjective would you use to describe an animal that resembles the *semedo*? It is a _____ iano animal.'

The questionnaire contained 34 questions of this type and was administered to 32 Spanish speakers.

Coronal softening was found to be completely unproductive in that it was not applied in a single answer. That is, no one transformed test items such as *semedo* into *semesiano* or *semeciano*. The phoneme /k/ was softened in only 30% of the answers (e.g., *semoca* > *semo/θ/ino*) and the phoneme /g/ in only 13%.[19] Of course, one additional factor that may have come into play with nonce words ending in velars is that they were placed next to written high vowels. Since *ce, ci* are pronounced [θe, θi] in the dialect of the subjects, and *ge, gi* are pronounced [xe, xi] the most likely explanation for many cases of softening is orthographic convention and not a productive softening process.

A similar study was carried out by Núñez-Cedeño (1993: 183–190). His goal was to determine the psychological reality of the coronal softening rules he suggests are responsible for alternation between stem-final /d, t/ and /s/ in forms such as *divi**d**ir, divisor,* and *división*. In his first experiment, he asked eight subjects to add the suffixes *-ión, -or, -ivo* and *-ble* to eight extant but uncommon Spanish verbs (*exordir, efundir, exaudir, enfurtir, cohonder, despender, luir, derruir)* and to one nonce form, *catir*. (According to Núñez-Cedeño's rules, the verbs *luir* and *derruir* contain an abstract stem-final /d/ in underlying representation from which the /s/ of the suffix *-sión* is derived. For this reason, they were included in his study.) In less than 10% of the cases did the subjects in his study 'soften' the consonant when they added *-ión* onto the verbal stem (*enfurtir* > *enfursión*). In over 60% of the cases, his proposed rule was not applied and the stem was left intact (e.g., *enfurtir* > *enfurtición*). The remaining cases involve various other odd changes.

In a second study, he presented the subjects with the verbal forms and a list of corresponding nominal forms with different morphophonemic alternations (e.g., *enfurtir* > *enfursión, enfurtión, enfurtición, enfusión).* The task of the subjects was to rate the forms on a scale of acceptability ranging from highly acceptable to highly unacceptable. In only 25% of the cases did subjects accept forms ending in *-ión* that demonstrated the change /t, d/ > /s/ as coronal softening would predict (e.g., *enfursión enfusión*). In 50% of the cases, the subjects accepted forms ending in *-ión* that maintained the stem-final /d, t/ (e.g., *enfurtición, enfurtión*).

In many regards, it is difficult to give these results a precise interpretation. The extremely small number of subjects and test items admit the possibility

that one item or subject may have severely skewed the overall results. It also means that the results do not lend themselves to reliable statistical analysis. Moreover, the results are conflated by test item as well as by subject. Since the results for individual items are not presented many questions are left unanswered. For example, half of the forms that maintained the /d, t/ in the stem were rejected by the subjects. I suspect that the rejected half consisted of nominals such as *enfurtión,* whose phonological shape is somewhat unusual in Spanish. The half that was judged acceptable were most likely of the *enfurtición* type. However, given the paucity of data presented I cannot support or refute my intuition about this matter. Methodological problems aside, Núñez-Cedeño is left to grapple with the fact that in more cases than not, the subjects did not apply the rules he posits. For Morin, these results come as no surprise since she argues that softening is not a productive process.

3.4 Depalatalization of /ñ/ and /ʎ/

The apparent alternations /n/∼/ñ/ and /l/∼/ʎ/[20] is exemplified in the following words:

donce/ʎ/a	'damsel'	*donce/l/*	'young nobleman'
aque/ʎ/a	'that'	*aque/l/*	'that'
e/ʎ/a	'she'	*e/l/*	'he'
be/ʎ/o	'beautiful'	*be/l/dad*	'beauty'
caba/ʎ/o	'horse'	*caba/l/gar*	'to ride a horse'
do/ñ/a	'Mrs.'	*do/n/*	'Mr.'
desde/ñ/ar	'to disdain'	*desdé/n/*	'disdain'
te/ñ/ir	'to dye'	*ti/n/te*	'dye'
re/ñ/ir	'to quarrel'	*re/n/cilla*	'quarrel'

In a formal analysis, Harris (1983) suggests that /ñ/ and /ʎ/ exist in the underlying forms but are depalatalized into /n/ and /l/ when they fall into a syllable coda during the intermediate stages of derivation (e.g., be ʎ + dad > be ʎ.dad > bel.dad). However, in some cases /ñ/ and /ʎ/ appear in syllable onsets in all stages of the derivation: *don.ce./l/es, *don.ce./ʎ/es* 'young noblemen'; *des.de./ñ/es, *des.de./n/es* 'you disdain.' This fact was accounted for by cyclic rule application the exact details of which are not relevant to the present discussion.

On the issue of depalatalization, Pensado (1997) asks:

> What is the need of the cycle when inflection seems perfectly faithful to surface forms? Do speakers actually set up abstract bases such as *desdeñ-* in

> order to account for derivational regularities (*desdén / desdeñar*) at the cost of complicating inflection (*desdenes*)? Are alternating forms systematically processed? (597–8)

Pensado studied depalatalization from a historical perspective and provides a great deal of evidence that there never was a process of depalatalization. Therefore, it could not have been passed down into Contemporary Spanish. As in the case of coronal and velar softening discussed by Morin, many apparent instances of depalatalization are actually due to borrowing. For example, *desdé/n/* and *donce/l/* were borrowed from Provençal. The postulated process, *desdé/ñ/- > desdé/n/* and *donce/ʎ/ > donce/l/* was never part of the historical evolution of Spanish.

Pensado also undertook a nonce word study to determine whether depalatalization is a productive process that is relevant for modern Spanish speakers. In one experiment, subjects were shown drawings of animal-like creatures and a machine that made them. The verbs used contained /ñ/ and /ʎ/ in syllable onsets while the related nouns contained /n/ and /l/ in syllable codas. In this way, the subjects were given evidence for depalatalization. For example,

> *Esto es un enapil. Esto es una máquina de enapillar*
> *Esto es un sirapén. Esto es una máquina de sirapeñar.*
> 'This is a ___. This is a machine to ___.'

They were then shown drawings containing several of the creatures and asked to fill in the blanks in three sentences:

Esto son dos _____*s.* (Task: form a plural)	'These are two ___s.'
Esto está bastante _____*ado.* (Task: form a past participle)	This is somewhat ___ed.'
Esto es muy _____*oso.* (Task: form an adjective)	'This is very ___ ous.'

The plurals would presumably be based on the nominal forms; therefore, *enapiles* and *sirapenes* would be expected. The expectation was borne out in that 81% and 87% of the plurals were *enapiles* and *sirapenes* respectively.

The formation of past participles and adjectives was more ambiguous. The subjects were divided into two groups prior to the administration of the questionnaire. One group received the stimulus sentence containing the noun *enapil* before the sentence containing the verb *enapillar*. The other group viewed the verbal stimulus before the nominal one. The groups also saw the sentences containing *sirapén* and *sirapeñar* in different orders. The order in

Table 1. Results from Pensado's study

	% enapiloso	*% enapilloso*	% other
enapil seen last	**68**	24	8
enapillar seen last	30	**52**	19
	% *sirapenoso*	% *sirapeñoso*	% other
sirapén seen last	**35**	27	38
sirapeñar seen last	27	**46**	27

which the fill-in-the-blank questions were presented was not varied for each group. It is interesting that the order or presentation influenced the outcome.

For example, if the last nonce word seen by the subjects was *enapil*, the tendency was to base the *-oso* form on *enapil* and not on *enapillar*. The responses in the 'other' category include odd answers such as *empiloroso, enapilleoso, siparoso,* and *sipariñoso* that do not demonstrate a simple combination of the stem and *-oso.*

Since past participles are based on verbal stems, *enapillado* and *sirapeñado* would be expected. However, many subjects gave *enapilado* and *sirapenado* and thus appear to have completely disregarded the allomorphy presented to them in the questionnaire. It is unfortunate that Pensado's data in this section are not clear enough that percentages for each outcome may be calculated.

Nevertheless, if depalatalization existed in the minds of the Spanish speakers who participated in Pensado's experiment they would have had enough material presented to them that they should have been able to apply it to these words just as it is supposedly applied to *caballo*∼*cabalgar* and *desdeñar*∼*desdén*. The large degree of inconsistency in their answers, coupled with the fact that many answers were based on the phonological shape of the last nonce word presented to them suggest that depalatalization is not a synchronically active process.

3.5 Intonation differences between English and Spanish

There is no doubt that the intonation patterns employed by Spanish speakers differ from those of English speakers. For many English speakers, Spanish sounds like a monotone machine gun. Spanish speakers, on the other hand may feel that English speakers produce a roller coaster of peaks and valleys in their speech. Such differences have led some to suggest that English uses four levels of pitch and Spanish only three (Stockwell & Bowen 1965; Whitley 1986[21]). In

the literature, these levels are spoken of in impressionistic terms rather than in terms of specific phonetic attributes. The idea is that English speakers employ a wider range of pitch than Spanish speakers.

Kelm (1995) set out to determine whether these alleged differences are quantifiable. He recorded ten native English and ten native Spanish speakers participating in a role play in their native language.[22] For each subject, he then measured the pitch of their conversation at 30 millisecond intervals. Breaks in the pitch contours were not considered. From these measurements he calculated the mean pitch for each speaker, the pitch range, and the standard deviation.

The Spanish speakers spoke at a mean pitch of 232 Hz while the English speakers spoke at 254 Hz. This difference is not significant. Greater relevance lies in whether English speakers reached higher highs and lower lows than Spanish speakers. That is, did they employ a larger *range* of pitch? The range of the English speakers was 150 Hz while that of the Spanish speakers was 110 Hz, however, this difference is not significant. Given these data the assumption that English speakers use four levels of pitch and Spanish speakers only three appears to be unfounded.

However, the most telling result from this experiment regards the standard deviation, which was 59 for the English speakers and only 39 for the Spanish speakers. This significant difference indicates that while speakers of both languages utilize the same absolute range of pitch, English speakers tended to vary the peaks and troughs more often than Spanish speakers. In other words, in a given period of time, English speakers vacillated their pitch more often than Spanish speakers. Perhaps this difference in variation gives the auditory impression that English speakers have a wider pitch range.

Kelm clearly cautions his readers that his acoustic findings will not necessarily correlate with results from perception. In addition, his findings may be limited to the speech of his 20 subjects. It is clear that more data need to be produced on this topic before any generalizations can be made. Of course, it is important to qualify Kelm's findings. One should note that while his study is not the definitive answer to the question of intonation differences, the fact that it is based on actual data and solid experimental method makes it many times more valuable than the previous studies founded solely on impressionistic observations.

3.6 Change-of-state verbs

Spanish has a number of ways of indicating changes of state. There are many reflexive verbs that express 'to become X' such as *enrojecerse* 'to become red,' and *enfermarse* 'to become sick.' In addition to these state specific verbs, a number of other verbs are used to indicate a change of state: *quedarse rojo* 'to become red,' *ponerse enfermo* 'to become sick.' There are in fact at least seven change-of-state verbs in Spanish that are commonly used: *llegar a ser, ponerse, volverse, quedarse, convertirse, transformarse,* and *hacerse.*

Researchers have attempted to distinguish the contexts in which these verbs appear (Coste & Redondo 1965; Crespo 1949; Eberenz 1985; Fente 1970). I will discuss four of these contexts. The first is that certain change of state verbs are thought to be used when the following predicate is nominal (*Se hizo capitán*, 'He became a captain.'), while others take adjectival predicates (*Te has puesto triste,* 'You've become sad.'). The speed at which the change occurs is the second factor governing which verb is used: Slow – *Llegó a ser una figura de importancia* 'He became an important figure;' Fast – *Al oír la sentencia del juez, se quedó perplejo* 'When he heard the judge's sentence, he became confused.'

The last two sentences exemplify a third criterion for distinguishing between change of state verbs. This involves the degree to which the change came about as the result of willful and active intent and effort. A person who becomes an important figure most likely reaches that status as the result of conscious effort and desire. Becoming confused, on the other hand, occurs passively and unexpectedly.

The fourth criterion is whether the sentence containing a change of state verb could be rewritten with the copulative *ser* or *estar.* For example, *Mi disgusto con él llegó a ser grande* 'My distaste for him became great' could be expressed, *Mi disgusto con él era grande* 'My distaste for him was great,' but not **Mi disgusto con él estaba grande*. In contrast, the sentence *te has puesto triste* would be rendered with *estar*, (*estabas triste)* and not with *ser* (**Eras triste).*

Many of the factors first discussed by Coste and Redondo (1965), Crespo (1949), Eberenz (1985), and Fente (1970) have now been incorporated into general grammars of Spanish (e.g., Butt & Benjamin 1994) as well as into pedagogical grammars (e.g., Rusch, Domínguez, & Caycedo Garner 1996). This has been done without verifying that the proposed factors actually differentiate the verbs. The original articles dealing with change-of-state verbs are suspect in one regard. Each article is written in much the same format. That is, a criterion for distinguishing the contexts in which one or more change-of-state verbs is defined, followed by a handful of sentences that exemplify that the criterion

works. Fente's treatment is somewhat more credible since his work is based on a database of 180 instances.

However, given the paucity of the data on which these studies were based, I decided to conduct a more thorough corpus-based study. I first extracted 1,283 change-of-state verbs *llegar a ser, ponerse, volverse, quedarse, convertirse, transformarse,* and *hacerse* from written and spoken sources and applied the four criteria discussed above to each verb (see Eddington 1999 for details). The literature on change-of-state verbs mentions more than four factors that may help distinguish between different change-of-state verbs. However, some of them are simply too difficult to apply to the items found in the corpora. Judging according to those criteria would require deciding whether a change is permanent or temporary, or whether it involves a change to an essential versus a non-essential property, or whether the change could be considered normal or unexpected.

Since these concepts do not lend themselves to objective quantification, only the four factors discussed above were considered: (1) was the verb followed by a nominal or verbal predicate? (2) did the change occur fast (i.e., within 24 hours) or did it conceivably take more than 24 hours (slow)? (3) did the change involve the active participation and volition of the animate experiencer? (4) could the change-of-state verb be replaced by *ser* or *estar*? Even using these four factors it was impossible to exactly determine a category for each of the 1283 instances, in which case, the category for the particular verb was left blank and not used in the calculations. Consider the sentence *se hizo tarde* 'It became/grew late.' The change-of-state verb in this sentence was categorized in this way: (1) *tarde* is an adjective; (2) the change takes less than a day to complete; (3) the change is the passive result of circumstance, not of active effort; (4) *se hizo tarde* could be replaced by *era tarde* 'It was late' (the verb *ser*), but not with the verb *estar*.

The results of the clearly classifiable cases are summarized in Table 2, in which boldface capital letters indicate a 66% or higher majority of cases with a particular feature, while lowercase, non-boldface indicates no clear majority.

Table 2. Results from the corpus study

	llegar a ser	*ponerse*	*volverse*	*quedarse*	*convertirse*	*transformarse*	*hacerse*
Noun/Adj.	n 51	**A**$_{100}$	a 64	**A**$_{99}$	**N**$_{99}$	**N**$_{99}$	a 58
Slow/Fast	**S**$_{89}$	**F**$_{88}$	**F**$_{69}$	**F**$_{94}$	f62	**F**$_{69}$	f 62
Active/Pass.	a 59	**P**$_{72}$	**P**$_{83}$	**P**$_{68}$	p 55	a 59	a, p 50
Ser/Estar	**S**$_{92}$	**E**$_{88}$	**S**$_{67}$	**E**$_{97}$	**S**$_{100}$	**S**$_{99}$	**S**$_{86}$

The previous studies are supported to a certain extent by the usage found in the corpus study. For example, Coste and Redondo (1965), Crespo (1949), and Fente (1970) all assert that *llegar a ser* is used when the resulting state is arrived at over an extended period of time. In the corpus, 89% of the cases that could be clearly judged did appear to change slowly. Fente and Coste and Redondo also correctly perceived the relationship between *ponerse* and the copula *estar* (88% correspondence), as well as the fact that changes signaled by *volverse* are usually (83%) of the sort that occur passively and not as the result of a conscious effort on the part of the grammatical subject.

However, not all of the criteria they discuss are as helpful as they suggest. For instance, Crespo, Coste and Redondo, and Fente suggest that *hacerse* is used with changes that are achieved as a result of voluntary and active effort. My judgments of the uses of *hacerse*, on the other hand, see active participation in only 50% of the cases. *Convertirse* and *quedarse* are cited by Fente as co-occurring with passive, involuntary change, yet in my corpus they appear in this context in only 55% and 68% of the cases, respectively.

These observations are not meant to be an exhaustive comparison of my results with those of the other researchers. They are merely cited to demonstrate that generalizations based on small numbers of examples are often incorrect or give the mistaken impression that there is little variation in actual usage. Of course, one difficulty with the corpus I studied is that it contains data from speakers from a wide variety of Spanish dialects. It is possible that the variation found may be due to dialectal differences in usage. To test this hypothesis I asked 32 Spaniards, 26 of whom were from Andalusia, to complete a questionnaire (see Eddington 2002a).

The questionnaire consisted of 30 questions in which the subjects had to determine which change-of-state verb or verbs sounded best to them. For example,

> *A mediados de la reunión, sonó la alarma de incendios. Por eso la reunión* →
>
> ___ *se hizo breve.*
> ___ *se quedó breve.*
> ___ *llegó a ser breve.*
> ___ *se volvió breve.*
> ___ *se transformó en breve.*
> ___ *se puso breve.*
> ___ *se convirtió en breve.*

'The fire alarm went off in the middle of the meeting. That is why the meeting was cut short.'

The 30 questionnaire items represented ten combinations of the four factors that are considered important in selecting the appropriate change-of-state verb. In the above example, *breve* is an adjective, the change occurred rapidly and was not the result of active planning and effort. If the sentence were rewritten with a copula it would be *la reunión fue breve* with the verb *ser*, not *estar.*

Subjects were free to choose more than one answer if they deemed both possibilities appropriate. On the average, subjects marked 1.5 answers per question. As Table 3 demonstrates, even for speakers of the same dialect of Spanish there is a great deal of variability in which verb is acceptable in each context.

When these data are considered along with those from the corpus study it paints a picture in which the uses of each verb are not neatly compartmentalized. As Fente notes, 'the semantic boundaries between these verbs are becoming blurred or have been blurred for a long time, and their uses … are jumbled and often confused' (translation mine-DE; 1970: 161). This is not to say that certain tendencies are not apparent, only that there is a great deal of overlap in the uses of each of the seven change-of-state verbs included in the study. The picture appears much more clear-cut for researchers who based their assumptions on only a handful of examples. Spanish change-of-state verbs represent another case in which the broad generalizations made are only partially valid because the data on which they were founded was entirely too small.

Table 3. Results from the questionnaire in number of responses

	llegar a ser	*ponerse*	*volverse*	*quedarse*	*convertirse*	*transformarse*	*hacerse*
Adj F Act E	16	**50**	**49**	0	2	3	4
Adj F Act Se	5	**26**	5	**75**	0	2	1
Adj F Pas Se	21	2	**38**	11	15	5	**42**
Adj S Act E	8	**67**	**35**	8	7	6	11
Adj S Pas E	10	15	**54**	**58**	12	6	15
Adj S Pas Se	24	10	**65**	1	17	14	**29**
Nom F Act Se	13	1	28	0	**60**	19	**45**
Nom F Pas Se	15	0	12	1	**72**	**18**	16
Nom S Act Se	**70**	0	11	0	31	9	**36**
Nom S Pas Se	16	1	35	2	**59**	**40**	27

Adj = adjectival predicate; **Nom** = nominal predicate; **F** = fast change; **S** = slow change; **Act** = active change; **Pas** = passive change; **E** = predicate expressed with *estar*; **Se** = predicate expressed with *ser*.

3.7 Conclusion

My goal in this chapter has been to show that many linguistic processes that are thought to occur in Spanish have been based on impressionistic observations. However, once empirical observations are brought to bear on these phenomena a very different picture often emerges. Linguists, as all other scientists, sometimes propose theories based on hunches or 'gut feelings.' It is lamentable that these impressions are too often presented as facts about the language without substantial supporting evidence. The examples discussed in this chapter should illustrate the vacuousness of assertions about phenomenon X in language Y without supporting empirical evidence.

Chapter 4

Frequency N Counts V

4. Introduction

Formal approaches to linguistics are generally conceived of in such a way that they do not incorporate a great number of factors that are relevant in speech production and comprehension. Consider the deletion of /d/ in intervocalic position. This results in *soldado, hablado* 'soldier, spoken' being pronounced as [soldao] and [aβlao] in many dialects of Spanish. A simple way of representing this formally is d > Ø / V__V. However, this rule does not explain that many factors determine the extent to which *d*-deletion occurs, both within a single dialect as well as within the speech of a single individual. The sociolinguistic literature is replete with social factors that a rule of the sort proposed above does not even attempt to incorporate. I will leave the role of social factors to the extant treatises of Spanish sociolinguistics. However, another difficulty with representing *d*-deletion formally is that the role of frequency is not addressed. The phone [d] tends to be deleted more often in frequently occurring words. The goal of this chapter is to demonstrate that frequency is an important factor in linguistic processes such as *d*-deletion, which will be discussed in more detail below.

It is important to define two measures of frequency: token frequency and type frequency. Consider the following word list:

inspección	*espejo*
inspección	*respeto*
inspección	*secuestro*
opción	*consideración*
oración	*consideración*

The token frequency is the number of times a unit such as a word or consonant cluster appears. Therefore, the token frequency of *inspección* and *secuestro* is three and one respectively. The token frequency of the cluster *-sp-* is five, and the token frequency of words ending in *-ción* is seven. Type frequency refers to the number of different units that contain a particular pattern. There are

only four types containing the suffix *-ción* namely *inspección, opción, oración*, and *consideración*. In like manner there are only three types containing *-sp-*: *inspección, espejo, respeto.*

4.1 Frequency as a factor in language processing

Perhaps one way to begin this chapter is to ask why the word *nuclear* is often rendered as [nuwkjələ˞] instead of [nuwkliə˞], especially since the latter is a pronunciation that reflects the spelling of the word more closely. It cannot merely be a random slip of the tongue made by an individual or it would have escaped virtually unnoticed. It is actually a pronunciation that is common enough that it has received media attention, especially when it has fallen from the lips of U.S. presidents George W. Bush and Jimmy Carter. Sheidlower (2002) explains that the reason is simple; with the exception of *cochlear* and a handful of other obscure words, [-liə˞] is not a frequent string of phonemes occurring at the end of English words. Words ending in [-jələ˞], on the other hand, are abundant: *angular, spectacular, particular, muscular, cellular, regular, circular, molecular*, etc. In a similar manner, Hay, Pierrehumbert, and Beckman (2003) found in their perception study that when low-frequency consonant clusters were misperceived, they were misperceived as more highly frequent consonant clusters.

Why would more frequent patterns replace less frequent ones? One way to explain it is that high-frequency phoneme clusters are more practiced. Let me illustrate this with a nonlinguistic example. When I drive out of my neighborhood I have the choice of turning left or right on the main road. I would estimate that I turn right about 90% of the time since that is the direction to the freeway. On a number of occasions when my mind has been preoccupied, I have turned right when I actually needed to turn left. However, I have never done the opposite and made a left-hand turn when meaning to go right. In other words, my errors tend to be in the direction of my most practiced pattern. Phonological patterns such as consonant clusters may also be a type of practiced pattern.

4.1.1 *Frequency as a factor in the processing of phonetic patterns*

Let me turn to Spanish data now. Brown (1999) discusses a process in which syllable-final /p/ becomes [k]: *séptimo* 'seventh' > [sektimo], *opción* 'option' > [oksjon], *Pepsi* > [peksi], *helicóptero* 'helicopter' > [elikoktero].[23] Theories based on articulatory effort, autosegments, and maximal acoustic differentiation fail to adequately account for this process. For this reason, Brown considered frequency. What she found is that syllable-final /k/ has a token frequency

14 times greater than syllable-final /p/ in the spoken language. As far as type frequency is concerned, syllable-final /k/ is seven times more frequent than /p/ in the same position.

Slips of the tongue or ear demonstrate this same tendency. One common variant pronunciation of *polígono* 'polygon, zone' is **polígano.* According to one frequency dictionary (Alameda & Cuetos 1995), there are 52 words that end in *-ono* in contrast to 274 that end in *-ano.* In other words, *-ano* is 5.3 times more common than *-ono*. As far as token frequency is concerned (Marcos-Marín 1992), *-ano* is 2.8 times more frequent than *-ono* (1,019 versus 358 occurrences respectively). Once again, we see that less frequent patterns tend to give way to more frequently occurring ones.

The frequency of phonetic patterns can introduce itself as an uncontrolled variable in phonetic studies. Consider the experiment by Pérez (1998). The question he set out to answer was whether the phonemes /p t k/ and /b d g/ in Spanish are distinguished in terms of a voicing contrast or duration. He recorded a native speaker reading the following nonce words containing plosives: *umpasa, umbasa, untasa, undasa, uncasa, ungasa*.[24] The idea behind the study was to acoustically modify the plosives so that /p t k/, which are generally of longer duration, are given the same duration as /b d g/ respectively. The sonority of /b d g/ was also 'pasted' into /p t k/, and the voicelessness of /p t k/ was 'pasted' into /b d g/. This was done by measuring the formants, fundamental frequency, intensity peak, and the band width of the glottal pulses that correspond to the voice bar of /b d g/. These were inserted into their voiceless counterparts, /p t k/. The same data were taken from the silence preceding the release of /p t k/ and inserted into /b d g/. This resulted in 12 modified nonce words:

1. /b/ with the lack of voicing derived from /p/.
2. /p/ with the voicing of /b/.
3. /b/ with the duration of /p/.
4. /p/ with the duration of /b/.
5. /d/ with the lack of voicing derived from /t/
6. /t/ with the voicing of /d/.
7. /d/ with the duration of /t/.
8. /t/ with the duration of /d/.
9. /g/ with the lack of voicing derived from /k/.
10. /k/ with the voicing of /g/.
11. /g/ with the duration of /k/.
12. /k/ with the duration of /g/.

Subjects listened to these modified nonce words along with the unmodified nonce words and transcribed what they heard orthographically.

Previous studies indicate that both voicing and length are important cues for distinguishing between voiced and voiceless stops (Benoît & Gurlekian 1992). However, it appears that duration is even more important than voicing itself (Martínez Celdrán 1991). The subjects in Pérez's experiment perceived /b d g/ in *umbasa, undasa*, and *ungasa,* which had been given the durations of /p t k/, as *umbasa, undasa*, and *ungasa,* in only 55% of the cases. They heard *umpasa, untasa*, and *unkasa* in 41% of the cases and perceived something else in 4% of the cases. The voiceless plosives /p t k/ whose duration had been modified to match that of their voiced counterparts were heard as /p t k/ at a rate of 84% and as /b d g/ in 15% of the cases.

Switching the voicing characteristics of the plosives did little to change the way most of them were perceived. The /b d g/ phonemes that had been 'devoiced' were still heard as /b d g/ at a rate of about of 95%, and as their voiceless counterparts at a rate of only 3%. In spite of being synthetically voiced, /p/ and /t/ were heard as /p/ and /t/ in the majority of cases (99% and 82% respectively). If this trend were consistent, one would expect the artificially voiced /k/ to also be perceived as /k/, but this was true only of 24% of the cases. The next most plausible outcome would be that the inserted voicing caused it to be perceived as /g/, but none of the subjects perceived it that way. What Pérez cannot explain in his study is why the 'voiced' /k/ in *unkasa* was heard as *untasa* 73% of the time.

I submit that the phonetic characteristics of the plosives are not the only factor that influenced the subjects' perceptions; the frequency of the consonant clusters that appear in the nonce words, (*-mp-, -mb-, -nt-, -nd-, -nk-, - ng-*), have also entered into the experiment as an uncontrolled variable. The token frequency of these clusters was calculated by counting the number of times they appeared in the Marcos-Marín (1992) corpus of spoken Spanish.

Cluster	Frequency
-mp-	9,733
-mb-	7,737
-nt-	46,235
-nd-	18,489
-nk-	5,782
-ng-	4,596

The cluster *-nt-* is eight times more common than *-nk-* and ten times as frequent as *-ng-*. The frequency of *-nt-* may explain why *unkasa* with a voiced /k/ may have been perceived as *untasa* so often.

Of course, one could counter that *umpasa* with an artificially voiced /p/ was not heard as *untasa* even though *-mp-* and *-mb-* are much less frequent than *-nt-* in the same way that *-nk-* is much less frequent than *-nt-*. However, the fact that voiced /k/ was heard as /t/ while voiced /p/ was not may be explained by the acoustic similarity of velars and dentals as opposed to labials and dentals. Consider the transitions that occur in the second and third formants of the vowels of the following syllables (Quilis 1993: 208–211):

[ba, pa]	F3	/----
	F2	/----
[ta, da]	F3	___
	F2	___
[ka, ga]	F3	/----
	F2	___

Labial consonants produce a transition that moves from lower to higher frequency in both F2 and F3. Dental consonants produce the exact opposite movement of both these formants, which explains why dentals and labials are highly dissimilar. Note, however, that the movement of F2 is in the same direction for both dentals and velars, which accounts for their tighter acoustic similarity. *Unkasa* with a voiced /k/ was misperceived as *untasa* for two reasons: (1) velars and dentals share the same F2 transition, and (2) *-nt-* is a much more frequent consonant cluster than *-nk-* or *-ng-*. On the other hand, *umpasa* with voiced /p/ was not misperceived as *untasa*. While it is true that *-mp-* and *-mb-* are much less frequent than *-nt-*, they are too phonetically distinct to be confused with *-nt-*.

4.2 Explaining epenthesis in terms of frequency

Vowel epenthesis before words beginning with /s/ plus a consonant (*sC-*) was a process that began in Late Vulgar Latin and continued into Spanish:

SPECULUM	>	*espejo*	'mirror'
STRICTU	>	*estrecho*	'narrow'
STUPPA	>	*estopa*	'burlap'

In Spanish, this process applied across the board and resulted in a lexicon devoid of words beginning in *sC-*. In one sense, epenthesis appears to be a process

that is synchronic as well as historical, since foreign words that are borrowed into Spanish undergo it:

smoking jacket	>	*esmóquin*
standard	>	*estándar*
stress	>	*estrés*

For this reason, one could assume that this epenthetic process has been in force for about two millenia. Given the high profile status of epenthesis as both a historical and contemporary process it has received a good deal of attention in a number of formal frameworks (e.g., Cressey 1978; Eddington 1992, 2001a; Harris 1983, 1987; Hooper 1976; Morgan 1984).

I would like to present some data that may not initially appear to be related to epenthesis. It involves a process in which *VsC-* clusters are converted into *esC-* clusters. Historically, this occurred to produce a number of words in Standard Spanish:

esconder	<	ABSCONDERE
escuchar	<	A(U)SCULTARE
espárragos	<	ASPARAGUS

The change from /a/ to /e/ is attributed to analogy with the large number of words beginning with *esC-* (Lloyd 1987:63). Although the standard language contains only a few words of this sort, (the three listed are not an exhaustive list), the process was much more common in Old Spanish than is evident in Contemporary Spanish. The following words and their derivatives are documented in Old Spanish:[25]

espital	<	HOSPITALE	'hospital'
esciéndolos	<	ADSCENDERE	'ascending them'
escurece/escuro	<	OBSCURUS	'it grows dark, dark'
espaviento	<	ASPA+VENTUS	'arm waving'
especto	<	ASPECTUS	'aspect'
espereza	<	ASPERITIA	'roughness'
Esrael	<	ISRAEL	'Israel'
estrología/estrólogo	<	ASTRUM	'astrology, astrologer'
estuto	<	ASTUTUS	'astute'

Errors committed by present-day Spanish speakers also evidence *VsC- > esC-*. The following words were found by searching Spanish language pages on the Worldwide Web:[26]

esbesto	<	*asbesto*	'asbestos'
escender	<	*ascender*	'ascend'
escensor	<	*ascensor*	'elevator'
escilar	<	*oscilar*	'oscillate'
escuridad	<	*oscuridad*	'darkness'
eslámico	<	*islámico*	'Islamic'
esmosis	<	*osmosis*	'osmosis'
espaviento	<	*aspaviento*	'arm waving'
especto	<	*aspecto*	'aspect'
espereza	<	*aspereza*	'roughness'
esqueroso	<	*asqueroso*	'disgusting'
Esrael	<	*Israel*	'Israel'
estentar	<	*ostentar*	'to flaunt'
estilla	<	*astilla*	'chip'
estmo	<	*istmo*	'isthmus'
estrología	<	*astrología*	'astrology'
estucia	<	*astucia*	'wisdom'
Esturiano	<	Asturiano	*'Asturian'*
estuto	<	*astuto*	'wise'

In order to account for these words along with those that undergo epenthesis, one would have to maintain that there are two separate processes at work: *sC- > esC-* and *VsC- > esC-*. I submit that both processes are actually the result of the same process. It is this process that turns *Pepsi* into [peksi] and *nuclear* into [nuwkjəlɔ˞], and that influences subjects to hear *unkasa* with a synthetically voiced /k/ as *untasa.*

> Patterns that range over large numbers of lexical items are highly reinforced or strengthened and apply more readily to new items [and to less frequent existing items-DE], while patterns that are found in a smaller number of items are correspondingly weaker and less apt to be productive. (Bybee 1988: 125)

Consider the data in Table 4, taken from Eddington (2001a). Historical epenthesis has produced a situation in which *esC-* is about five or six times more frequent than *asC-, isC-, osC-*, and *usC-* combined. It is no wonder that misperceptions and slips of the tongue (and pen) tend toward *esC-*. When Spanish speakers are confronted with words such as *scanner* and *stress,* whose initial consonant cluster frequency in Spanish is zero, they bring such words into alignment with Spanish phonotactic patterns by adding an /e/. In terms of frequency, the other vowels are simply not big contenders for epenthesis in this position. In fact, /e/ is the most frequently occurring vowel in Spanish regard-

Table 4. Frequency of *VsC-* clusters

	esC-		*a, i, o, u sC-*	
Type Frequency	2,367	(82.3%)	447	(17.7%)
Token Frequency	21,549	(85.3%)	3,707	(14.7%)

Table 5. Frequency of *esC-* versus *seC-* clusters

	esC-		*seC-*	
Type Frequency	2,367	(78.8%)	637	(21.2%)
Token Frequency	21,549	(75.8%)	3,885	(24.2%)

less of position (Guirao & García Jurado 1990) so it is even less surprising that it is the vowel of choice.

One may ask why it is that *scanner* is not modified by inserting the most frequent vowel after the /s/ instead of before, yielding **secáner*. After all, this would break up the consonant cluster to form an open syllable, which according to some would be the preferred state of affairs since open syllables are unmarked. Table 5 demonstrates that *esC-* is an even more common pattern than *seC-*.

Given the fact that *sC-* clusters are nonexistent in Spanish, foreign words with this structure have a strong tendency to epenthesize in order to bring them into line with Spanish phonotactics. The situation is somewhat different in the case of words such as *osmosis* and *astuto*. There is nothing illicit about their phonological structure and they are generally left untouched; however, their structure is infrequent, which explains the slight tendency for them to be replaced with *esC-*.

4.3 Vosotros and vos imperatives

The Latin VOS imperative suffix -TE evolved into *-d* as the *vosotros* imperative: *venid* 'come,' *predicad* 'preach.' The historically related *vos* imperative is marked by the lack of a morpheme that indicates that *-d* was lost: *vení*, *predicá*. Another reduction that occurred involved the clitic pronoun *vos,* which was reduced to *os*. What is interesting about the change from *vos* to *os* is that it was first manifest when it followed the *-d* of the *vos(otros)* infinitive: *venidvos* > *venidos* (Lathrop 1984: 155). The reason given for this change is that the cluster *-dv-* was phonotactically unusual.

A number of other changes have occurred in this context that appear to have been motivated by the tendency to metathesize uncommon consonant

clusters into more frequent ones (see Hume 2001, 2004 for examples of this tendency from other languages). For example, in Golden Age Spanish the /d/ of the imperative and the /l/ of the clitic *le(s)* underwent metathesis (Menéndez-Pidal 301): *dadle* 'give him/her' > *dalde; ponedlo* 'put it' > *poneldo*. A number of other words developed into Spanish due to the low frequency of *-dC-* clusters and pressure from more frequent clusters:

LEGITIMU	>	*lindo, *lidmo*
CATENATU	>	*candado, *cadnado*
SEROTINU	>	*serondo/seroño, *serodno*
RETINA	>	*rienda, *riedna*
CAPIT(U)LU	>	*cabildo, *cabidlo*
SPAT(U)LA	>	*espalda, *espadla*
TIT(U)LARE	>	*tildar, *tidlar*

In Contemporary Peninsular Spanish, an *-r* final *vosotros* imperative form has arisen: *apagadla* 'turn it off' > *apagarla*; *decidme* 'tell me' > *decirme*. The Royal Spanish Academy considers this form uncultured speech (Real Academia 1985:460), but it is fairly widespread and has been documented in Andalucía, La Rioja, Aragón, and Navarra (Alvar et al. 1980: map 1837; Alvar et al. 1973: map 1.719).

The change from *-d* to *-r* in *vosotros* imperatives and to *-Ø* in *vos* imperatives may also be attributed to consonant cluster frequency (Eddington 1991). Consider the possible clusters that arise when the imperative morpheme *-d* is combined with the consonants of the clitic pronouns: *dejádnoslo* 'leave it with us,' *dádselo* 'give it to him,' *bajadlo* 'take it down,' *ayudadme* 'help me.' The type frequency of the clusters in Table 6 was taken from Alameda and Cuetos (1995) and the token frequency from Marcos-Marín (1992). As can been seen, replacing /d/ with /r/ in these contexts results in more usual consonant clusters.

As far as the *vos* imperatives are concerned, the lack of any /d/ at all means that there is no consonant cluster. Instead, the highly frequent *-Vn-*, *-Vs-*, *-Vl-*, and *-Vm-* are the result. In sum, the processes of deletion and metathesis that

Table 6. Frequency of consonant clusters in *vos* and *vosotros* imperatives

Cluster	Type Freq.	Token Freq.	Cluster	Type Freq.	Token Freq.
-dn-	4	4	*-rn-*	1,014	2,292
-dl-	9	13	*-rl-*	2,265	2,628
-ds-	12	50	*-rs-*	1,792	3,515
-dm-	118	289	*-rm-*	1,680	5,178

applied to form the contemporary *vos* and *vosotros* imperatives, as well as the change from *vos* to *os* are due to highly frequent clusters ousting uncommon clusters. Of course, word final *-d* and intervocalic *-d-* are frequently subject to deletion anyway which provides additional motivation for the formation of the phonetic form of the *vos* imperative.

4.4 Word frequency

To this point, I have presented cases in which the frequency of phonetic patterns plays a role in the direction of speech errors and sound change. Bybee (2001) provides an example of how the frequency of whole words is also a factor. She inspected a corpus of New Mexican Spanish and counted the number of times /d/ was retained and deleted intervocalically. She found that *d*-deletion occurred in about 23% of the cases overall. In words with a token frequency of 123 or less, *d*-deletion occurred 10% of the time. Words with a token frequency of 124 or higher, on the other hand, suffered *d*-deletion in 24% of the cases. What is more, deletion occurred in 58% of the words containing the first conjugation past participle forms *-ado, -ados, -ada, -adas*, and only in 29% of the second and third conjugation past participle forms *-ido, -idos, -ida, -idas*. This is relevant because the first conjugation has a much higher type and token frequency than the second and third conjugations combined. A rule such as d > Ø / V__V simply is not able to capture these sorts of data while appeals to the influence of frequency are. Of course, there may be some phonetic motivation for deletion as well (Bybee 2001: 150). The vowel [a] involves a lower jaw position than [i]. Therefore, in the sequence [að] the tongue is pulled farther away from the teeth than in the sequence [ið] which makes in more difficult to reach the target approximant pronunciation.

4.5 The frequency of word combinations

In a study of English, Vogel Sosa and MacFarlane (2002) considered word pairs containing the preposition *of*. Their data were divided into four groups depending on how frequent the pairs of words were – their collocational frequency. For example, *kind of* has a high-frequency while, *sense of* has a low-frequency. The task of their subjects was to listen to sentences and press a key as soon as they heard the word *of*. Interestingly, the subjects failed to recognize the word *of* in 55% of the test cases. They also found that the time required to recognize high-frequency collocations such as *kind of* was greater than the time required to respond to lower frequency collocations such as *sense of*. What the

data suggest is that high-frequency collocations are processed not as separate words, but as a holistic units. For this reason, more time is required to parse the unit into its component words and recognize *of*. The words that comprise low-frequency collocations, on the other hand, are more likely to be processed as separate units, so the *of* is more salient as an individual word.

Muller (2002) applied this same experimental methodology to Spanish collocations containing *de* 'of,' such as high-frequency *dentro de* 'inside of' and low-frequency *casa de* 'house of.' Unlike the results from the English study, the outcome of the Spanish experiment showed no effect of collocational frequency. Muller explains that the reason for the difference may rest in the fact that the frequencies of the English collocations were much higher than in the Spanish ones. That is, all of the high-frequency word pairs in Vogel Sosa and MacFarlane's experiment had a frequency of 267 occurrences per million or above.[27] Of the 11 high-frequency word pairs used by Muller, seven had a frequency of less than 267 instances per million and the remaining four occurred very close to the 267 occurrence cut-off established in the English study (i.e., 307, 305, 281, and 276 times per million). Muller hypothesizes that

> in order for collocations such as these to become lexicalized into single morphemes a much greater frequency of co-occurrence (and phonological reduction) is required. In other words, there may be a certain threshold of co-occurrence, which we must pass in order to begin processing two items as a single unit. (2002:15)

In order to test this, further research that utilizes higher frequency collocations in Spanish is in order.

Although the collocational frequency was not a factor in Muller's study, some frequency effects were found. Muller also used a measure of frequency called transitional probability, which is calculated by dividing the collocational frequency by the number of times the first word in the pair occurs in all contexts. For example, *serie de* 'series of' occurs 155 times, while *serie* is found a total of 181 times; therefore the transitional probability that *serie* will be followed by *de* is 86% (155/181). This sort of probability was found to be relevant for Spanish speakers. This means that upon hearing words with a high transitional probability, such as *serie,* the subjects were expecting *de* to follow and recognized *de* faster. In the case of *casa de,* whose transitional probability is only 12%, the chances that *de* will be the next word were slim, which is responsible for the slower responses.

4.6 Conclusion

My purpose in presenting the frequency data in this chapter is to demonstrate that if one's goal is to study how language is processed, it is imperative to consider the effect of frequency. High-frequency words and phoneme combinations are processed differently from those of lower frequency. Of course, if an analysis only claims to describe linguistic structure, frequency is not an important factor.

Chapter 5

Linguistic processing is exemplar-based

5. Introduction

In Chapter 1, I discussed a number of reasons why the formal mechanisms of many linguistic analyses may not be considered to relate to actual linguistic cognition. If such accounts are assumed to underlie language production and comprehension, they would mean that children sort through the input they receive and subconsciously find patterns and generalizations in that data and then convert those generalizations into rules, constraints, parameter settings, etc., that are stored as separate entities in the mind. These entities would then be called on to perform in subsequent processing.

Early generative approaches assumed that the lexicon contained only the minimal amount of information that could not be derived by general rules. It contained morphemes and not whole words. This meant that a great deal of processing had to be performed on the input and output data. Some early researchers (e.g., Halle 1973; Jackendoff 1975) did suggest that all known words could be stored as whole entities in the mind. However, the majority of formal work still assumes that computation has the greatest role in processing, while the contents of the lexicon are secondary. Nevertheless, it is heartening to see that more contemporary formal approaches are beginning to acknowledge the influence of fully formed lexical items: lexical conservatism (Steriade 1997, 1999), correspondence theory (Benua 1995; Burzio 1996; McCarthy 1995; McCarthy & Prince 1994a, b), and base-identity (Kenstowicz 1996).

In fact, a number of language researchers have become disenchanted with models of language that espouse rule-based mechanisms as the principal mechanism of linguistic cognition. Due to the dissatisfaction with such models, several non rule-based models have recently been proposed (e.g., Bybee 1985, 1988, 1995; Goldsmith 1993; Lakoff 1993; Stemberger 1985, 1994). Other models have been developed under the connectionist architecture (e.g., Cottrell & Plunkett 1991; Plunkett & Marchman 1991; Rumelhart & McClelland 1986; Seidenberg 1992; Seidenberg & Bruck 1990; Seidenberg & McClelland 1989). Another model, one that has received less attention in the literature,

is Skousen's Analogical Modeling of Language (1989, 1992, 1995). It is this framework that I wish to explore in more detail.

5.1 Processing by exemplars

The idea that language is exemplar-based turns the storage versus processing paradigm on its head. It asserts that people go through life storing all of the linguistic input they receive in all of its redundant glory and with all of its messy, irrelevant detail. This means that speakers do not need to subconsciously find systematic correspondences and generalizations in the data, make rules or constraints out of them, then discard the input they are based on. Instead, the generalizations exist in the linguistic experiences they have stored in their long-term memory. Behavior that appears to be rule-based may be explained by assuming that people have recourse to their past experience that is stored in highly organized matrices in terms of similarity. In short, linguistic cognition entails enormous amounts of storage and little processing.

Prototype theory is similar to exemplar models. In some models the most representative exemplar, the prototype, is stored and used in processing. In others, an abstract prototype is generalized from the exemplars and stored. However, psycholinguistic research demonstrates that exemplar models generally outperform prototype models (Chandler 2002).

There is also evidence to suggest that most words are stored as wholes, not merely as combinations of morphemes (Alegre & Gordon 1999; Baayen, Dijkstra, & Schreuder 1997; Butterworth 1983; Bybee 1985, 1988, 1995, 1998; Manelis & Tharp 1977; Sereno & Jongman 1997; Stemberger 1994). Of course, this claim has not gone uncontested (e.g., Pinker & Prince 1988; 1994), and the debate continues. Nevertheless, psycholinguistic studies demonstrate that not only are words stored as types, but individual tokens of the same word may be stored in long-term memory (Goldinger 1997; Palmeri, Goldinger, & Pisoni 1993). Words also appear to be stored with all of the phonetic detail present in the speech signal, rather than in a form that has all redundant and irrelevant phonetic details abstracted away (Brown & MacNeill 1966; Burton 1990; Bybee 1994, 2001; Fougeron & Steriade 1997; Pisoni 1997).

In one study, Kolers and Roediger (1984) found that when subjects saw or heard a word they had previously experienced, if the word was presented the second time with the voice of a different speaker, or in a different tempo, or different font, or with different spacing, those differences affected the subjects' responses. In like manner, Spanish speakers were found to process samples of Spanish belonging to a different dialect more slowly than samples belonging

to their own dialect (Reiter-Boomershine 2004). Findings such as these are only possible if actual tokens are stored and not if supposedly irrelevant details are removed.

The exemplar approach to language entails the storage of not only word tokens but combinations of words as well (Bybee 1998; Pawley & Syder 1983). In fact, the productive aspect of syntax can be explained as storage of whole sentences and recombinations of fragments of stored sentences into new sentences (Becker 1983; Bod 1998). Psychologists have also explained learning as an exemplar-based process (Barsalou 1999; Logan 1988; Nosofsky & Palmeri 1997). Ellis (2002) has demonstrated that language acquisition may be explained in terms of storage of past linguistic experience and direct reference to those stored memory tokens.

One of the most difficult challenges for formal models is that they do a very poor job of explaining frequency effects. Numerous studies have shown that high-frequency words are accessed more rapidly than low-frequency words (e.g., Allen, McNeal, & Kvak 1992; Scarborough, Cortese, & Scarborough 1977), and that they are less subject to error than low-frequency items (e.g., MacKay 1982). A number of other frequency effects were also presented in Chapter Four. Frequency effects are found in language acquisition data (Ellis 2002) and form the basis for phonotactic judgments (Frisch 1996, Frisch, Large, Zawaydeh, & Pisoni 2001).

The frequency of a word also has implications for morphological processing (Sereno & Jongman 1997) and for phonological processing (Bybee 2001). For example, Bybee (2000) demonstrates that the deletion of word-final /t, d/ in English is influenced by word frequency. Deletion applies more often to highly frequent words such as *told* (> [t^{h}ol]) than it does to lower frequency words such as *meant* (> [mɛ̃n]). These findings naturally fall out of a mental architecture in which individual word tokens are stored (Bybee 2000; Hooper 1981; Pierrehumbert 2001).

5.2 Exemplar-based models

One advantage of assuming that language processing is exemplar based is that it lends itself to explicit empirical test. That is, it is not difficult to estimate with some degree of accuracy what words exist in a speaker's mind. It would be foolish to assume that each Spanish speaker passes through life knowing all the words in the *Diccionario de la Real Academia Española.* However, the most frequently occurring words are most likely part of every Spanish speaker's mental storehouse, as are the most frequent word combinations; these can be put into a

computer database and used for experimentation and hypothesis testing. Data from corpora of natural languages may also be used as approximations of what a speaker's past linguistic experience may be.

A number of computer algorithms have been devised to test the ability of exemplar theory to model cognitive processes (Aha, Kibler, & Albert 1991; Daelemans, Zavrel, van der Sloot, & van den Bosch 2001; Medin & Schaffer 1978; Nosofsky 1988, 1990; Pierrehumbert 2001; Riesbeck & Schank 1989; Skousen 1989, 1992). There are numerous differences among the models, some of which have been discussed elsewhere (Daelemans, Gillis, & Durieux 1994; Chandler 2002; Shanks 1995). What is most important is what they have in common. Consider this simplified description: If the task is to predict whether consonant deletion will apply to a given word or between a given pair of words, the database would contain information about words (or word pairs) taken from natural language usage such as a corpus of utterances. For each entry in the database, a category variable would specify whether deletion did or did not occur. The algorithm's task is to take a word as a test case and determine its similarity to the other words in the database (the determination of similarity is where the algorithms vary most radically from each other). If the word is found to be most similar to a database item (or items) demonstrating deletion, then the word is predicted to undergo deletion.

According to these models, there is no independent rule of deletion apart from the stored memory traces of the words themselves. The probability that deletion will or will not apply to a word involves inspection of the database of past experiences. The only computational part of an exemplar model is a search and compare metric coupled with the ability to apply the behavior of the items found in the search to the word in question. For instance, if a search finds that the most similar item to *pelado* is *helado*, and the majority of cases of *helado* are pronounced with a deleted [ð], then the predicted outcome is *pelao.* Of course, the common phonetic tendency to undershoot a target pronunciation and delete sounds in certain contexts will play a role and increase the probability of [ð] deletion as well. This is a fairly simplified example, since some exemplar algorithms consider the similarity of many, possibly thousands of items in the database when making predictions and do not necessarily consider only one. Models that base their predictions on one or a few database items are referred to as nearest neighbor models.

One objection often raised at this point is that searching through the enormous amount of stored data in the mind is simply not feasible. At present, the inner workings of the brain remain a mystery to science so it is impossible to say how such a search is actually implemented. However, the speed at which

such searches are performed is easily demonstrated. If I were to ask you to name the first country you can think of that begins with the letter *F*, you could do so. Yet in the process of searching, you probably would not go through a list of countries in your mind's eye. I would also predict that most people would have imagined France or Finland and not Faroe Islands or Fiji (unless you happen to be Faroese), since the former are more commonly encountered. I could also ask you to name a Spanish word ending in *s,* or to name someone you know who has curly hair, or who stutters. This sort of information is easily and quickly searched and accessed from the storehouse of past experiences. In like manner, when a person accommodates his or her speech to fit the patterns of a particular speech community, that person uses the stored knowledge about those patterns in order to modify his/her own speech.

5.3 Analogical modeling of language

The influence of one word on another is commonly referred to as *analogy*. Traditionally, analogy has been used to account for exceptional outcomes. That is, when an outcome does not obey the general rule, a form that is semantically or phonetically similar to the exceptional one is sought that is then said to influence the exceptional form in such a way that it does not develop according to the general process. What makes this sort of analogy suspicious is that it ultimately serves to patch up the inability of rules to derive all forms. In addition, no limits are set regarding what forms can serve as analogs nor on how similar two forms must be in order for analogy to become a factor. In contrast to the traditional notion of analogy, exemplar-based models assume that both regular and irregular forms are attributable to the analogical influence of other forms. These models also recognize that irregular isolate forms, such as *fue* 's/he went' from *ir* 'to go,' cannot be predicted.

I would now like to discuss one particular exemplar model called Analogical Modeling of Language (henceforth, AM). The reader is referred to Skousen (1989, 1992, 1995) for specific details of the algorithm; I will briefly outline the tenets of the theory. In order to understand AM it is useful to compare it to the more familiar rule model. Rule models derive surface forms from underlying forms by the application of rules, while constraints in optimality theory apply to candidates produced by GEN. In AM, two things are needed in order to predict forms: a database of fully specified words[28] and a mechanism for searching and comparing those words. The behavior of the words most similar to the word in question generally predicts the behavior of the word in question, although the behavior of less similar words also has a small chance of applying.

The basic algorithm is the following: (The given context is the linguistic item whose behavior is being predicted).

> We first search for actual examples of that context and then move outward in contextual space looking for nearby examples. In working outward away from the given context we systematically eliminate variables, thus creating more general contexts called *supracontexts*. (Skousen 1995: 217)

The probability that a word is chosen as an analog for the given context depends on three derived properties (Skousen 1995: 217):

(1) *proximity:* The more similar the example is to the given context, the greater the chances of that example being selected as the analogical model.

(2) *gang effect:* If the example is surrounded by other examples having the same behavior, then the probability of selecting these similarly behaving examples is substantially increased.

(3) *heterogeneity:* An example cannot be selected as the analogical model if there are intervening examples with different behavior that are closer to the given context.

These derived properties are important since they constrain what examples can constitute analogs, as well as decide between competing analogs. These are precisely the factors that are lacking in traditional appeals to analogy.

Suppose that the task is to predict the nominal form of the verb *tentar* 'to tempt.' Would it be *tención, tentición, tensión,* or *tentación*? The algorithm would find all verbs that begin with /t/ and consider them together. It would also compare all verbs with /e/ as the nucleus of the penult syllable. Other groupings would contain the verbs beginning with /te/, /ten/, or /t/ and whose theme vowel is /a/, and so on until all possible groupings of all variables are considered. These groups, called *subcontexts*, are then inspected in order to calculate disagreements.

A disagreement occurs when not all of the verbs in the subcontexts bear the same relationship to their nominal form. For example, if the subcontext of verbs ending in *-tar* contained *representar* 'to represent' and *adoptar,* 'to adopt' there would be a disagreement because the relationship between the verbs and their corresponding nominals, called the behavior, is different. *Adoptar* takes the suffix *-ción* and loses the stem-final /t/ and stem vowel /a/, yielding *adopción*. *Representación,* on the other hand, maintains the stem vowel and /t/ of the verb. Under certain conditions, members of a subcontext containing disagreements will be eliminated from consideration. In this example, the algorithm specifies that verbs ending in *-tar* do not form a cohesive enough group from

which to draw analogs on which the nominal form of *tentar* may be predicted. Bear in mind that the more similarities a given verb has to *tentar*, the more subcontexts it will appear in which means that it has a greater chance of not appearing in subcontexts that are eliminated due to disagreement. In general, a word with more in common with the given context exerts more analogical influence on the given context than a word that has less in common. I will not go into exact details about precisely what conditions must be met in order for a subcontext with disagreements to be excluded (see Skousen 1989 for details).

Once all the irrelevant subcontexts have been eliminated, the remaining verbs constitute the analogical set. The words from this set can then serve as analogical models for a given context. For the purposes of this example, let us assume that the analogical set for *tentar* contains three items and that AM has calculated the extent of similarity (or predicted probability) of each as specified below:

A	*representar*	'to represent'	(*representación*)	50%
A	*excitar*	'to excite'	(*excitación*)	30%
B	*adoptar*	'to adopt'	(*adopción*)	20%

Notice that the first two verbs in the analogical set have the same sort of behavior/relationship (A) with regard to their nominal forms. According to Skousen (1989:82), there are two ways in which the contents of the analogical set can influence the outcome. The first is that a word[29] can be randomly selected from among those in the analogical set and the outcome for that word can be applied to the given context. In this case, there is a 33% chance of choosing any one of the three words. Since two of them have the same behavior, the probability of behavior A is 66% (33+33). The second possibility is to determine which outcome is most frequent among the words in the set and assign that outcome to the given context. Since the probability of A is 80% (50+30) behavior A would apply.

One possible mechanism for nominalizing *tentar* can be conceived of as a sort of proportional analogy: *representar* is to *representación*, and *excitar* is to *excitación* as *tentar* is to __? One could assume that the process involves adding *-ación* to the stem *tent-*, or adding *-ción* to the stem plus theme vowel *tenta-*. Another possibility would be to assume the entire infinitive form is considered and that the final *-ar* has to be deleted and replaced by *-ación*. The point here is that the exact procedure speakers employ in order to modify the nominal of *tentar* so that it bears the same relationship that *represenatar* and *excitar* bear to their nominal forms is irrelevant as long as the output is the same. In fact, the exact mechanism may vary from one speaker to another. Another important

point is that the procedure employed to produce the correct outcome is devised on the fly; it is not recalled from memory as a preestablished or prefabricated rule, although it is possible that repeated calculation of an analogical set allows the set itself to be stored in memory thus eliminating the need to perform a search each time.

One important thing to keep in mind is that AM calculates the most relevant model or relationship but does not necessarily specify how that relationship is to be used. Much of the AM literature assumes that the relationship is in the form of a proportional analogy, but non-proportional analogy and product-oriented analogy have also been considered (Skousen 2002: 42–43). Assume that the task is to predict the past tense of the nonce word *spling* in English (see Bybee & Moder 1983; Bybee 2001: 126–129). If the most relevant analogs are calculated to be *spin* and *sting* then the speaker could apply this proportional analogy: *spin*∼*spun, sting*∼*stung, spling*∼*?*

The analogy is based on the fact that *spling* shares /I/ and initial /s/ with both forms, shares /st/- with *sting* and /ŋ/ with *sting*. The relationship between the present and past forms is such that where there is /I/ in the present the vowel /ʌ/ appears in the past, therefore, if the speaker applies a proportional analogy the past tense would be *splung*. What happened, then, to the verb *strike* whose past tense form was originally *striked*, but has developed the newer past tense *struck* by analogy? Assume that the most similar verbs are *slink, sting*, and *spin*. The analogies would be: *slink*∼*slunk, sting*∼*stung, spin*∼*spun, strike*∼*?* Proportional analogy cannot apply to this set because the /I/ ∼/ʌ/ relationship does not hold between [strajk] ∼[strʌk] as it does between the other members of the analogical set. Therefore, *struck* did not develop by application of the /I/ ∼/ʌ/ relationship, but only by an inspection of the past tense side of the relationship. That is, the analogy was product-oriented. The analogical set contains present tense forms that are similar to *strike,* and the past tense of these verbs all have the vowel /ʌ/, therefore the past tense *struck* arose based on the past tense form of the verbs. Another case of non proportional analogy is the past tense *branged* sometimes used by children as the non-standard past tense of *bring*. There is no single proportional analogy that can explain this form. However, this form is arguably influenced by two different patterns in the lexicon: *ring*∼*rang, sing*∼*sang*, and the regular *-ed* pattern. The formation of *branged* reflects an attempt to apply the two different analogical patterns to one form.

One fundamental difference between exemplar approaches such as AM and rule approaches is the way each divides contextual space. Contextual space may be thought of as a mapping of forms according to how similar they are. More similar forms are closer to each other in contextual space than less similar

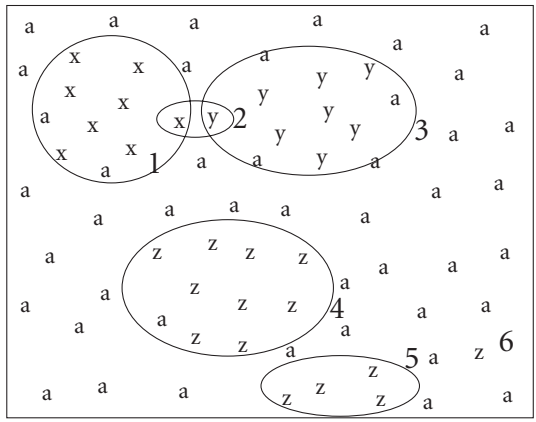

Figure 1. Two-dimensional representation of multi dimensional contextual space

forms. Similarity may be based on semantic, morphological, phonological, or other traits. According to rule models, contextual space is divided with sharp boundaries between forms that behave differently. A given form either meets the structural description of a rule or it does not. A given form either violates or does not violate a constraint; there is no middle-of-the-road and there are no shades of gray. From the analogical standpoint, the existence of partial overlap and gradience is a reality that must be dealt with. Consider Figure 1.

Each letter represents a word in contextual space that exhibits a certain behavior or relationship to another word. The words in Figure 1 belong to one of four groups based on their behavior: a, x, y, or z. The closer any two letters are, the more similar the words (or other linguistic unit) they represent are.

In general, words that behave similarly tend to share linguistic traits and fall closer to each other in contextual space. The major premise of AM is that a given context, if unknown or novel to the speaker, will behave in accordance with the neighbors that surround it. Therefore, if a given context falls in the middle of Circle 1, the majority of its neighbors have the behavior associated with Group X. The chance that one or more members of Group X will be chosen as analogical models for the given context are very high in this case. Of course, if the given context already exists in the mental lexicon and if access to it is not impeded by noise in the system, the closest analog is the word itself making the probability 100% that the word's own behavior will be chosen.

Consider now the members of Group Y. There are members that are not completely surrounded by other Y members. In particular, the member of Group Y in Circle 2 has a neighbor from Group X. AM predicts that under con-

ditions of imperfect memory, in the formation of neologisms, and in language acquisition or language change, there will be some tendency for that member of Y to exhibit the behavior of Group X because it has a close neighbor from X. An example of this is found in English-speaking children who occasionally use **brang* as the past tense of *bring* (Bybee & Slobin 1982). The /-Iŋ/ of *bring* makes it extremely close to verbs such as *sing∼sang, drink∼drank*, and *shrink∼shrank*. The historical change of the past tense of *dive* from *dived* to *dove,* can also be explained in terms of the analogical influence of neighbors such as *drive∼drove*.

5.4 An analogical simulation of Spanish gender assignment

The relationship between grammatical gender and phonological word shape has been approached from various points of view: pedagogy (Bergen 1978; Bull 1965; Teschner & Russell 1984), bilingualism (Clegg 1997; Poplack, Pousada, & Sankoff 1982; Smead 2000; Zamora 1975; Zamora Munné & Béjar 1987), description (Natalicio 1983; Teschner 1983; Rosenblat 1952), generative linguistics (Harris 1985b, 1991a; Klein 1983, 1989), acquisition (Brisk 1976; Pérez-Pereira 1991), & dialectology (García 1998). In this section, I would like to demonstrate the sort of empirical findings that are manifest when gender assignment is assumed to be an analogical process.

5.4.1 *Gender assignment according to analogy and rules*

The first question to grapple with is how analogy compares with rule models. The extant generative analyses (e.g., Harris 1985b, 1991a; Klein 1983, 1989) are not useful for this purpose; they strive to describe gender in terms of a rule system that derives a word's final phoneme(s) given the word's inherent gender and a set of abstract assumptions about the word's underlying structure. In other words, they are not designed to predict a word's gender given its phonological properties. Bull's (1965) pedagogical rules serve as a more appropriate point of comparison. Based on an extensive dictionary search, Bull concludes that most words ending in *-a, -d, -ción, -sión, -tis,* and *-sis* are feminine, while words ending in any other phoneme or combination of phonemes are masculine.

As a test set, I extracted all 2,416 single gender[30] nouns from the Juilland and Chang-Rodríguez (1964) frequency dictionary of Spanish (see Eddington 2002c for details), therefore, this database represents the most frequent nouns in the language. In order to be processed by AM's algorithm, the phonemes of the final syllable of each word were encoded. When Bull's rule is applied to

these 2,416 words a success rate of 95% is achieved. In the analogical simulation, each word was treated as if its gender was unknown and gender assignment was based on analogy to similar words in the database. The simulation yielded a 95.5%[31] success rate which is nearly identical to that of Bull's rules.

Although both approaches performed equally on the database items, analogy was found to be superior on a different task involving assigning gender to unknown words. A group of 31 Spanish speakers was asked to determine the gender of 118 antiquated Spanish words that have fallen out of contemporary usage, such as *sorce* 'small mouse' and *bocacín* 'part of a wagon.' The gender of the most frequent response was used for comparison. Bull's rules corresponded with the majority responses on 75% of the test words, while analogy correctly predicted 81% of them (Eddington 2002c).

5.4.2 *Analogy and the acquisition of gender*

Brisk (1976) noticed that children make more gender errors on feminine words than on masculine words (17.7% vs. 11.5% respectively). However, children whose abilities in Spanish were least developed made more errors that entailed giving masculine words feminine gender when compared with more advanced speakers. This sort of developmental phenomenon can be modeled by assuming that speakers with more advanced abilities have larger vocabularies.

I performed a series of analogical simulations using databases of varying sizes. The purpose of the simulations was to calculate the number of errors that would occur (i.e., masculine nouns misassigned feminine gender and feminine nouns misassigned masculine). The first step was to organize the database of 2,416 words by descending order of token frequency and then to divide the database into ten data sets, each containing about 241 words. The first data set contained the 241 most frequent nouns. The second data set included all of the items in the first one, plus the next 241 most frequent words, and so on until the tenth data set comprised all 2,416 items. This progression of data sets not only corresponds to the fact that language acquisition entails increasing the size of one's mental lexicon, but that more frequent words are learned first and less frequent words at a later stage. The gender of all of the 2,416 items was predicted in each of the ten simulations. Comparing the number of errors on masculine and feminine nouns in the database is a valid procedure since there are almost identical numbers of masculine and feminine items (masc. N = 1,207; fem. N = 1,209).

The outcome of the ten simulations is summarized in Figure 2. Errors on masculine nouns outnumber those committed on feminine nouns in the simulations that used smaller data sets composed of high frequency words.

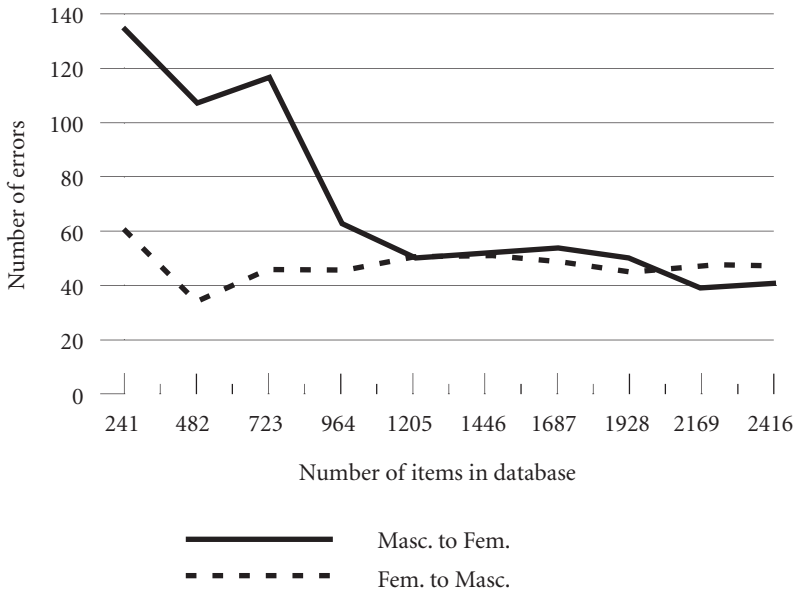

Figure 2. Errors by size of database

However, once the database contained 2,169 words, the number of errors on feminine nouns became slightly higher. This mirrors quite closely the acquisitional data presented by Brisk. According to her, the split that is observed for more advanced speakers occurs because masculine nouns outnumber feminine nouns in children's vocabularies. While this may be true in her study, it does not explain the outcome of the simulations. In the simulations in which errors on feminine nouns outnumber errors on masculine nouns (2,169 and 2,416 items) the percentages of masculine and feminine data set are roughly equal (2,169: 51% masc. 49% fem.; 2,416: 50% masc. 50% fem.). It is unclear how Bull's rules or any other rule-based model of gender could account for these data.

5.4.3 *Markedness and gender*

According to generative theory, the masculine gender must be explicitly specified as the unmarked option. The notion of frequency and markedness are intimately entwined because it is often the case that the unmarked entity is also the more frequent. However, there are reasons besides dominant frequency for considering the masculine the unmarked or default gender in Spanish (see Prado 1982 for an extensive list). For instance, both children and adults favor

the masculine when asked to assign gender to unknown nouns (Clegg 1997; Natalicio 1983; Perez-Pereira 1991; Smead 2000). I suggest that the structure of the nouns themselves is responsible for establishing the masculine as the unmarked member. Plunkett and Marchman (1991) observe that unmarked or default status does not necessarily depend on numerical superiority. Instead, items belonging to the marked category tend to cluster in groups sharing many characteristics. Unmarked items, on the other hand, have less in common and tend to be spread out across contextual space. Bull's rules for gender assignment reflect this state of affairs; words ending in *-a, -d, -ción, -sión, -tis,* and *-sis* are feminine (marked), while words ending in any other phoneme(s) are masculine (unmarked). From an analogical perspective, what this means for gender is that a random throw of the dart onto a map of nouns organized according to phonological similarities has a much higher probability of landing in a neighborhood of masculine nouns, even if they do not dominate feminine nouns numerically. Analogy accounts for and identifies the unmarked member without having to specify that it is unmarked beforehand.

5.5 An analogical simulation of Spanish nominals in *-ión*

The purpose of this section is to demonstrate how the relationship between Spanish verbal and nominal forms can be modeled within AM and how the analogical approach contrasts with rule-based accounts of this phenomenon. The study is based on 939 verbs having corresponding nominal forms that end in *-ión*. The words were extracted from the Alameda and Cuetos (1995) frequency dictionary of Spanish. These verb/noun pairs demonstrate the 14 most frequent verb/noun relationships. I will only discuss a few of these relationships. However, a more detailed treatment is available (Eddington forthcoming). The characteristics of the relevant forms will be discussed as they are introduced.

For the simulation, the verbal form of all 939 pairs was encoded in terms of 12 variables that specify the phonemic makeup of the word and the syllabic structure of the final two syllables minus the final /r/ infinitive marker.

Infinitive	Variables
	210987654321
percutir	per0ku0==0ti
disminuir	dis0mi0=0nui
poseer	0po0se0===0e

A number of questions arise regarding the selection of database items and variable selection. Is a database of 939 items too many or too few? Is it a fair approximation of what Spanish speakers know? This database contains types and not tokens. Would a database generated on the basis of token frequency be more representative?[32] The variables in the database represent phonemes and are organized according to syllables. Would it be better to consider phonetic features or acoustic qualities rather than phonemes? Perhaps some alignment of the variables other than according to syllables would be more psychologically plausible. All of these are valid questions that have yet to be answered. Some progress toward answering these important questions has already begun (Eddington 2004b). Nevertheless, as will be seen, the variables used in this simulation appear to work well even though they may not be optimal.

These nominal and verbal forms have become part of the Spanish lexicon in a number of different ways. Some have been part of the language since it developed from Vulgar Latin. Others were borrowed from Latin at a later stage. Some nominal forms were derived more recently from verbs. However, it is unreasonable to assume that speakers have access to the historical origin of the forms which is why all are included. I do have serious misgivings about this database in another regard because it implies that all the verbs in the database are tightly linked with their corresponding nominal form in the mental lexicon.

It is a well-known fact that derivational relationships between words such as these are much less robust than inflectional relationships (Bybee 1985). Chances are high that a good number of Spanish speakers do not strongly relate verb/noun pairs such as *sujetar/sujeción* 'to subject/subjection.' If the simulation were of an inflectional process, such as predicting the third-person plural subjunctive form on the basis of the first-person singular present tense form,[33] there would be greater certainty that those inflectional forms are strongly associated in the mental lexicon.

It is also not hard to imagine that speakers may know a nominal form such as *desnutrición* 'malnourishment' without knowing the verb *desnutrir* 'to malnourish.' If taken as a model of a speaker's mental lexicon, the simulation assumes that the speaker knows all of the nominal and verbal forms in the database described below. Quite a few of the nominal forms (n=270, e.g., *despolarización* 'depolarization,' *sonorización* 'voicing') have an extremely low frequency. Would all such low-frequency words be known and available as possible analogs for other words? These are the reservations I have about this simulation. However, in spite of these difficulties, a simulation of verbal/nominal correspondences still seems to be an appropriate way of demonstrating the characteristics of an exemplar model and contrasting it with the rule approach.

5.5.1 *Group A*

The most common way to nominalize a Spanish verb is by affixing the suffixes *-ión*, *-ción*, or *-sión* to the verbal stem. The choice of suffix, its exact placement in regard to the stem, and a number of other morphophonemic alternations make nominalization a fairly complex phenomenon. About 75% of the database is made up of verb/noun pairs that have what I call the Group A relationship. The vast majority of these verbs are first conjugation *-ar* verbs, and in reality, the theme vowel /a/ is the only phonetic trait shared by most members; only a small number of verbs are of the *-ir* variety. Nominals in Group A retain the entire verbal stem and theme vowel, and take the suffix *-ción.*

admirar	'to admire'	*admiración*	'admiration'
competir	'to compete	*competición*	'competition'
masticar	'to chew'	*masticación*	'chewing'
narrar	'to narrate'	*narración*	'narration'

Because of the size and phonetic diversity of this group, its members are scattered throughout contextual space. The sheer number of Group A members implies that members of most other groups will have several neighbors from Group A. Therefore, if the nominal form of an A-type verb is not remembered or not known, its neighbors will in most cases influence it to behave as a member of Group A.

A simulation was carried out in which each database item was removed from the database and an analogical set was calculated that contained the word's neighbors. The relationship between the verbal and nominal form was predicting on the basis of the neighbors in the analogical set. In other words, each verb is treated as if its corresponding nominal form were unknown or produced for the first time. Under these circumstances, AM correctly predicted 91% of the nominals in the database. It is important to remember that a major premise of AM is that all morphologically related words have individual representation in the mental lexicon. As a result, it does not pretend to be able to correctly derive all outcomes. It does claim to mirror actual natural language usage such as acquisition, neologisms, slips of the tongue, historical changes, dialectal development, and so on. Predicting the behavior of the Spanish nominals with AM gives insight into the workings of the model. Moreover, certain groups of nominals are particularly interesting in that they allow the analogical approach and the rule-based approach to be contrasted and compared.

The considerable size of Group A has specific consequences. Many members on the fringes of other groups have Group A neighbors (Figure 1). This means that leakage from other groups will often be in the direction of Group

A. Neologisms also feel the pull of Group A verbs; in Spanish, all new additions to the verbal lexicon take *-ar* morphology. This fact coupled with the wide dispersal of Group A words across contextual space means that most nominal neologisms ending in *-ión* are also swept into the Group A pattern.

5.5.2 *Group G*

Group G is the only group that takes the suffix *-ión*. It appears immediately following the stem with no intervening theme vowel. Of the 20 members, 14 end in *-sar.*[34] The 6 other members are *unir* 'to unite,' *reunir* 'to meet,' *desunir* 'to divide,' *opinar*, 'to think,' *rebelar* 'to rebel,' and *coercer* 'to coerce.'

expresar	'to express'	*expresión*	'expresion'
rebelar	'to rebel'	*rebelión*	'rebellion'

With the exception of *reunir*, the members of this group are isolated from each other in contextual space. As a result, the behavior of a given member cannot be predicted on the basis of any other member. In fact, AM's prediction places them in Group A. Therefore, errors and language change are expected to be in the direction of Group A behavior: *anexar* > **anexación* instead of *anexión* (where orthographic *x* is /ks/).

From a rule-based perspective, AM apparently fails to capture the significant generalization that verbs ending in *-sar* take the suffix *-ión*. However, a closer examination of the data reveals that no real generalization exists. While there are 14 Group G verbs ending in *-sar*, there are also 10 Group A words that end in *-sar* (e.g., *acusar* 'to accuse,' *conversar* 'to converse,' *cesar* 'to cease'). AM's apparent failure to see the stem-final /s/ and theme vowel /a/ as the unifying variables of the *-sar* subgroup is attributable to the simple fact that they are not unique to Group G. Of course, in a rule-based approach it is always possible to use a diacritic of some sort to distinguish *-sar* verbs with Group A behavior from *-sar* verbs with Group G behavior. However, in my view, the use of diacritics is undesirable because it allows one to artificially break up contextual space into neat but ad hoc subdivisions. Diacritics are also questionable as psychologically real elements.

5.5.3 *Group I*

All 25 members of Group I end in *-tar*.[35] The suffix *-ción* appears in the nominal in place of the theme vowel and the stem-final /t/.

ejecutar	'to execute'	*ejecución*	'execution'
infectar	'to infect'	*infección*	'infection'
objetar	'object'	*objeción*	'objection'

Of these 25 members, AM predicts that only 6 will have Group I behavior. The remaining 19 are influenced heavily by members of Group A (see Table 7). This is highly reminiscent of the distribution of Group G.

The six members that are correctly assigned to this group (*infectar, inyectar, insertar* 'to insert,' *intentar* 'to try,' *inventar* 'to invent,' and *desinfectar* 'to disinfect') influence each other because they have several traits in common. In addition to ending in *-tar,* each begins with the prefix *in-* and has /e/ as the nucleus of a penult syllable that is closed by one consonant and begins with one consonant (in.CeC.tar).

With the exception of this subgroup, AM is unable to correctly predict the behavior of words belonging to Group I. On the other hand, obtaining the correct outcome is a simple process in rule-based models (Harris 1969: 148–52; Núñez-Cedeño 1993: 151–3). Accordingly, rule application is able to derive the correct nominal form of any verb ending in *-tar*, with the exception of *explotar* 'to explode,' whose nominal takes *-sión*. However, the apparent superiority of the rule-based approach is actually due to the apparent homogeneity of the data.

Harris suggests that *explotar* is the only exception when in reality there is a substantial group of *-tar* verbs whose nominal form would be incorrectly derived by the application of rules. Besides *explotar*, the nominal of *conectar* 'to connect' is *conexión*. Also, 96 verbs that end in *-tar* (e.g., *habitar* 'to inhabit,' *representar* 'to represent,' *excitar* 'to excite,' *exaltar* 'to exalt') have Group A behavior. It would surely be unfeasible to consider 98 forms as exceptions to a rule devised to account for only about 25 forms. Nevertheless, it is always possible in a rule analysis to mark the words in Group I with an abstract diacritic that has no basis in the morphological, semantic, or phonological characteristics of

Table 7. Sampling of Group I outcomes

Infinitive	Gloss	Prob. of I	Prob. of A
**adoptar*	'to adopt'	0.2	99.8
**secretar*	'to secrete'	10.9	89.1
**detectar*	'to detect'	2.8	89.4
infectar	'to infect'	81.6	18.1
inyectar	'to inject'	100	0

* missassigned nominal form.

the words. This is Núñez-Cedeño's solution. He uses the diacritic [±special], that allows *-tar* verbs with Group A behavior to be distinguished from the *-tar* verbs in Group I that do not have Group I behavior.

Such abstractions do not figure into this simulation, which is based on observable surface characteristics. From the standpoint of analogy, there is no real generalization to be made concerning *-tar* verbs. Furthermore, AM makes the empirical prediction that under conditions of imperfect memory many verbs of this type will take the nominal form associated with Group A words.

5.5.4 *Group K*

The suffix *-sión* appears after the stem of the Group K nominals. Five of the verbal forms end in *-uir* and the remaining three in *-Ver.*

corroer	'to corrode'	*corrosión*	'corrosion'
desposeer	'to dispossess'	*desposesión*	'dispossession'
poseer	'to possess'	*posesión*	'possession'
concluir	'to conclude'	*conclusión*	'conclusion'
excluir	'to exclude'	*exclusión*	'exclusion'
incluir	'to include'	*inclusión*	'inclusion'
ocluir	'to obstruct'	*oclusión*	'obstruction'
recluir	'to imprison'	*reclusión*	'imprisonment'

All of the forms ending in *-uir* are correctly predicted by AM to belong to Group K. *Desposeer* and *poseer* serve as the major sources of analogy for each other, which allows them to be successfully linked to Group K as well. *Corroer* is the only member that is sufficiently distant from the others that it cannot be assigned Group K behavior based on analogy to other members.

5.5.5 *Group L*

The unifying factor in Group L is the appearance of *-ción* after the stem of the nominal with no intervening theme vowel. There are two subgroups: words ending in *-uir* and words ending in *-traer*.

construir	'to construct'	*construcción*	'construction'
destruir	'to destroy'	*destrucción*	'destruction'
instruir	'to instruct'	*instrucción*	'instruction'
reconstruir	'to reconstruct'	*reconstrucción*	'reconstruction'
atraer	'to attract'	*atracción*	'attraction'
contraer	'to contract'	*contracción*	'contraction'
distraer	'to distract'	*distracción*	'distraction'
extraer	'to extract'	*extracción*	'extraction'
sustraer	'to subtract'	*sustracción*	'subtraction'

All the *-traer* words are predictable on the basis of other such verbs of the group. However, only two of the four *-uir* verbs (*construir, instruir)* are successfully associated with Group L. The major analog of *reconstruir* is *construir,* but enough similarities are found with Group A verbs that 50.5% of the analogical force is from Group A and only 42.6% from Group L. *Destruir* appears to be even more isolated (2.7% L, 90.7% A).

5.5.6 *Overlap between -uir verbs*

Groups J, K, and L each contain subgroups of words whose verbal form ends in *-uir*. In addition, *fruir* 'to enjoy' and *intuir* 'to sense' belong to Group A. Therefore, an important aspect of any account of how nominal and verbal forms are related is that it is able to distinguish between these similar forms. The *-uir* members of Group K have a great deal in common. Since they end in *-cluir* their behavior may be predicted on the basis of their similarity to the other members of the group. The same is true of the members of Group L, all of which end in *-struir*. The members of Group J bear less similarity to each other since they do not all end in the same word-final phonemes. Nevertheless, they are all predictable on the basis of the similarities they have with other members of the group. Although they do not share exactly the same word-final phonemes, an analysis of the analogical set constructed for each Group J word reveals a group similarity that is not immediately apparent. Not only do they all end in *-uir* but they have a common syllable structure. All Group J verbs have an open penult syllable with the nucleus /i/. These similarities are responsible for the cohesiveness of Group J. In any event, AM also correctly associates *intuir* with Group A. Therefore, the only case of true misassignment made by AM is *fruir* that is assigned to Group L instead of Group A.

The nominalization of the different *-uir* verbs can be handled in a rule-based approach as well, although it requires a bit of abstraction. Harris (1969) and Núñez-Cedeño (1993) propose rules that derive the nominal forms from the verbal forms. They effectively separate the *-uir* verbs in each category by positing abstract underlying forms containing /d/ and /g/ for the *-uir* verbs of Groups K and L respectively. (They make no mention of the two Group A forms).

	Example	Underlying Form
Group J	*instituir*	/instituir/
Group K	*incluir*	/inklu**d**ir/
Group L	*destruir*	/destru**g**ir/

Their inclusion of abstract phonemes in the underlying representation is directly traceable to the perceived need in generative analyses to derive all surface forms from unique underliers. For example, their rules are meant to derive not only the nominal *destrucción* from the verb *destruir,* but the agentive *destructor* and the adjective *destructivo*. The same rules are applied to derive *lector* 'reader' and *lectura* 'reading' from *leer* /leger/ 'to read' even though *leer* does not end in *-struir*. The point here is that instead of simply treating *leer* as an exception to the generalization about verbs ending in *-struir*, rule-based analyses are compelled to derive the morphological relatives of *leer* with the same rules that derive the morphological relatives of *-struir* verbs. For this reason, they do not entertain the possibility that *-struir* could constitute the context for the rules, hence the need to posit an abstract /g/ in the stem.

In contrast to the rule-based approach, distinguishing between the *-uir* verbs in each group is a straightforward task in the analogical approach. All Group L verbs end in *-struir.* This high level of similarity enables AM to predict each member's behavior on the basis of the other members of Group L. As mentioned above, all Group J verbs. This sort of similarity is seen by analogy but poses a problem for a rule-based analysis since there is no unique context for all these verbs. For this reason, the rule-based accounts must resort to abstract diacritics.

5.5.7 *Group M*

All 38 Group M verbs are of the *-er* or *-ir* conjugation and are characterized by a stem-final /d/ or /t/. The corresponding nominals take *-sión* and do not include the stem-final /d/ or /t/.

aludir	'to allude'	*alusión*	'allusion'
emitir	'to emit'	*emisión*	'emission'
persuadir	'to pursuade'	*persuasión*	'pursuasion'
transmitir	'to transmit'	*transimisión*	'transmission'

Only 22 of these forms find a enough analogical pull from other members of the group that their behavior is predicted to be that of Group M. The strongest pull the remaining members feel is from their Group A neighbors.

As far as Group M is concerned, AM predicts the correct nominal form for only about half of the verbs. This suggests that the stem-final /t/ or /d/ is apparently not a sufficiently unifying trait. However, according to Harris (1969: 143–153) and Núñez-Cedeño (1993: 160–70) a stem-final /d/ or /t/ in an *-ir* or *-er* verb triggers the application of rules that derive all the nominal forms in Group M from their respective verbal forms. Harris cites only a handful of

exceptions to his rule. It would appear that the success rate of the rule-based approach far outweighs that of AM in accounting for the behavior of Group M. However, appearances are deceiving since there are actually more than a mere handful of exceptions to the rule.

Besides the *-tender* verbs Harris mentions as exceptions (*atender* 'to tend to,' *contender* 'to contend'), the nominal of *deglutir* 'to swallow' is *deglución* and not **deglusión,* as the rules would derive. *Competir* 'to compete,' *partir* 'to divide,' *repetir* 'to repeat,' and *repartir* 'to distribute' exhibit Group A behavior as do *medir* 'to measure' and *rendir* 'to render.' Several other verbs with Group A behavior that are not included in the database would also be incorrectly affixed by the rules: *expedir* 'to expedite,' *vender* 'to sell,' *hundir* 'to sink,' *perder* 'to lose,' *comedir* 'to exercise restraint.' In addition to these exceptions, section 5.5.8 discusses seven additional forms that would be incorrectly derived by the proposed rules.

Since verbs ending in *-dir, -der, -tir,* and *-ter* are not consistently tied to the same morphophonemic alternation between their verbal and nominal forms, the analogical approach does not find a relationship between all verbs with stem-final /d/ or /t/ and Group M behavior. In other words, AM's inability to consistently predict the behavior of Group M words reflects the fact that no single unique variable unites the group and thus no broad generalization exists. Such evidence again clarifies that the only way to salvage a rule-based analysis would be to through the use of some sort of diacritic to separate the *t*-stem and *d*-stem words that undergo the rule from those that do not.

5.5.8 *Group N*

Six of the seven verbs in Group N have a stem ending in *-fundir,* the exception being *escindir*. The nominal forms take the suffix *-sión* and the stem-final consonants /nd/ are absent:

confundir	'to confuse'	*confusión*	'confusion'
difundir	'to disseminate'	*difusión*	'dissemination'
efundir	'to effuse'	*efusión*	'effusion'
escindir	'to divide'	*escisión*	'division'
fundir	'to fuse'	*fusión*	'fusion'
infundir	'to instill'	*infusión*	'instilling'
transfundir	'to transfuse'	*transfusión*	'transfusion'

The behavior of 4 members of this group is predictable on the basis of their similarity to other members of the group. Given its differing phonological makeup, it is not surprising that the nominal of *escindir* is not predictable on

the basis of other Group N words. However, AM associates *difundir* and *fundir* more closely with Group A than with Group N despite the fact that it shares *-fundir* with five other Group N words. The analogical influence of *fundar* is responsible for this.

Although Harris's and Núñez-Cedeño's analyses do not include this group of verbs, their nominal forms could of course be derived by rules. However, some sort of mechanism would be required so that the rules that apply to *d*-stem verbs in Group M do not apply to the *d*-stem verbs in Group N. Of course, the Elsewhere Condition (Kiparsky 1973) could be invoked to handle this situation. In this case, the rules that derive the nominals from the stem *-fundir*[36] would take precedence over those that derive the nominals from the simple *d*-stem verbs because the former context is more restrictive than the latter.

From the analogical perspective it appears that the Elsewhere Condition is merely an epiphenomenon. In fact, this same point was made by Daugherty and Seidenberg (1994) for connectionist models. There are no instances in which AM assigns a Group M behavior to a Group N word. This is because most Group N words have a great deal in common. The fact that they share the stem *-fundir* allows them to exert a large analogical force on each other. At the same time, the fact that they share stem-final /d/ with verbs from Group M (as well as with some verbs in other categories) is simply not enough similarity for them to be drawn in to displaying Group M behavior. There is no need to invoke the Elsewhere Condition; it simply falls out as a natural result of making analogies based on similarities. The more similarities that exist, the greater the possibility of analogical influence between members of the same group.

5.5.9 *Error analysis*

As a whole, the simulation was able to correctly predict the nominal form of 91% of the verbs when they were treated as novel items. While this number is impressive, the errors made are also telling since AM claims to be able to predict language change, neologisms, and slips of the tongue. To test this I conducted a search of the internet using Google for all of the 81 'erroneous' forms predicted. Using the internet for linguistic analysis is becoming more common (e.g., Rainer 2003) and it has some advantages and disadvantages. On the on hand, it allows for rapid inspection of massive amounts of data. On the other hand, information about the author (e.g., age, sex, country of origin, etc.) is not always readily apparent. Nevertheless, all of the nominals cited below were found on web pages that appear to be written by native Spanish speakers, or at least highly proficient non native speakers. None of the uses ap-

pear to be jocular, nor are they discussed as examples of improper usage or dialectal peculiarities.

Of the 81 'erroneous' forms, 41 are actually attested on Spanish-language web pages.

adapción	*detectación*	*fundición+*	*redactación*
adoptación+	*difractación*	*impulsación+*	*retransmitición*
afectación+	*difundición*	*intervenición*	*revisación+*
anexación+	*disecación+*	*invadición*	*secretación*
asumición	*dispersación*	*objectación*	*sujetación*
cesión+	*editación*	*opinación*	*televisación*
conjuntación	*ejecutación*	*optación+*	*transmitición*
convertición	*exceptación+*	*presumición*	*unición*
decidición	*exentación*	*proyectación*	
desertación	*expandición*	*rebelación*[37]	
destruición+	*expresación*	*recolectación*	

The 30 items not marked with a cross can be considered either slips of the keyboard or neologisms. A word such as *unición,* which appeared in the context *Unición Europea* 'European Union,' could represent an attempt by the writers who used it to nominalize the verb *unir* when the word *unión* was either missing from the mental lexicon or temporarily unavailable from memory. This demonstrates that *unir* is not a central member of its group, but falls closer to Group A (see Figure 1, example 6). This explains why it takes Group A morphology. Exemplar-based models predict the direction that this sort of slippage takes in actual usage.

The 11 words marked with a cross are interesting for another reason. They appear as entries in an online dictionary of the Real Academia Española.[38] These words may be considered to be doublets, that is, two nominals related to the same verb but with differing morphology. What makes these words interesting is the century in which they are first attested. I found this by searching for all possible spelling variants of these words on *Corpus del Español.*[39] As Table 8 indicates, AM predicts the newer nominal form related to the verb,[40] hence, it predict the direction of the language change.

One interesting case is *destruición*. The verb *destruir* originally existed alongside its nominal *destrucción*. However, analogy shows that *destruir* is closer to words with Group A behavior than it is to other member of its group *construir* and *instruir* because its nominal is predicted to be *destruición*. Sometime in the 15th century, the influence of Group A on *destruir* gave it this new nominal. The analogically based *destruición* coexisted alongside *destrucción* un-

Table 8. Attested neologisms predicted by analogy

Newer word predicted by AM		Word in original database	
fundición	16th	*fusión*	14th
adoptación	19th	*adopción*	15th
impulsación	RAE	*impulsión*	15th
afectación	16th	*afección*	15th
revisación	20th	*revisión*	16th
anexación	RAE	*anexión*	17th
disecación	RAE	*disección*	18th
cesión	13th	*cesación*	16th
optación	20th	*opción*	16th
exceptación	RAE	*excepción*	14th
destruición	15th	*destrucción*	13th

RAE = appear in the online dictionary, but not in *Corpus del Español*.

til sometime in the 19th century when the former disappeared in careful writing. However, the analogical influence never disappeared completely, which is why *destruición* may be found on the internet in contemporary contexts.

5.5.10 *Summary of results and conclusion*

The purpose of this section has been to investigate the ability of AM to account for the morphophonemic alternations that hold between Spanish verbal forms and their related nominals ending in *-ión*. The real test of AM is not its ability to correctly assign behavior to all forms when they are treated as novel items. In the first place, all words are assumed to be stored in memory and unless a word is temporarily forgotten, the probability that an item will be correctly predicted is 100%. Analogy accepts gradience and fuzzy boundaries, which necessarily entails that some members of a group are separated in contextual space from other members of the same group. This presents no problem for speakers whose mental lexicon contains the forms and who are able to recall them. However, the above exercise consisted of removing the verbal form from the lexicon and predicting its nominal form on the basis of other verbal forms. Under these circumstances, outlying members will be less influenced by the other members of their group, which explains the 40 misassignments made by the model.

Analogy presents several advantages over rule-based models. The first advantage is that it makes its predictions based on surface-apparent traits. That is to say, there is no need to posit abstract underlying elements or assume any sort of diacritic marks. Along the same lines, there is no need to assume that speakers formulate and store rules or constraints. A local generalization may

always be calculated by reference to stored lexical items if the need arises. Perhaps the greatest advantage is that claims based on storage of exemplars may be tested empirically with an algorithm such as AM. Analogy is also able to predict the sort of slippage between categories that is found in actual language use. About half of the errors made by the model actually exist either as doublets or as attested mistakes or neologisms made by Spanish speakers.

5.6 Accounting for dialectal differences through analogy

One issue that often arises in the linguistic literature is how to account for differences in dialects of the same language. I will use diminutive formation in Spanish to exemplify this. The most frequent diminutive suffix in Spanish is *-ito/a*, which generally attaches to nouns and adjectives, but sometimes affixes to adverbs and gerunds as well. Diminutives express a number of different notions such as familiarity, affection, disdain, or physical size. The allomorphs of *-ito/a* and their distribution have received a great deal of attention (Ambadiang 1996, 1997; Colina 2003; Crowhurst 1992; Elordieta & Carreira 1996; Harris 1994; Jaeggli 1980; Miranda 1999; Prieto 1992). There are essentially three allomorphs: *-ito/a, -cito/a,* and *-ecito/a*. They may attach to entire words (*animal* 'animal'> *animalito; pan* > 'bread' *panecito, pancito; grande* 'large' > *grandecita; tigre* 'tiger' > *tigrecito)* or to stems stripped of their final vowel (*tigre* > *tigrito; casa* 'house' > *casita*). As can be seen, there is more than one way to form a diminutive for a number of words. In my previous corpus study of diminutives (Eddington 2002b), I searched several corpora totaling about 51 million words and identified about 2,460 different diminutives. About 60 of these diminutives have alternative forms. For example, *viento* 'wind' has two possible diminutives, *vientecito* and *vientito*. The diminutive of proper name *Jorge* is attested as both *Jorgecito* and *Jorgito*. A great deal of variation is also found in words containing the diphthongs [je] and [we] in the final syllable of the stem. The diminutive of *hierro* 'iron' is either *hierrecito* or *hierrito*, while the diminutive of *muerto* 'dead' is either *muertecito* or *muertito*. Another place where variation occurs is in words ending in [jo] and [ja]. *Indio* 'Indian' yields either *indiecito* or *indito*, while *rubio* 'blond' has two competing forms *rubiecito* and *rubito*.

A number of formal mechanisms have been proposed to account for these sorts of dialectal variants. For example, Crowhurst (1992) suggests that in some dialects a minimal word template exists that is composed of two bisyllabic feet. Other dialects do not have this template. This difference is thought to account for the dialectal variation in diminutive formation (see also Prieto

1992). Crowhurst explains the alternation between diminutives such as *dientito* 'tooth' and *dientecito* by proposing that in the former, the diphthong of the stem is resyllabified in the course of the derivation in such a way that each of its components belong to separate syllables. In the case of *dientecito*, no such resyllabification occurs. Colina (2003) approaches the question from the standpoint of Optimality Theory. According to her analysis, all variation may be accounted for by assuming that the constraints that govern diminutive formation are ranked differently in each dialect.

As I argued in Chapter 1, formal mechanisms such as constraints may be excellent tools for describing linguistic phenomena, but they do not relate to the actual mechanisms speakers use to process language. I believe than a process of analogy more closely captures actual performance mechanisms. Using AM (Eddington 2002b), I predicted the diminutive formation of the 2,460 base words found in the corpus search. When these words were treated as if their diminutive form were previously unknown, 96% of them were given an attested diminutive form. I will use this database to show that dialectal differences in diminutive formation may be explained in terms of storage and analogy to stored base/diminutive sets.

Consider a dialect[41] in which diminutives containing [je] and [we] in the final syllable of their stems are of the sort *dientito* and *muertito*, in contrast to *dientecito* and *muertecito.* Assume that every diminutive known by a speaker of this dialect is stored in his/her mind and is connected to the base word on which the diminutive is formed. Provided that the speaker is not having memory problems, s/he will use the same diminutive form the next time it is needed. What is of greater interest is how the speaker will form the diminutive of a word whose diminutive form s/he has never heard or produced before. This may be simulated using the database from the previous study. One difficulty with the database is that it contains diminutives produced by speakers of many different dialects of Spanish. However, it may be altered to simulate the dialect in question. To this end, I modified the entries in the database so that every word with [je] and [we] in the stem of the base word took a diminutive of the type *dientito* and *muertito*. I then ran AM so that it predicted the diminutive of 21 words containing [je] and [we] that do not appear in the database. The 2,460 words in the database were available as analogs. These results appear in Table 9 under Dialect A.[42]

In order to simulate a dialect, such as Peninsular Spanish, in which the diminutives of *diente* and *muerto* are *dientecito* and *muertecito*, rather than *dientito* and *muertito,* the database was revised to reflect this sort of relationship between base words and diminutives. A comparison of Dialects A and B in

Table 9. Probabilities of variant forms in two simulated dialects

Base word	DIALECT A		DIALECT B	
	Prob. of *-ito*	Prob. of *-ecito*	Prob. of *-ito*	Prob. of *-ecito*
cuerdo	100	0	10.17	89.83
diestro	99.95	0	14.94	84.98
fuero	100	0	0.12	99.88
mueble	100	0	1.14	98.86
pienso	99.98	0	8.41	91.37
pliegue	99.99	0	3.50	96.48
riego	100	0	0.21	99.78
ruego	100	0	0.12	99.87
siervo	99.99	0.01	1.88	99.08
trueno	100	0	6.14	93.86
	Prob. of *-ita*	Prob. of *-ecita*	Prob. of *-ita*	Prob. of *-ecita*
cuelga	100	0	1.13	98.87
cuerda	100	0	9.21	90.61
fiebre	97.67	1.68	0.30	97.37
friega	100	0	16.97	82.98
huerta	100	0	6.97	92.86
niebla	100	0	0.25	99.75
nieve	92.09	7.70	0	92.09
prueba	100	0	1.10	98.90
suerte	96.12	1.11	0.33	95.69
sierva	100	0	0.65	99.30
tuerca	100	0	7.52	92.40

Table 9 clearly demonstrates that analogy can account for dialectal variation. Children who are raised speaking Dialect A produce diminutives of this type because they have access to the forms that already form part of their mental lexicon. They use these same forms to analogize on when the formation of a diminutive is required, and analogy will predict forms similar to those they already know. The possibility also exists that in a particular dialect there is a degree of variation between forms such *dientito* and *dientecito,* in which case analogy would be expected to predict both forms to a certain degree.

Consider another variation in diminutives that has occupied scholars. Words ending in [jo] and [ja] may form two types of diminutives. For example, *rubio* may yield *rubiecito* or *rubito*. The original database was modified so that all words ending in this way took either diminutives of the type *rubito* (Dialect A) or of the type *rubiecito* (Dialect B). The diminutive form of the words in

Table 10. Probabilities of variant forms in two simulated dialects

Base word	DIALECT A		DIALECT B	
	Prob. of *-ito*	Prob. of *-ecito*	Prob. of *-ito*	Prob. of *-ecito*
cambio	99.96	0	28.57	71.38
diario	100	0	16.68	83.32
egipcio	100	0	3.88	96.12
labio	100	0	17.37	82.63
precio	99.48	0.38	4.59	95.24
	Prob. of *-ita*	Prob. of *-ecita*	Prob. of *-ita*	Prob. of *-ecita*
ansia	100	0	38.04	61.81
biblia	100	0	6.24	93.76
copia	100	0	36.60	63.40
familia	100	0	3.19	96.81
lluvia	99.98	0	0.04	99.86
patria	100	0	27.47	72.52

Table 10, which are not in the database, were then predicted by analogy to all 2,460 items in the database.

The point I wish to make it that there is no need to postulate differences in constraint orderings or rule orderings in order to account for dialectal differences in diminutive formation. Children simply learn the diminutives used in their community and those diminutives themselves are the models for subsequent diminutive formation.

5.7 Conclusions

There are a number of advantages to accounting for language processing via analogy. First, simulations may be performed that are empirically testable and robust. Analogy assumes only storage of known forms and the ability to find similarities and apply similar behavior. Second, it does not require speakers to glean generalizations from the data and formulate them into systems of rules or constraints. Some of these formal systems are extremely challenging for trained students of linguistics to understand and devise, yet if taken as actual performance mechanisms (e.g., Bromberg & Halle 2000), they are assumed to be subconsciously arrived at by native speakers with no formal training in linguistic analysis.

Chapter 6

Diphthongs, syllables, and stress
Beyond formalisms

6. Introduction

The processes of diphthongization, syllabification, and stress assignment have held a prominent position in the literature on Spanish phonology. They have received treatment in formal frameworks ranging from classical generative to autosegmental and Optimality Theory. There are numerous interesting phenomenon in Spanish phonology. A discussion of these processes, to the exclusion of others, is included in this chapter simply because they have been the most studied from an empirical standpoint.

6.1 Diphthongization

One of the most widespread phonological alternations in the Spanish language is the alternation between unstressed [o] and the stressed diphthong [wé], and between unstressed [e] and its stressed counterpart [jé]. These alternations appear in inflectional morphology: *c[wé]nto∼c[o]ntámos* 'I count∼we count;' *v[jé]ne∼v[e]nís* 's/he comes∼you come.' It is also prevalent in derivational morphology: *b[wé]no∼b[o]ndád* 'good∼goodness;' *pim[jé]nta∼pim[e]ntéro* 'pepper∼pepper shaker.'

These alternations may be traced to the open vowels /ɔ/ and /ɛ/ in Romance. In Spanish, they became diphthongs when stressed, but closed to [o] and [e] elsewhere. The situation is somewhat more complex in the contemporary language since /ɔ/ and /ɛ/ have disappeared. In the first place, not every stressed [o] or [e] is realized as a diphthong: *c[ó]ses∼c[o]sémos* 'you sew∼we sew;' *t[ó]sen∼t[o]séis* 'they cough∼you cough.' Secondly, certain processes produce an alternation between stressed and stressless diphthongs: *p[we]blíto∼p[wé]blo* 'small town∼town;' *m[jé]do∼m[je]dóso* 'fear∼afraid.' Thirdly, some apparently diphthongizing vowels fail to appear with a diphthong even though they are stressed: *c[wé]sta∼c[o]stár* 'it costs∼to cost,'

but *c[ó]sto* 'cost' not **c[wé]sto; d[jé]z∼d[e]cimál* 'ten∼decimal,' but *d[é]cimo* 'tenth' not **d[jé]cimo.*

Given this messy situation, the main thrust of formal diphthongization analyses has been to show how cases of stressed mid-vowels and of stressless diphthongs may be accounted for. These involve various systems of diacritic marks on the vowels, empty vowel slots, and cyclic rule application (e.g., Carreira 1991; García-Bellido 1986; Halle, Harris, & Vergnaud 1991; Harris 1969, 1977, 1978, 1989b). None of these diacritics is surface apparent, which is why these analyses escape any sort of empirical verification or refutation. An additional difficulty with diacritics is that the are essentially ad hoc mechanisms that assume discrete, not overlapping categories while performance studies show that linguistic categories have fuzzy boundaries. In my view, the most interesting question is how Spanish speakers process the diphthong alternations, not whether linguists can fit the Spanish data into one formal framework or another.

Bybee and Pardo (1981) devised an experiment to determine whether diphthongization is a psychologically valid process or is merely a historical remnant. Their study involved testing the extent to which diphthongization is applied to new words. They presented speakers with paragraphs containing nonce words in this format:

> *La mamá de Osito lo bierca mucho. Cada vez que se enferma lo bierca. Ayer, Osito se enfermó y su mamá lo ___.*
> 'Little Bear's mother nonce word him a lot. Every time he gets sick she nonce word him. Yesterday Little Bear got sick and his mother _____ him.'

If the subjects had some sort of internalized rule of diphthongization and applied it, they should answer by changing the stressed diphthong in the nonce verb into a stressless mid-vowel (i.e., *bercó)*. However, 73% of the subjects preferred to leave the diphthong intact in this sort of question, and responded *biercó.*

Other test items contained morphological variants of the nonce words that contained both stressed diphthongs and stressless mid-vowels. For example,

> *Osito muena sopa todos los días. Le gusta mucho monar la sopa. Ayer en la tarde se ___ un plato grande.*
> 'Little Bear nonce word soup every day. He likes to nonce word soup. Yesterday afternoon he _____ a big dish of it.'

Under these circumstances, 76% of the subjects gave responses such as *monó* that avoided stressless diphthongs. However, only 36% of the subjects con-

sistently used stressless mid-vowels and stressed diphthongs as rules would predict. I conducted a similar experiment and found essentially the same results (Eddington 1998). Bybee and Pardo also discovered that the subjects were more apt to apply the diphthong/mid-vowel alternation [jé]∼[e] than they were the alternation [wé]∼[o], which is not consistent with rule approaches that handle both alternations with the same mechanism. We can conclude from these studies that the diphthongization alternation is somewhat productive, but not to the extent that formal analyses would predict. Although the subjects in the studies did not unwaveringly apply an internalized rule of diphthongization, they were aware of the alternation to some extent since they could apply it correctly in many cases.

One question that remains is how speakers know which mid-vowels alternate with diphthongs (*c[wé]sta*∼*c[o]stár*) and which do not (*t[ó]sen*∼*t[o]séis*). The kinds of diacritics that have been proposed to differentiate between diphthongizing and non-diphthongizing mid-vowels (see Section 2.3.6.2) are abstract in that they do not exist in the surface forms of the words. Is there some surface-apparent way of knowing if a mid-vowel alternates with a diphthong, or is it simply a matter of memorizing specific words and their morphemic relatives on a case-by-case basis? I concur with Bybee and Pardo that to a large extent the alternation is tied to specific words with a small degree of productivity.

Albright, Andrade, and Hayes (2001) attempted to determine whether there are any surface clues that allow diphthongizing and non-diphthongizing vowels to be distinguished. They did so by computational means, beginning with a database of 1,698 Spanish verbs containing mid-vowels. The stressed and stressless allomorphs of each verb (e.g., [empes-], [empjés-] of the verb *empezar* 'to begin;' [kos-], [kós-] of the verb *coser* 'to cook') were treated as related pairs by the computer algorithm. The algorithm generated rules that are specific to one pair of allomorphs, as well as more general rules that apply to many pairs. Rules are evaluated in terms of how many different lexical items they apply to. The reader is referred to Albright and Hayes (1999) for a detailed description of the algorithm. A total of 3,346 rules were derived from the database of verbal stems pairs, but the authors insist that most of the rules do not have the chance to apply, because in their model more general rules take precedence over more specific ones.

The algorithm calculates the probability that a stem will appear with a mid-vowel or a diphthong. For instance, it finds that there are a number of phonologically similar diphthongizing verbs such as *cerrar* 'to close,' *enterrar* 'to bury,' *encerrar* 'to enclose,' and *desterrar* 'to exile.' Based on these verbs, the

model predicts that verbs ending in *-errar* have a 92% likelihood to be diphthongizing. On the other hand, no verb with the structure *-echar* (e.g., *echar* 'to throw,' *aprovechar* 'to take advantage of,' *sospechar* 'to suspect') is diphthongizing, which results in an extremely low predicted probability that the [e]∼[jé] alternation will apply to verbs with this phonological shape, and a high probability that stressed [é] will appear.

To test the predictions of the model against Spanish speakers' intuitions the researchers devised 33 nonce verbs and presented them to Spanish speakers in the context of a paragraph, much as in the study by Bybee and Pardo (1981). The nonce verbs and the subjects' responses were given orally. The verbs were all presented with an inflection in which the mid-vowel was stressless. For example, the subjects were provided the verb *lerrámos* and asked to provide forms in which the test vowel was either unstressed *(lerrádo, lerrár*) or stressed (*lérro* or *liérro*). The crucial test was whether the stressed form would yield a diphthong or not. After the subjects provided an answer, the experimenter read the test sentence back to them with both possible answers (e.g., *lérro* and *liérro*). They were then asked to rate each answer on a scale of one to seven in terms of how good each sounded.

Correlations were calculated between the percentage of times subjects produced diphthongs and the probability predicted by the algorithm. A significant correlation (r = .510) was obtained. An even greater correlation resulted (r = .838) with the speakers' introspective ratings of how good each possible answer sounded. It is important to note that the contexts [o]∼[wé] and [e]∼[jé] in which diphthongization was likely to occur were calculated separately. Albright et al. found that the correlation diminished from .510 to .424 when the contexts for both alternations were merged. This supports the findings of Bybee and Pardo (1981) that there are two processes of diphthongization even though from a formal standpoint the two may be conflated.

Albright et al. assume that the ability of the subjects to use phonemic context to differentiate between diphthongizing and non-diphthongizing stems is developed in the course of acquisition:

> Children comb through the data, looking for generalizations about phonological environments. When the data don't pattern cleanly, the result is a rather messy set of conflicting learned generalizations. We further hypothesize that tacit knowledge of these generalizations persists into adulthood and can be detected experimentally. (2001:118)

It is hard to imagine that children need to calculate 3,346 rules in order to handle 1,698 verbs. If something analogous to this magnitude of rule formulation

is carried out during acquisition, imagine how many millions of rules must be devised in order to cope with the thousands of phonological, morphological, and syntactic processes in the Spanish language. Prototype effects of the sort that Albright et al. report for diphthongization are not limited to language, however. Psychologists have shown that such effects are central to all of human cognition. If the massive rule induction model of Albright et al. is taken as a general performance model of psychology, that would mean that the subconscious mind would spend an enormous amount of time calculating rules and probabilities for all of life's experiences in childhood that would be stored and later accessed in adulthood. In my view, it is more reasonable to assume that words (and other past experience) are stored, and calculations of contextual similarity are only performed when needed, such as when one is asked to inflect nonce verbs.

Nevertheless, the study by Albright et al. demonstrates that speakers use phonemic material in verb stems in order to decide whether a nonce word has a diphthongizing stem or not. In Eddington (1996, 1998), I showed that suffixes are also correlated with diphthongization. I considered the diminutive, superlative, and augmentative suffixes *-(c)ito, -zuelo, -(c)illo, -ísimo*, and *-azo,* all of which are fairly productive. In actual usage, diphthongizing stems maintain their diphthongs when one of these productive stems is added: *pueblo > puebl(ec)ito, fuerte > fuertezuelo, vieja > viej(ec)illa, buena > buenísima, bueno > buenazo.* Diphthongless versions of words with these suffixes, such as *bonísimo*, do appear in some dictionaries but are extremely stilted and do not occur in normal speech. The only common exception is *caliente > calentito*. Because these more productive suffixes are free to attach to almost any noun or adjective, a representative sample of words containing these suffixes is not likely to be found in any dictionary.

In addition to the more productive suffixes, I considered the less productive suffixes *-al, -(i)dad*, *-ero*, *-oso*, and *-ista*. I searched a dictionary for words with these suffixes that have diphthongizing stems.[43] These results were presented to five college-educated Spanish speakers who rated each word on the likelihood that the average Spanish speaker would be familiar with it. After removing the uncommon words from the results of the dictionary search, I calculated for each suffix the likelihood that the diphthongizing stem occurs with a mid-vowel instead of a diphthong. Usage suggests that the more productive suffixes rarely occur with mid-vowels (*pueblo > *poblito, rueda > *rodezuela, nueva > *novísima),* but this is not evident in a dictionary search.

When a large number of existing words is considered it becomes apparent that rule-based analyses are hard pressed to account for all of the forms.

Table 11. Co-occurrence of diphthongizing mid-vowels and certain suffixes

Suffix	% of mid-vowels	Example	Gloss
-ero	100	*herrero*	blacksmith
-al	100	*dental*	dental
-(i)dad	100	*novedad*	news
-oso	83	*vergonzono*	shameful
-ista	50	*huelguista*	striker
-(c)ito	low	*cuerp(ec)ito*	dim. of body
-zuelo	low	*cuentezuela*	dim. of bill
-(c)illo	low	*huesillo*	dim. of bone
-ísimo	low	*ciertísimo*	very certain
-azo	low	*buenazo*	good-natured

For example, Harris (1969, 1977, 1989b) suggests that the surface alternation between mid-vowels and diphthongs is the result of different morphological composition in the deep structure. A word containing a diphthong, such as *viejito*, is derived from an underlying form with this morphological structure: [[vej] ito]. Words such as *vejez*, that have no surface diphthong appear in underlying representation as [vej + ez] and not *[[vej] ez]. The difficulty with this proposal is its inability to account for variability in extant forms such as between *calientito*∼*calentito* 'dim. of hot,' and *fervientísimo*∼*ferventísimo* 'dim. of fervent.' It is possible to derive *calientito* from [[calent] ito], and *calentito* from [calent + ito], but their morphological structure is actually identical. In reality, the only motivation Harris gives for assuming different morphological structure for words such as *viejito* and *vejez* is that they differ in regards to diphthongization. Postulating different morphology, in this case, can be regarded as a sort of ad hoc abstract diacritic mark (Hooper 1976: 45).

In a later framework, Halle, Harris, and Vergnaud (1991) view diphthongization as a process that is related to certain suffixes. They present an analysis in which diphthongizing stems ending in the suffixes *-oso* and *-(i)dad* pass through the cyclic rule system in such a way that no diphthong is derived (e.g., *vergonzos*o, 'shameful,' *bondad,* 'goodness,' **vergüenzoso, *buendad)*. Other suffixes such as *-ez* and *-(c)ito* are thought to yield stems that contain diphthongs (e.g., *viejita* 'dim. of old,' *piernita* 'dim of leg,' **vejita, *pernita)*. However, in formulating their analysis, Halle et al. failed to consider the existence of words such as *calentito,* which ends in *-(c)ito* and yet has no diphthong, as well as *mierdoso* 'shitty,' *aspavientoso*, 'affected person,' and *miedoso* 'afraid' that have diphthongs yet end in *-oso.* This is clearly another case in which a formal analysis fails because of the meager database on which it is founded.

Thus far, I hope to have demonstrated that there is a relationship between certain suffixes and diphthongization, although the relationship involves some variability. The interesting question is whether the relationship is psychologically relevant for Spanish speakers. I conducted two studies to this end. In Eddington (1996), 51 native Spanish speakers were shown one nonce word and one neologism ending in each of the ten suffixes studied. Their task was to choose between a word with and without a diphthong. For example, they were asked which word would best fit the definition 'having the quality of honey,' *mieloso* or *meloso.* In Eddington (1998), 69 Spanish speakers were shown four neologisms ending in each of the ten suffixes. They again had to choose which response they preferred:

> *Por supuesto, trabajar con estiércol no era agradable. Pero si el único trabajo que había consistía en estercolar los campos, él trabajaría de* _____.
>
> *(a) estiercolero*
>
> *(b) estercolero*
>
> 'Of course, working with manure wasn't pleasant. But if the only work there was consisted of manuring the fields, he would work as a _____.'

Since the tasks were so similar, I combined the responses from the two questionnaires. These are compared with the words from the dictionary search that were judged to be known by most Spanish speakers.

It is not hard to see a marked preference for mid-vowels over diphthongs in nonce words and neologisms ending in *-al, -(i)dad*, and *-ero*. This corresponds with the paucity of common words that end in these suffixes and have stem diphthongs. Words ending in *-oso* and *-ista* demonstrate more variation between mid-vowels and diphthongs, both in the subjects' preferences and in

Table 12. Comparison of diphthongization in common words and in survey responses

Suffix	% of mid-vowel responses	% of mid-vowels in common words
-ero	72.5	100
-al	66.7	100
-(i)dad	64.8	100
-oso	59.8	83
-ista	50.0	50
-ísimo	49.6	low
-zuelo	47.9	low
-(c)ito	26.4	low
-azo	24.6	low
-(c)illo	22.5	low

common extant words. The diminutive, superlative, and augmentative suffixes are the most likely items to appear with diphthongized stems.

When taken together, the experiments reviewed above indicate that diphthongization is not the sort of black-and-white process that rule-based approaches assume. The variation in diphthongization exemplified in words ending in *-oso* and *-ista* is a case in point. However, the variability in the subjects' answers is not random, but follows the same degree of gradience found in existing words in the language. Speakers are able to use not only suffixes but the phonological shape of the stem (Albright et al. 2001) in formulating their preferences. They are able to extrapolate diphthongization correlations from existing words and apply them to novel items in a systematic fashion.

This finding, however, does not necessarily entail the application of some sort of diphthongization process each time a known word, such as *puedo* 'I can,' is produced. If all known words are stored as wholes in the mental lexicon, such processing would be unnecessary. Bear in mind that storage of whole words does not imply that words are stored as isolated entities. Instead, words have massive networks of connections to other words via semantic, phonological, syntactic, and associative links (Bybee 1985, 1988). In this regard, the diphthongization alternation is represented as a recurring phonological pattern that is found in many interrelated lexical items. This pattern may be extracted when the need arises, such as when the phonological shape of novel words needs to be computed.

6.2 Syllables

Early generative research into phonology did not consider the syllable a relevant linguistic unit. However, a major shift took place in the early 1980s that emphasized the utility of the syllable in phonetic processes. For example, the process of /s/ aspiration was originally thought to occur if the /s/ were either in pre-consonantal or word-final position. However, these two contexts may be collapsed, and the analysis rendered simpler, when one observes that pre-consonantal and word-final position both correspond to the rime position of the syllable. A number of proposals emerged about how words and phrases are to be parsed into syllables in Spanish (e.g., Harris 1983, 1989a, 1993; Hualde 1991; Núñez-Cedeño 1986; Núñez-Cedeño & Morales-Front 1999). Amidst all of the excitement over the newly rediscovered phonological unit it seems that one important factor was overlooked – whether the syllable is a significant unit for the speakers of the language themselves and not merely a unit convenient for formal linguistic analysis.

What makes the syllable difficult to delimit is that there are no phonetic properties that correspond to syllable boundaries, which makes the syllable a strictly phonological entity. One piece of evidence for the reality of syllables is that Spanish speakers demonstrate a high degree of consistency in deciding how to divide words into syllables and in counting the number of syllables in a word (e.g., Jiménez-González & Ortiz-González 1994). However, some of this consistency may be due to the explicit training that literate speakers receive when being taught to read. Schnitzer (1999) found some disagreement among illiterate Spanish speakers in regards to how to syllabify certain words. For instance, some of his subjects divided *abstinencia* 'abstinence' *abs.ti.nen.cia* and others *ab.sti.nen.cia*. Hualde (1999) and Hualde and Prieto (2002) also provide some evidence that syllable boundaries are variable where a sequence of vocoids can be interpreted either as a diphthong or a hiatus. That is, a word such as *barriada* 'neighborhood' may be interpreted either as *ba.rria.da* or *ba.rri.a.da.* Although most speakers' intuitions coincide regarding where syllable boundaries fall in Spanish, the cases of variation suggest that the syllable is not a totally discrete category.

6.2.1 *The bigram frequency trough hypothesis*

Seidenberg (1987, 1989) and Seidenberg and McClelland (1989) studied bigram frequencies, (i.e., how often two phonemes occur together) in English words and hypothesized that the syllable has no independent existence. Instead, they suggest that syllable boundaries are an epiphenomenon. Consider the word *anvil.* The letters AN occur together 289 times in the frequency count they used, the letters NV only 5 times, and the letters VI 324 times.[44] The syllable boundary must fall between N and V, because those letters have a bigram frequency of only 5 and are surrounded by bigrams with higher frequencies. Therefore, syllable boundaries may be explained in terms of bigram frequency alone. Since syllable boundaries coincide with a bigram frequency trough there is no need to assume the existence of the syllable as an entity.

Rapp (1992) tested the bigram frequency trough hypothesis in English by briefly presenting multicolored words to the subjects and then asking them to remember what color a specific letter was. Some of the stimulus words were coded so that the colors of the letters corresponded to the syllables in the word. For example, SIG, the first syllable of SIGMA, was green and MA was yellow. In these cases, when the subjects were asked what color G was, most correctly indicated that it was green. In other cases, the colors and syllables did not correspond so that SI was green and GMA was yellow. If the subjects were asked to decide what color the letter G was, and GMA was yellow, they would have

trouble because G has the color that corresponds to the second syllable, yet G belongs to the first syllable. This mismatch caused the subjects to perceive that G was green even though it was actually yellow. In other words, syllable structure influenced the subjects' color perception. In her experiment, Rapp included test words in which the syllable boundaries coincided with bigram frequency troughs and words with syllable boundaries that did not coincide with bigram frequency troughs. Her results indicate that the subjects' responses were influenced by syllable boundaries whether or not those boundaries coincided with troughs. In other words, syllables have an existence independent of bigram frequency troughs. Of course, a study on the English language does not indicate the role of the syllable in Spanish. In fact, a number of other studies suggest that the syllable is not a relevant processing unit in English (Bradley, Sánchez-Casas, & García-Albea 1993; Cutler, Mehler, Norris, & Seguí 1986; Schiller1999), nor in Dutch (Schiller 1998), but is significant in processing French (Ferrand, Segui, & Grainger 1996; Mehler, Dommengues, Frauenfelder, & Seguí 1981).

Carreiras, Álvarez, and de Vega (1993) tested the bigram frequency trough hypothesis with Spanish speakers. They employed two experimental paradigms. In the naming experiments, subjects were shown extant words and nonwords on the computer screen one at a time. Their task was to name the words as fast as possible. The time required to name each word was measured. The other experimental paradigm was a lexical decision task in which the subjects saw Spanish words and nonwords and had to press one key if the string that appeared on the screen was a word and another key if it was a nonword. This also allowed reaction times to be measured.

None of the words that were used in the experiments had bigram frequency troughs that corresponded to syllable breaks. If the syllable is merely an illusory entity resulting from bigram frequencies no syllable effects would be expected. In contrast to this position, Carreiras et al. found significant syllabic effects under both experimental conditions. More specifically, words composed of low-frequency syllables elicited quicker responses than words composed of high-frequency syllables. At first, this may seem counterintuitive. However, Carreiras et al. (1993:770) explain that

> high-frequency syllables are shared by a larger number of words than low-frequency syllables … Assuming that it takes more time to select a word from a large set than from a small one, then it follows that high-frequency syllables should produce slower processing times than low-frequency syllables

Remember, none of the stimulus words contained bigram frequency troughs that coincided with syllable boundaries, nevertheless, syllable frequency effects were found. The negative correlation between syllable frequency and reaction times has been found in other studies as well (e.g., Álvarez, Carreiras, & Taft 2001). As far as visual word recognition is concerned, the frequencies of both syllables of disyllabic words make their influence felt, but the first syllable of the word appears to exert the most influence (Álvarez, Carreiras, & de Vega 2000).

6.2.2 *Syllabic priming effects*

Priming is said to occur when the presentation of one word, the prime, affects the reaction time of a word presented later, called the target. Domínguez, de Vega, and Cuetos (1997) carried out an experiment in order to determine whether syllabic overlap between prime and target can influence the subjects' reaction time to the target. In the first experiment, called a lexical decision task (LDT), the subjects were presented Spanish words and nonwords and had to respond by pressing one of two keys to indicate whether the word was a word or nonword. All of the target words directly followed the presentation of the primes. The test was structured so that the subjects only saw each target once. Consider the examples below, where the syllable boundary is indicated here with a period. (No period appeared during the experiment.)

Prime	Target	Relationship
nor.ma	*nor.te*	same first three letters, same syllable structure
no.ria	*nor.te*	same first three letters, different syllable structure
sa.via	*nor.te*	different first three letters, different syllable structure
man.do	*nor.te*	different first three letters, same syllable structure

If the priming effect were caused by orthographic overlap alone, both *nor.ma* and *no.ria* should influence the response time to *nor.te* equally, but this was not the case; primes such as *nor.ma* caused slower reaction times to targets such as *nor.te* but not to *no.ria*. Therefore, the effect must be due to syllabic overlap as well. If abstract syllable structure devoid of segmental content exerted its influence, one would expect *man.do* to prime *nor.te* because they have the same syllable structure, however, this was not the case. A slowing of reaction times such as the one that occurs when target and prime share the same syllable is called inhibitory priming. In the above experiment, *nor.ma* had already been seen before the presentation of *nor.te* which means that *nor.ma* was still partially activated in the mind when *nor.te* was presented. Since *nor.ma* and *nor.te* are syllabically similar, recognition of *nor.te* competes with partially activated *nor.ma* which slows down the time required to recognize *nor.te* as a legitimate

Spanish word. The same degree of competition is absent when *no.ria* primes *nor.te* suggesting that word recognition in Spanish is carried out by syllabic similarity not merely according to orthographic similarity.

In the second experiment, Domínguez et al. (1997) replicated the first study but manipulated the syllable frequency of the primes. That is, primes were used whose initial syllable frequencies were higher than those of the target stimuli.[45] The reaction time to targets such as *car.ta* slowed if they were preceded by primes such as *car.ne*, as in the first study. The difference here is that orthographic effects also emerged; reaction times to targets such as *car.ne* were also influenced by primes such as *ca.rro* that share the same first three letters, but not the same syllable structure. That orthographic overlap can influence a lexical decision task implies that visual word recognition is sensitive to orthographic similarity, which is not surprising. However, syllabic overlap was also found to be relevant. The fact that the overlap slows reaction time is explained in this way (Domínguez, de Vega, & Cuetos (1997:415):

> The role of syllables is to activate a pool of competing lexical candidates for selection. These candidates share the first syllable. At a later stage, as new syllables are processed, an inhibitory mechanism operates in the lexicon to reduce or suppress the activation of those lexical candidates that partially match the input syllable description. Finally, a single candidate – the one selected – will remain activated.

This suggests that at least under certain conditions, words visually presented are parsed into syllables and that one port of access to the mental lexicon is syllable-based.

In some experimental paradigms, priming results in speeded reaction times rather than slowed reactions. For example, Carreiras and Perea (2002) used primes in which only part of the word was visible:

Prime	Target	Relationship
*ca*****	*ca.si.no*	*ca* is the first syllable of the target.
*cas****	*ca.si.no*	*cas* is not the first syllable of the target.
*car****	*car.tel*	*car* is the first syllable of the target.
*ca*****	*car.tel*	*ca* is not the first syllable of the target.

Under these conditions, facilitative priming occurred when a prime's visible letters corresponded to a target's initial syllable (*car**** / *car.tel*), but reaction times were not speeded in cases such as *ca***** / *car.tel* where the orthographic overlap does not coincide with the syllabic overlap.

Not all experiments of this sort rely exclusively on visually presented words. Costa and Sebastián-Gallés (1998) first trained subjects to respond verbally to a set of pictures by naming them out loud. Later, during the experiment, the subjects' reaction times were measured as the interval between the appearance of the picture on the computer monitor and the beginning of the subjects' verbal responses as registered by a voice key. The primes in this study were presented to the subjects auditorily 150 milliseconds after the picture appeared on the screen.

Prime	Target	Relationship
mo.ra	mo.no	same initial phonemes, same syllable structure
cu.ña	mo.no	different initial phonemes, same syllable structure
mos.ca	mo.no	same initial phonemes, different syllable structure
cul.pa	mo.no	different initial phonemes, different syllable structure

Phonemic overlap produced speeded responses. That is, the subjects named the picture of the monkey (*mo.no*), for example, more quickly if they heard either *mo.ra* or *mos.ca* during the presentation of the picture. Overlapping syllable structure also facilitated picture naming, but to a lesser degree than phonemic overlap. It is not surprising that *mo.ra* primed *mo.no* since they share initial phonemes and syllable structure. However, even words such as *cu.ña* primed words such as *mo.no* when the only thing they have in common is abstract syllable structure.

Another variant of the picture naming paradigm involves naming pictures as well as written words that appear on the screen (Costa & Sebastián-Gallés 1998). Prior to the presentation of the target pictures, the subjects had to name a series of four to seven visually presented words, all with the same syllable structure. Note that the primes and targets do not demonstrate phonemic overlap.

Primes	Target	Relationship
ces.ta, bol.so, sal.to, car.ta	pin.za	same syllable structure
ces.ta, bol.so, sal.to, car.ta	pi.no	different syllable structure
ce.sa, bo.lo, sa.la, ca.pa	pi.no	same syllable structure
ce.sa, bo.lo, sa.la, ca.pa	pin.za	different syllable structure

The idea behind this task is that after having named a number of words with the same syllable structure, naming a picture with the same structure involves the reuse of syllabic material, that will cause the picture to be named more quickly. When the structure of the primes and target picture do not match, no such reuse of syllable structure can occur and naming will not be speeded.

This is exactly what the researchers found. Taken together, these experiments indicate that the syllable is an important unit in Spanish language processing.

6.2.3 *Phoneme and syllable monitoring*

Measuring the priming effect that one word has on another is not the only method for testing the reality of the syllable in speech processing. Another method is known as monitoring. In this experimental paradigm, the subjects' task is to press a button as soon as a particular phoneme or sequence of phonemes is perceived. In Pallier, Sebastián-Gallés, Felguera, Christophe, and Mehler's study (1993), the subjects were asked to press one key if they heard a certain sound in a word and another key if they did not. If they were to monitor for /g/ for example, they heard words with and without /g/. The words containing /g/ appeared in either the third (*doG.ma*) or fourth position (*san.Gre*). Words with the phoneme in the fourth position were included as distractors. The relevant test words were those with the phoneme in the third position. Although the sequential position of the phoneme in the test words was identical, the syllable the phoneme belongs to was varied: in *seg.men.to,* it appears in the first syllable, and in *sa.gra.do* in the second.

The subjects were divided into two groups. One group heard a majority of target words containing the assigned phoneme in the coda of the syllable, as in *seg.men.to.* The majority of the target words the other group heard had the phoneme in the onset of the second syllable, as in *sa.gra.do*. If syllable structure is irrelevant, there should be no difference between the groups since most of the target phonemes appear in the third serial position. On the other hand, if syllable structure is a factor, the group that received most of the targets in the second syllable should become accustomed to finding it in that position, and respond more quickly to those words than to the words containing the target phoneme in the first syllable. The opposite effect would be expected for the group hearing the target phoneme in the coda of the first syllable. The outcome of the study indicates that subjects reacted more quickly to words having the target phoneme in the syllabic position they expected to find it in.

The subjects' reaction times ranged from about 600 to 750 milliseconds. In the psycholinguistic literature, such reaction times are thought to indicate that lexical access has occurred. That is, the subjects appear to have had enough time to actually look up the word in the mental lexicon, which could be interpreted to mean that syllabification takes place after the word is recognized and that the syllable was not used as a unit that aided initial recognition. In an experiment by Sebastián-Gallés, Dupoux, Segui, and Mehler (1992), syllabic effects disappeared in short reaction times. To speed reaction times, Pallier et

al. (1993) replicated the study with two modifications. First, the instructions given to the subjects emphasized speed. Second, the subjects pressed a key only when they detected the target phoneme; no key was pressed to indicate the lack of the target phoneme in a test item. Under these test conditions, reactions were faster, ranging from about 320 to 370 milliseconds. In spite of faster reactions, the influence of the syllable remained significant. This suggests that the subjects were able to detect whether the target appeared in the expected syllable position at a point in processing before the word has been completely recognized and accessed from memory.

Rather than having subjects monitor for individual phonemes, Bradley, Sánchez-Casas, and García-Albea (1993) had them respond if they perceived the test word to contain a specific sequence of phonemes that could represent a syllable or not. Test words and target sequences demonstrated phonological overlap as well:

Target Sequence	Test Word	Relationship
pa	*pa.lo.ma*	same initial syllable
pa	*pal.me.ra*	different initial syllable
pal	*pa.lo.ma*	different initial syllable
pal	*pal.me.ra*	same initial syllable

It should come as no surprise that the subjects recognized *pal* in *pal.me.ra* more quickly than in *pa.lo.ma*. In the same way, they recognized *pa* in *pa.lo.ma* faster than in *pal.me.ra*. The most logical explanation is that their perception was based on syllabic overlap and not on segmental overlap. Sebastián-Gallés et al. (1992) performed a similar experiment. They also obtained a syllabic effect, but only in an experiment with average reaction times of about 626 milliseconds. When the experimental design elicited much faster reactions (about 378 milliseconds), the syllable effect disappeared.

Evidence that Spanish speakers make use of syllables in language processing is plentiful, but not unanimous. Although a small minority of studies show non significant effects, the body of evidence is positive. It is important to note that this evidence comes from studies involving a number of different paradigms including visual word recognition, picture naming, and phoneme monitoring tasks. The variety of experimental paradigms lessens the possibility that the results are due to the particular experimental task rather than a general linguistic processing strategy.

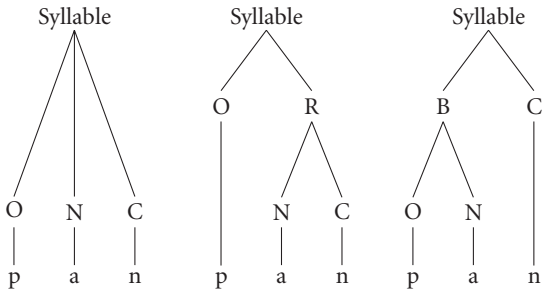

Figure 3. Three representations of syllabic constituents

6.2.4 *Internal syllable structure*

The evidence presented above suggests that Spanish speakers make reference to the beginnings and ends of syllables, but it says nothing about the internal structure of syllables. A number of structures have been proposed in the literature. As an illustration, I will present three major structures, bearing in mind that many more are possible when every nuance of formal representation is taken into consideration.

The first represents a flat structure. It is the most simple because it does not group any syllabic positions into constituents. This follows Saporta and Contreras (1962) and Núñez-Cedeño (1985) most closely. The second reflects Harris's (1983, 1989a) proposal which posits the existence of the rime (R) that comprises the nucleus (N) and coda (C). The third is the reverse of the second. However, to my knowledge, a syllable containing a body (B) composed of the onset (O) and nucleus has not been proposed for Spanish but has been for Korean (Derwing, Yon, & Cho 1993; Yoon & Derwing 2001). In sum, the body structure entails a close relationship between the onset and nucleus, while the rime structure implies that the nucleus and coda are more closely tied. The flat structure does not indicate any particular kinship between the nucleus and other syllabic elements. A number of methods have been employed to test the internal structure of the syllable. I review these below, although I forewarn the reader that the evidence is problematic and inconclusive.

One way to differentiate the three proposed syllable structures is to observe how speakers divide syllables. In English, blends tend to support a rime structure. A blend occurs when parts of two words are combined to form a new word. For example, *breakfast* + *lunch* = *brunch*. This show that the syllables [brɛk] and [lʌnʧ] group the nucleus and the coda together ([ɛk] and [ʌnʧ]) rather than the onset and the nucleus ([brɛ] and [lʌ]). The constituents

of the rime stay together in blends yielding *brunch* [brʌnʧ] rather than **brench* [brɛnʧ]. The blending evidence is consistent with other data on the English syllable (Treiman 1986).

As Bertinetto, Agonigi, Cioni, García Lecumberri, and González Parra (1999) have observed, blends like *brunch* are extremely hard to come by in Spanish. But, if one combines *portugués* and *español* what is the resulting form? *Portuñol* occurs 1,640 times in a search of the internet using Google, and *portañol* only 15 times. In *portuñol*, the syllable *-tu-* keeps its onset and nucleus together, while the less common *portañol* strips onsets from nuclei of the syllables *-t/u-* and *-p/a-*, leaving *-ta-*. *Portuñol* may be preferred for another reason that has nothing to do with syllable structure; *portuñol* involves the combination of whole, uncut syllables: *por.tu.~~gues.es.pa.~~ñol*.

Perhaps clearer cases of syllable constituency could be encountered in the formation of blends that involve breaking syllables. For instance, syllable breaks are found in both *inglañol* and *ingleñol*. The former appears 24 times on the internet and its variant *englañol* twice for a total of 26. *Ingleñol* on the other hand, appears 123 times, while its spelling variant *engleñol* is not found, for a total of 123. The crucial syllables here are *-gles* and *-pa-*, which are both broken between the onset and nucleus in *inglañol* and *englañol* (*gl/es, p/a*), suggesting that the rime is a unit. Nevertheless, *ingleñol* is slightly more common and it appears to involve the removal of the coda *-s* from *-gles*. Given its more frequent occurrence, can one posit the body as a significant syllabic constituent in Spanish? I think not because another possibility is that whole syllables are strung together (*in.gles.ñol*) and the unattested consonant cluster *-sñ-* is simplified to *-ñ-*. The fact that *inglañol* alternates with *ingleñol* in the same way *portuñol* and *portañol* are attested could also be interpreted to mean that neither the body nor the rime exist. According to the flat representation of the syllable there are no internal constituents, which could account for the fact that breaks occur both before and after the nucleus.

The purpose of this diatribe has been to demonstrate just how difficult it is to reach any sort of conclusions on the basis of a handful of forms. Bertinetto et al. (1999) also grappled with the problem of Spanish syllable structure and carried out two experiments. In one study, their subjects had to replace certain parts of a series of nonce words with new material. After a brief training session for each type of replacement (n=12), the actual test items were presented. Consider one of the tasks in which they were instructed to replace the word-initial CV- with another CV-. Training and testing were carried out in this fashion. First, they were shown the letter or letter combinations they were to use in the replacement (e.g., *BU*). They were then presented an auditory nonce word

(e.g., *len*). They performed the substitution and gave their responses orally. According to the training, the correct response would be *bun*.

Subjects were given new instructions and trained for each replacement position. Six of the replacement positions are directly relevant to the issue of Spanish syllable structure. The underlined CV units indicate the locus of replacement in each task:

Rime	Example	% Correct
CVC	Insert UL into *men* yielding *mul*	85.0
CVCCVC	Insert ES into *lambur* yielding *lesbur*	52.5
CVCCVC	Insert OR into *casfén* yielding *casfór*	70.8
Body	**Example**	
CVC	Insert BU into *len* yielding *bun*	84.9
CVCCVC	Insert TU into *discar* yielding *tuscar*	65.4
CVCCVC	Insert TI into *palsor* yielding *paltir*	53.3

The idea behind this experimental paradigm is that the errors should correspond to the syllable structure. If the rime is a syllabic constituent, fewer errors should be made when the replaced segments correspond to the rime. If the body is a significant unit, replacing segments corresponding to the body of the syllable should be easier, producing fewer errors. A statistical analysis reveals no significant difference in the number of errors produced on the rime versus the body replacements. These findings could be interpreted in favor of a flat syllable structure that consists of neither a body nor a rime.

In a second experiment, Bertinetto et al. (1999) employed a breaking task. The subjects heard two nonce words, each with a CVC structure. They were then shown two words that could result when the words are combined. The subjects chose between the possibilities. For instance, given *res* and *gan* they could choose either *ren* or *ran. Ren* is derived by stripping the coda from each input word while leaving the bodies intact. *Ran*, on the other hand, is expected if the onsets are separated and the rimes left intact. 55% of the responses respected the rime by not dividing its members; 32% demonstrated a preference for not splitting the body; 13% of the cases gave no answer. Analysis demonstrates that the tendency for preserving the rime over the body is statistically significant.

I would like to suggest that the outcome of this experiment is not as easily interpreted as it may first appear. More specifically, syllable and trigram frequency may have arisen as confounding factors. I compared each of the two pairs of possible responses in terms of their syllable frequency. For instance,

according to Álvarez, Carreiras, and De Vega (2000), the syllable *ren* has a frequency of 42, while the syllable frequency of *ran* is 73. In 8 of the 12 pairs of test responses the rime response is composed of a more frequently encountered syllable. I also consulted the Alameda and Cuetos (1995) frequency dictionary. There, *ren* was found to occur as a word-final trigram in 93 words and *ran* in 903. In 8 of the 12 pairs, the response that respected the rime appears in more Spanish words. In short, the subjects' preferences may not have had much to do with how syllables are broken, but with which of the two responses looked more like other Spanish words. Only further investigation will be able to tease these two possible influences apart.

On the whole, it is safe to conclude that the body of evidence points to the syllable as a unit that plays a role in processing the Spanish language. However, the evidence adduced thus far does not permit any conclusions to be drawn regarding what the internal constituency of the Spanish syllable may be.

6.3 Stress

The issue of Spanish stress assignment has occupied a good number of researchers working within the generative paradigm (e.g., Bailey 1995; Den Os & Kager 1986; Harris 1969, 1983, 1989a, 1992, 1995; Hooper & Terrell 1976; Lipski 1997; Roca 1988, 1990, 1991, 1997; Saltarelli 1997; Whitley 1976). As discussed in Chapter One, such analyses may relate to the linguistic competence of an ideal speaker-hearer, but do not necessarily have relevance to actual mental processes, though they are often described in performance-oriented terms. This section will center on analyses that are concerned with performance aspects of Spanish stress assignment.

6.3.1 *Words ending in* -n

The generalization in Spanish is that non-verbs ending in *-n* have final stress. However, Aske (1990) noticed that while about 90% of commonly occurring non-verbs ending in *-in, -on, -an,* and *-un* were stressed on the final syllable, only about 62% of non-verbs ending in *-en* had final stress. Could this subpattern play a role in linguistic cognition, or do speakers follow a rule to the effect that *-n* final non-verbs receive final stress? Aske devised 6 final *-en* nonce words and 6 nonce words ending in an *n* preceded by another vowel. He then embedded them in sentences in which they appeared as non-verbs and asked Spanish speakers to read the sentences while he noted where the subjects stressed the test words. He noted that 96.8% of the responses to words ending in an *n* preceded by *a, i, o,* and *u* were given final stress, while only 55.6% of the responses

to *-en* words received final stress. This led Aske to hypothesize that stress is not based on internalized generalizations in the form of a rule, rather, it is determined on the basis of the stress patterns that are found in the words of the language. That is, stress placement is analogical.

Face (2004) augmented Aske's production evidence with evidence from perception. He devised 10 nonce words ending in *-an, -on*, and *-en.* These words were then produced using a speech synthesizer in such a way that the fundamental frequency was held constant, the duration of each syllable was identical, and intensity did not vary from syllable to syllable. In this way, acoustic cues about which syllable was stressed were eliminated. These synthesized words were presented to subjects who determined which syllable they perceived to carry the stress.

Unlike Aske, who gave all nonce words as non-verbs. Face's study presented the words in isolation, which opens the possibility that final *-en* and *-an* words could be construed as verbal forms; penultimate stress is the norm in those cases. The results corroborate Aske's findings somewhat in that *-en* words were given final stress less often than *-an* and *-on* words (59%, 68%, and 75% respectively). However, the fact that either a verbal or non-verbal interpretation was possible makes close comparisons impossible.

6.3.2 *Stress is assigned analogically*

Both Face and Aske assert that the subjects in their studies used, not an internalized rule of stress placement, but analogy to determine the accentuation of the *n*-final nonce words they were presented. This premise forms the basis of a number of simulations I performed (Eddington 2000b). The simulations were carried out using the analogical algorithm that was presented in 5.3 (Analogical Modeling of Language; Skousen 1989, 1992, 1995). The database of exemplars that serves as an approximation of the contents of the mental lexicon consisted of the 4,970 most frequent Spanish words. These included inflected forms, uninflected forms, and verb-plus-clitic combinations. Each word was converted into a series of variables that included the phonemes of the three final syllables, as well as variables indicating the person and tense of the verbal forms. Each entry was also marked to indicate which syllable was stressed. In the simulation, each word was removed from the database and its stress was determined analogically according to the stress placement of similar words in the database.

The simulation predicted the correct accentuation in 94% of the cases, which indicates that an analogical model of accentuation is able to recognize regular stress patterns without computing a global generalization about the data. In fact, correct assignment of words stressed on the antepenultimate

syllable reached 40%. While this may not appear particularly impressive at first glance, it is important to note that this was achieved without incorporating any diacritic marking into the antepenultimately stressed words. This contrasts with formal accounts of Spanish stress placement which must incorporate some sort of diacritic mark on all antepenultimately stressed words in order to correctly assign stress to them. The analysis of the errors produced during the simulation is also telling because 80% of them involved regularizing irregular stress.[46] In his nonce study, Face (2000) also found that most errors in the perception of stress placement involve regularization.

Other empirical data support stress assignment by analogy. For example, in a study of stress acquisition, Hochberg (1988) elicited words with different stress patterns from preschoolers. Children named various objects shown to them in a picture book. In another task, they were asked to repeat nonce words that they heard. The types of errors the children made on both tasks were tabulated. The children made fewer errors on penultimate stress, followed by final stress, while the highest error rates occurred on antepenultimately stressed words. This exact hierarchy of difficulty was evident in the errors produced by the analogical simulation (Eddington 2000b).

Hochberg also noticed that the error rate on regularly stressed words remained virtually unchanged for all of her subjects ages three to five. However, the five-year-olds produced significantly fewer errors on irregular items than did the four-year-olds. The difference between the mental lexicon of a younger versus an older child is arguably the size of the vocabulary. In order to imitate this difference I ran one simulation with analogs drawn from only the most frequent half of the database items and another using all 4,970 items. Nevertheless, stress was predicted for all 4,970 original database items. When the error rates of the two simulations are compared, the percentage of errors made on regularly stressed words did not differ significantly, but significantly fewer errors were made when the entire database was available. This again corresponds to Hochberg's developmental data.

Face (2003) devised an experiment to further test the notion that accentuation is governed by analogy. He chose ten consonant-final words such as *hospital* with final stress and removed the final consonant yielding [ospita]. These modified words were synthesized so that phonetic cues such as intensity, duration, and frequency were equalized across all syllables. These same words with no phonetic stress cues were also synthesized so that the final consonant was replaced with /s/, yielding [ospitas]. The resulting 20 test words were then presented to Spanish speaking subjects, who indicated which syllable they perceived to receive primary stress.

If the subjects applied a rule to the effect that vowel-final words and those ending in *-s* are stressed on the penultimate syllable, one would expect the test words to be given penultimate stress. On the other hand, analogy would predict that not only will the influence of other vowel-final and *s*-final words make their influence felt but test words [ospita] and [ospitas], for example, should also feel a strong pull toward final stress based on the high degree of similarity they bear to the extant word [ospitál]. The outcome of Face's experiment clearly demonstrates the influence of analogy; 59% of the vowel-final, and 37% of the *s*-final words were perceived as having final stress. This contrasts starkly with the application of a rule, which would predict penultimate stress in most instances. Penultimate stress was given in only 29% of the vowel-final cases and 45% of the *s*-final cases. This serves as further evidence that analogy, not global generalization underlies stress placement.

6.3.3 *Quantity sensitivity and stress placement*

In many languages of the world there is a difference between light syllables which end in vowels and heavy syllables which end in consonants as far as stress assignment is concerned. In general, heavy syllables attract stress. Whether or not Spanish is a quantity sensitive language is much debated. This issue is not a purely formal topic because syllable weight is directly observable. For this reason, quantity sensitivity has not escaped the attention of experimentalists. Arguing from within a formal framework, Harris (1983, 1992), Núñez-Cedeño and Morales-Front (1999), and Lipski (1997) assert that syllable weight is an important factor. Roca (1990), on the other hand, feels that Spanish is not a quantity sensitive language.

One evidence that quantity sensitivity is relevant in Spanish is cited by Harris (1983, 1992). He claims that Spanish has a constraint based on syllable weight that disallows words with a heavy penultimate syllable from being stressed on the antepenultimate syllable. In other words, there are no native Spanish words of the type **teléfosno* or **típisco*. Of course, foreign words such as *bádminton* and *Wáshington* are exceptions to this generalization. Núñez-Cedeño and Morales-Front (1999) analyze /rr/ as a sequence of two flaps that is divided between two syllables. In this way, *cachorro* would be syllabified *ca.chor.ro* making the penultimate syllable heavy and explaining why words such as **cáchorro* are nearly unattested in Spanish. In my view, one thing that weakens their assertion about /rr/ is that in many Spanish dialects /rr/ is not a trill, but a sibilant or fricative (Lipski 1994), which makes it a phone that is difficult to divide neatly into two parts. Even in dialects where the trill exists

it would be surprising if native speakers who were asked to divide words into syllables divided *cachorro* as *ca.chor.ro.*

The question to be asked from a performance perspective is whether this lexicon-internal constraint has psychological validity. To this end, Alvord (2003) asked a group of native Spanish speakers from various countries to judge a set of nonce words in terms of whether they could be Spanish words or not. Some words in the questionnaire violated Spanish phonotactic patterns (e.g., *ponsinmrú*). Others were possible words, such as *distropa,* that follow legitimate phonotactic patterns. The real test words in the experiment were antepenultimately stressed words containing heavy penultimate syllables, and penultimate syllables followed by /rr/ (e.g., *ránlinta, tíldorra*).

It should come as no surprise that 96% of the impossible words were rejected and 93% of the possible words were judged acceptable. However, 97% of the nonce words containing penultimate syllables followed by /rr/ were perceived as possible Spanish words and 94% of the words containing a heavy penultimate syllable were judged as acceptabe. These high levels of acceptability were given in spite of the fact that all of these crucial test words were stressed on the antepenultimate syllable. Spanish speakers appear to have little problem accepting words that violate the constraints proposed by Harris, and Núñez-Cedeño and Morales-Front. In fact, Spanish speakers have no problem accepting words with preantepenultimate stress either in spite of the fact that these violate supposed constraints (Face & Alvord 2003). This not only casts serious doubt on the psychological reality of the constraint, but also weakens the claim that syllable weight plays a role in Spanish accentuation.

Face (2000, 2002) and Bárkányi (2002) also examined the issue of syllable weight. In Bárkányi's experiment, the subjects' task was to assign stress to orthographic representations of invented words. She found that not every subject assigned final stress to every nonce word ending in a closed syllable, nor did they all assign penultimate stress to every word ending in a vowel. For this reason, she concludes that syllable weight is not an active factor in Spanish stress placement.

In Face (2000), subjects heard recorded nonce words that had been manipulated so that each syllable nucleus was of identical length and intensity. The subjects' task was to determine where they perceived the stress. Face submitted his results to statistical analysis and found that heavy syllables attracted stress. However, there was a flaw in the study in that the syllable nuclei were all of the same length, while the duration of each syllable was not. In order to remedy this undesirable situation Face (2002) performed another similar study. This time the nonce words were manipulated so that each syllable, not just each syllable

nucleus, was of equal length and intensity. Under these conditions, the weight of the final syllable was influential in the subjects' perception of stress placement, but the weight of other the syllables was not. This lead him to conclude that Spanish is not a quantity sensitive language.

Waltermire (2004) followed up Face's experiments by presenting the same 60 nonce words from Face (2000) to 41 native Spanish speakers. The speakers were asked to indicate where they felt the words should be stressed. However, in Waltermire's study the words were presented in written rather than auditory form. Nevertheless, the outcome obtained by Waltermire mirrors that of Face (2000) quite closely; syllable weight was a significant factor in the subjects' choice of which syllable was accentuated. The weight of the final syllable was extremely influential, while the weight of the penultimate and antepenultimate syllable exerted only a modicum of influence.

Using the nonce words from Face (2002), I performed a study identical to Waltermire's (Eddington 2004a). These 40 nonce words were given in written form to 38 university students, all of whom were natives of Spain.[47] With the exception of three who did not give their age, the participants were between the ages of 17 and 26. The questionnaire revealed the percentage of responses in which each possible accentuation was given. For instance, the item *bolnala* was given antepenultimate stress by 16% of the subjects, penultimate by 68%, and final by 16%. The results from the 40 test items from this study were combined with the 60 from Waltermire's in order to correlate the findings with a number of computer simulations which made predictions about the stress placement in terms of percentages.

The purpose of running analogical simulations with the nonce words from the two questionnaires was to determine what factors most influence the subjects' choice of stressed syllable. To this end, the stress placement of the 100 nonce words devised by Face (2000, 2002) was determined by analogy using seven sets of variables. The word *bolnala* will serve to highlight the difference between each of seven sets of variables:

1. Syllable weights alone (*bolnala* = closed, open, open)
2. Syllable weights and final phoneme (*bolnala* = closed, open, open, a)
3. CV tier alone (*bolnala* = C,V,C,C,V,0,C,V,0)
4. CV tier and final phoneme (*bolnala* = C,V,C,C,V,0,C,V,0,a)
5. Phonemic representation alone (*bolnala* = b,o,l,n,a,0,l,a,0)
6. Phonemic representation and syllable weights (*bolnala* = b,o,l,n,a,0,l,a,0,closed, open, open)

7. Phonemic representation and CV elements
(*bolnala* = b,o,l,n,a,0,l,a,0,C,V,C,C,V,0,C,V,0).

All 4,970 database items were available as possible analogs for the nonce items. The results of the simulations were correlated with the data from Waltermire's study, and my own study that used Face's 2002 nonce words (see Table 13). These were used to the exclusion of the data gathered by Bárkányi and Face (2000), since the latter do not provide their subjects' responses on each individual test word, making it impossible to calculate correlations between their experimental findings and the analogical simulations.

The motivation for testing different sets of variables was the following. Quantity sensitivity refers to the presence or absence of a syllable-final consonant. However, there are several ways in which heavy and light syllables could be processed. One is that the string of phonemes composing a word is parsed into syllables. The phonemic content is then completely abstracted away and calculation of stress placement is carried out with reference to syllable weights alone as in (1). Of course, given the important status of the final consonant it may be that the final phoneme needs to be considered as well, as in (2). Another possibility is represented in variable set (3). There, stress placement does not make reference to syllable weights per se, but only to elements on the CV tier (Clements & Keyser 1983). Set (4) includes both the elements on the CV tier and the word-final phoneme as well.

Another possibility that needs to be explored is that the actual phonemic content of the word is what is used in determining stress placement. That is, syllable weights and CV tier elements are derived directly from the phonemic makeup of a given word. Thus, it may be that the phonemic make up of the words, not abstractions such as syllable weights and CV elements, is responsible for the ability of these abstract representations to predict accentuation. That is,

Table 13. Correlations between nonce words and questionnaires with different variable sets

Variable sets	Correlations
(1) Syllable Weights Alone	.764
(2) Syllable Weights Plus Final Phoneme	.897
(3) CV Tier Alone	.503
(4) CV Tier Plus Final Phoneme	.834
(5) Phonemic Representation	.648
(6) Phonemic Representation Plus Syllable Weights	.695
(7) Phonemic Representation Plus CV Tier Elements	.649

any effect of these abstract representations may be merely epiphenomenal. In order to test this hypothesis variable set (5) was included. Variable sets (6) and (7) were designed to determine whether syllable weights or CV tier elements work in conjunction with the phonemic content of the word in assigning stress.

Syllable weight, when considered alone (1), results in higher correlations with the subjects' responses than both phonemic (5) and CV representation (3). When the final phoneme is included along with the syllable weights (2) the highest correlation is found. Considering the phonemic information along with weights (6) or CV elements (7) does not improve correlations. If stress placement is carried out analogically, this suggests that the subjects used an abstract calculation of syllable weight apart from the sequence of phonemes and CV elements that composed each test word. Additionally, some phonemic information such as the final segment also appears to have played a role.

6.4 Conclusions

The outcomes of various psycholinguistic studies indicate that the syllable is a unit that plays a role in processing the Spanish language. The evidence regarding what the internal constituency of the Spanish syllable may be is much more ambiguous and requires further study before any conclusions are warranted. In terms of linguistic performance, stress placement of nonce words has been shown to be tightly related to the patterns of stress in existing Spanish words. In other words, accentuation is calculated on the basis of analogy, not rules or constraints. Is Spanish a quantity sensitive language? That is, is the weight of the syllables a significant factor in stress placement? The production experiments cited above suggest that an abstract notion of syllable weight, separate from the presence or absence of particular phonemes in the coda, in fluences the subjects' accentuation preferences. The weight of the final syllable was particularly influential in the experiments, while the weights of the penultimate and antepenultimate syllables (Waltermire 2004) exerted a much smaller influence.

Chapter 7

Morphology in word recognition

7. Introduction

The experimental study of morphology has focused principally on the role of morphology in visual word recognition, which explains the title of this chapter. The major experimental paradigm utilized in studies of word recognition is the lexical decision task (LDT) introduced in Section 6.2.2. The task involves deciding if a string of letters presented on the computer screen is an extant word or not. Subjects press one key if the string is recognized as a legitimate word, and another key if it is not. In this way, reaction times may be measured. Of course, some of the stimuli are real words and others are not. Relationships between a pair of words are determined if reactions times to a target word, such as *Madrid*, are affected by the prior presentation of a prime word such as *madrileño* 'inhabitant of Madrid'. If reaction time is speeded, the effect is said to be facilitatory; if it is slowed the priming is said to be inhibitory. A control condition is included in which there is no relationship between the prime and target such as *Madrid/gustar* 'to please'. The average reaction time in the control condition is subtracted from the average reaction time in the test condition in order to calculate whether there has been any inhibition or facilitation. LDTs allow experimenters to measure the effect of orthographic similarity between words such as *donde* and *donar* 'where/donate,' the effect of semantically related words such as *médico/hospital* 'doctor/hospital', and the morphological relationship between words such as *Madrid* and *madrileño*.

7.1 Orthographic and semantic priming

A number of studies in several different languages have shown that the presentation of a prime such as *bribe* will inhibit the reaction time to a target such as *tribe* that shares orthographic features (Drews & Zwisterlood 1995; Forster, Davis, Schnoknecht, & Carter 1987; Fowler, Wolford, Slade, & Tassinary 1981; Grainger, Cole, & Segui 1991; Henderson, Wallis, & Knight 1984; Stolz & Feldman 1995). Other studies yielded facilitatory effects (Forster 1987; Hillinger

1980; Napps & Fowler 1987; Sereno 1991), and still others demonstrated no effects at all (García Albea, Sánchez-Casas, & Igoa 1998; Martin & Jensen 1988; Murrell & Morton 1974; Slowiaczek & Pisoni 1986).

Purely semantic priming has also been demonstrated. For example, Martin and Jensen (1988) showed that synonyms such as *swap* and *trade* demonstrate a priming effect. Semantic priming between word-pairs such as *sea/ocean*, and *mouse/cheese* is also attested (Balota 1983; Dannenbring & Briand 1982; Henderson, Wallis, & Knight 1984; Napps 1989). A semantically related prime can facilitate a target even in the masked condition when the subjects are not aware of having seen the prime at all (Balota 1983; Fowler et al. 1981; Marcel 1983). One thing that is clear in these studies is that semantic priming is fairly short lived. With longer lags between the presentation of prime and target, or when many intervening test items appear between the prime and target, the effect of the prime on the target disappears (Becker 1980; Bentin & Feldman 1990; Fischler 1977a, b; Henderson et al. 1984; Monsell 1985; Napps 1989).

7.2 Morphological priming

Orthographic and semantic priming are interesting effects, but the goal of this chapter is to discuss the sort of morphological priming that occurs between words such as *hijo* and *hija* 'son/daughter' that share the morpheme *hij-*. Presenting orthographic and semantic priming first raises a relevant question. Morphemes are units that share both orthographic/phonemic and semantic elements. Most would agree that *enseñar* and *enseñanza* 'to teach/teaching' share a morpheme, while *voy* and *fui* 'I go/I went' do not.[48] The argument is that the former pair shares both phonemic shape and meaning, while the latter only have common semantic characteristics that renders them incapable of having a morpheme in common. Therefore, the crucial question is this: Are morphological relationships merely the confluence of orthographic/phonemic and semantic traits? If this is the case, morphology is not a separate realm of psycholinguistic processing as Seidenberg (1987) asserts. Demonstrating that morphological relationships are different from orthographic/phonemic and semantic relationships would lay a firm groundwork for the reality of morphology as an independent linguistic dimension.

As already discussed, semantic priming produces very short lived facilitation. Orthographic priming has been shown to inhibit, facilitate, or not influence recognition at all. It is difficult to imagine how the additive effects of these priming effects could result in the sort of robust, long-lived facilitation that has been measured when one morphemic relative primes another (Feldman 1994;

Napps 1989). Morphological priming differs from semantic and orthographic priming in other ways as well. For example, Marslen-Wilson, Komisarjevsky, Waksler, and Older (1994) employed a cross-modal paradigm in which the subjects heard the primes, but responded to visual targets. In the experiment, the degree of phonemic overlap was varied. Phonologically transparent pairs such as *happy/happiness* and *friend/friendly* have no change in the phonology of the base word. Other pairs such as *elude/elusive* and *serene/serenity* do no overlap completely in their phonemic structure. In spite of these differences, the degree of phonemic overlap did not affect the degree of priming obtained (however, see Kempley & Morton 1982 for some counter-evidence).

In studies of visual LDT, regular pairs such as *manage/manages* and *clear/clearly* have been compared with spelling change pairs such as *creep/crept* and *clear/clarify* (Fowler, Napps, Feldman 1985; Napps 1989). The varying degree of orthographic overlap did not produce varying degrees of priming, suggesting that the morphemic effects cannot be due to orthographic effects. This point is forcefully driven home in a study carried out in Serbo-Croatian, in which both the Roman and Cyrillic alphabets were used to represent the language orthographically (Feldman & Moskovljevic 1987). Primes were presented in one alphabet and targets in the other. Under these conditions, significant morphological priming occurred. One may argue that the ability of *manage* to prime *management* is due to the fact that the subjects remember seeing the string *manage*. However, in the Serbo-Croatian study, the effect of episodic memory was eliminated, yet priming effects still emerged. The test words must have been accessed from the mental lexicon in order for the priming effects to be manifest.

Nevertheless, the ability of episodic memory to influence priming studies has been demonstrated in other studies. Stanners, Neiser, Hernon, and Hall (1979) found that pairs such as *lift/lifting* primed better than irregular inflections such as *shake/shook*. These results were obtained with an average lag of 10 items between the prime and the target. However, Fowler, Napps, and Feldman (1985) suggest that the experiment by Stanners et al. may have been influenced by episodic sources. That is, subjects may have remembered seeing a word or its relative on an earlier trial, and regularly inflected words are more orthographically similar than inflected words with stem allomorphy.

In order to eliminate any possible episodic influence Fowler et al. increased the average number of items that were presented between the between the prime and target from ten (as in Stanners et al.) to 48. At an average of ten, nonwords such as *flup* primed for inflected nonwords such as *flupping*. This effect for nonwords must be due to episodic sources and not to morphemic re-

latedness. However, at an average of 48 intervening items, no priming occurred between nonwords. At the same time, priming did occur for inflectionally related words; *signing* primed for *sign*, *sang* for *sing*. Fowler et al. (1985) concluded that once episodic sources are controlled for, priming is seen to occur between related words with or without stem allomorphy because all words that have a common morpheme are organized together in the lexicon.

In short, the degree of orthographic overlap does not affect morphological priming. Can the same be said of degree of semantic overlap? Feldman and Stotko (cited in Feldman 1992) compared closely related words such as *create/creation* with less semantically similar words such as *create/creature*. In spite of the differences in semantic relatedness, words of both types primed equally well. Rastle, Davis, Marslen-Wilson, and Tyler (2000) compared the priming effects of three groups of words, all of which are orthographically similar: 1) words that are both semantically and morphemically related, such as *adapter/adaptable*; 2) words that have a historical morphological relationship, but are not synchronically related in terms of meaning *part/apartment*; 3) words that are semantically but not morphemically related, such as *scream/screech*. Each group of words was also matched on the average degree of orthographic overlap.

The results indicate that pairs of primes and targets such as *adapter/adaptable* and *part/apartment* primed equally well, and more fully than *scream/screech*. This suggests that it is the morphemic relationship that is responsible for the priming. More importantly, pairs such as *adapter/adaptable* and *scream/screech* are all semantically related, yet morphemically related pairs such as *adapter/adaptable* prime much more fully than do pairs such as *scream/screech*.

One interpretation of these results is that they constitute firm proof that morphological relationships may not be reduced to the sum of orthographic and semantic relationships. It would be easy to imagine that the difference is due to the fact that pairs such as *scream/screech* simply bear much less semantic resemblance to each other than *adapter/adaptable*. However, the authors used two different measures of semantic similarity in order to closely match their stimuli.

In keeping with the theme of degree of semantic relatedness, it is fairly uncontroversial that inflectionally related words are more closely related semantically than derivationally related words (Bybee 1985). Fowler et al. (1985) found no differences in degree of priming between inflectionally related words such as *manage/manages*, and between derivationally related words such as *manages/management*. In contrast, there is evidence in Italian and Serbian that sug-

gests inflectionally related words exhibit more robust priming than derivationally related words (Feldman 1994; Laudanna, Badecker, & Caramazza 1992). Additionally, in a cross-modal task in English (Marslen-Wilson et al. 1994) priming only occurred when the relationship between prime and target was semantically transparent, as in the case of *friendly/friend* and *govern/government*. Semantically opaque pairs such as *depart/department* and *author/authority* did not prime.

There are a number of ways to interpret these contradictory findings. One is to assume that morphological relationships are not completely separate from semantic relationships; varying the semantic component can, under certain conditions, influence the morphological relationships. Another possibility is that the results are due to differences in experimental paradigm and/or differences in the way one language deals with morphology compared to another (Marslen-Wilson et al. 1994).

7.3 Morphological effects in Spanish

Having set the stage in general terms, I now turn to evidence from Spanish. Domínguez, Cuetos, and Segui (2001) performed a LDT in which prime and target pairs were either semantically related such as *saña/odio* 'viciousness/hate', or morphologically related such as, *loco/loca* 'crazy man/crazy woman'. There was a robust priming effect for the morphologically related words, but only a small effect for the semantically related pairs. What this demonstrates is that morphological priming is not merely semantic priming. However, it does not take into consideration the fact that *loco/loca* are orthographically as well as semantically related.

In a similar experiment, Domínguez, Cuetos, and Segui (2000) contrasted morphologically related pairs (*loco/loca*) with purely orthographically related pairs such as *foco/foca* 'headlight/seal'. It should not be surprising that the priming was much more robust for the morphologically related words than for the orthographically related words. Similar stimuli were used by Barber, Domínguez, and de Vega (2001) who measured changes in voltage taken from the scalp (ERP) while the subjects were involved in a LDT. These measurements showed marked differences in changes in voltage when morphological versus orthographic words were responded to.

Taken together, these experiments serve to show that morphological priming is not merely the result of orthographic similarity. However, the morphologically related words used are also semantically related; therefore, these studies do not control for possible semantic influence on the outcomes.

Table 14. Results of Domínguez et al. (2002)

Relationship	Examples	Glosses	SOA in milliseconds		
			32	64	250
Morphological	*loco/loca*	'madman/madwoman'	+29*	+68*	+53*
Orthographic	*rato/rata*	'while/rat'	+27*	+28*	–12
Semantic	*luto/pena*	'mourning/sadness'	–3	+14	+45*

* significant at $p < .05$

García-Albea et al. (1998) compared the priming effect of inflectionally related (*hijo/hija*) and derivationally related pairs (*rama/ramo* 'branch/bouquet'). All of these test words are semantically and orthographically similar. However, the inflectionally related words are arguably more semantically close than the derivational pairs, nevertheless, both yielded equally strong priming effects. Therefore, the locus of priming is arguably morphological and not semantic/orthographic in nature.

Domínguez, Segui, and Cuetos (2002) compared morphological, orthographic, and semantic priming at three SOAs (stimulus onset asynchrony: the amount of time that elapses between the presentation of the prime and target). They calculated facilitation by subtracting average reaction time to unrelated pairs such as *loco*/*cera* 'madman/wax'. Their findings are summarized in Table 14.

At 32 ms SOA, orthographic and morphological priming are identical and it is impossible to determine if they are not actually one and the same effect. By 64 ms, it is clear that morphological priming is not the same as orthographic. Although semantic priming is not significant at this SOA, the question is whether the 28 ms orthographic facilitation combined with the 14 ms semantic facilitation (totaling 42 ms) is not responsible for what appears to be morphological facilitation. The authors compared the morphological facilitation of 68 ms with the combined orthographic and semantic facilitation of 42 ms and found them to be statistically equivalent, suggesting that at 64 ms SOA, morphological priming could indeed be conceived of as a mere combination of the other two types of priming.

There is an alternative explanation as well. Morphological priming is significant at all SOAs. However, at no SOA are both semantic and orthographic priming significant at the same time. This is especially apparent at the 250 ms SOA, where the effect of orthographic similarity tends toward inhibition while semantic similarity demonstrates strong facilitation. According to this view, orthographic and semantic facilitation do not add up to the robust morphological facilitation they are thought to underlie. The authors remained neutral as

to whether these results nullify the 'semantic + orthographic = morphological hypothesis' or not.

One difficulty with many of the above-cited experiments is that they successfully demonstrate that morphological effects are not the same as orthographic effects, yet they do not control for possible semantic effects of the stimuli that are morphologically related. Allen and Badecker's (1999) experiments are important in that they controlled for orthographic overlap and eliminated any effect semantic overlap could have. This was done by comparing stimuli with orthographically identical stems such as *mor-ía/mor-os* 's/he died/Moors' and *col-ar/col-as* 'to strain/tails'. Such pairs are called stem homographs (Laudanna, Badecker, & Caramazza 1992) because their morphological stems are orthographically identical, yet they have no semantic characteristics in common. If orthographic overlap can be eliminated as a factor in stem homograph priming, any remaining priming effect must be purely morphological. Allen and Badecker did this by including stimuli that overlapped orthographically, but not morphologically. For example, *col-ar* and *colm-ar* 'to top off' both begin with /kol/ yet the initial morphemes differ (*col-* versus *colm-*). Of course, hyphens indicating morphological breaks were not seen by the test subjects.

Their first experiment included three test conditions:

Relationship	Prime	Target	Glosses
Stem homograph	*mor-ía*	*mor-os*	's/he died/Moors'
Orthographic	*moral*	*mor-os*	'moral, mulberry/Moors'
Unrelated	*sill-a*	*mor-os*	'chair/Moors'

The unrelated condition served as a control condition from which inhibition or facilitation could be calculated for the other two conditions. Orthographically similar items were recognized 63 ms more slowly, and with 0.6% more errors than the unrelated words. The stem homographs, on the other hand, took an average of 124 ms longer to recognize and produced 7.8% more errors. The differences between the two conditions are significant. In both experimental conditions, the amount of orthographic overlap (three letters) and semantic overlap (= 0) is identical.

Allen and Badecker explain the difference by assuming that there are two levels of representation. At the 'form level' words are parsed into constituents. Those constituents then activate elements at the 'morphosyntactic/morphosemantic level.' For instance, the prime *mor-ía* is recognized by parsing it into parts. The morpheme MOR- is thus activated. Therefore, when the target *mor-os* is presented, the two form-level identical morphemes, MOR-

of *mor-os* and MOR- of *mor-ía,* compete with each other and hamper quick recognition. The same competition does not arise between the morphemes MOR- of *mor-os* and MORAL of *moral.*

In a second experiment, they tested the ability of allomorphic variants of the stem morpheme to prime. If a morphosyntactic/morphosemantic level truly plays a part in word recognition, all of the allomorphic variants of a verb such as *mor-ir* (e.g., *mor-ía, mor-imos, muer-e, mur-ió* 's/he was dying, we die, s/he dies, s/he died') should be subsumed under the abstract morpheme MOR-. The question here is whether there will be competition between MOR- of *mor-os* and MOR- of an allomorph of the verb *mor-ir* that is not orthographically identical to MOR- of *mor-os.* The following are sample stimuli:

Relationship	Prime	Target	Glosses
Stem homograph	*muer-e*	*mor-os*	s/he dies/Moors
Orthographic	*mir-an*	*mor-os*	they look/Moors
Unrelated	*sill-a*	*mor-os*	chair/Moors

Under these conditions, an insignificant degree of orthographic inhibition was found, and 1.2% more errors were made on orthographic relatives than on unrelated words. However, stem homographs required 116 ms more time to recognize than the unrelated words and induced 5.8% more errors. Purely orthographic overlap can by no means explain the differences. Instead, the allomorphic variant *muer-e* must activate the abstract morpheme MOR- that corresponds to the form level words *muer-e, mor-ir, mor-ía* etc. It is this MOR- that competes with the MOR- of *mor-os* and results in slower reaction times. Appeals to form-level similarities cannot account for these findings.

One could counter that the diphthongizing alternation evident in pairs such as *muer-e, mor-ir* is marginally productive in Spanish (Eddington 1996) and could have influenced the results. Nevertheless, the test stimuli included many different allomorphic variants such as the stem homograph prime/target pairs *plazc-a/plac-a* 's/he pleases/plaque' and *quep-a/cab-o* 's/he fits/cape'. These alternations are not productive by any stretch of the imagination.

If a morpheme is the confluence of semantic and orthographic/phonological characteristics, then one cannot speak of a separate level of morphological processing. Most studies find morphological priming while controlling for either semantic or orthographic/phonological influences, but not for both at the same time. Allen and Badecker, on the other hand compare words that have no semantic overlap and where orthographic overlap is controlled for. *Mor-ía, moral,* and *mor-os* all share the first three graphemes. However, *mor-ía* primes

mor-os while *moral* does not. This must be because what is being compared is not the overall orthographic similarity of the words, but the similarity of the morphological stem. *MOR-ía* overlaps the stem of *MOR-os* but *MORAL* does not. Therefore, it appears that there is a separate morphemic level of representation and that this level is not explainable in terms of the additive effects of orthographic and semantic similarity.

7.4 Morphology as associations between lexical items

The evidence presented thus far demonstrates that the additive effects of orthographic and semantic similarity do not equal the strong morphological effects encountered. However, one possibility that has not been discussed in the literature is that, while orthographic and semantic influences by themselves are not particularly strong or long-lived, when they are combined the dynamics of the relationship changes dramatically because both types of similarities reinforce each other. A parallel can be drawn with chemistry. At room temperature, oxygen and hydrogen are gases, yet when joined the two gases do not merely combine to form a different kind of gas, but a liquid. In the case of orthographic and semantic relationships, the different state they create could be conceived of as morphology.

There are a number of models that hold that all words have separate representation in the mind. However, according to these models, words are not stored as lone items, but as entities with complex networks of interconnections between each other (Bybee 1985, 1988, 1995; Drews & Zwisterlood 1995; Fowler et al. 1985; Grainger, Cole, & Segui 1991; Lukatela, Carello, & Turvey 1987). In particular, words belonging to families of morphological relatives are thought to be stored and linked to each other. It is this sort of lexical organization that allows the access of one member of the family to activate the other members (Feldman 1992; Lukatela, Gligorijevic, Kostic, & Turvey 1980; Segui & Zubizarreta 1985). Presumably, two of the most important types of links between stored items are semantic and orthographic/phonological.

This concept contrasts with the models presented by others (e.g., Stanners et al. 1979; Marslen-Wilson et al. 1994) that explain priming effects as due to access to the same lexical entry (e.g., the stem), or who explain it in terms of overlapping representations (Rastle et al. 2000) or obligatory parsing into morphemic constituents (Taft & Forster 1975). One bit of evidence against parsing into morphemes is that monomorphemic words do not take more time to recognize than words containing several morphemes (Manelis & Tharp 1977). Perhaps it was the difficulty of reconciling storage versus parsing that led other

scholars to investigate the possibility that both morphemic parsing and full word storage operate during linguistic processing (Baayen & Schreuder 1999; Caramazza, Laudanna, & Romani 1988).

As far as the Spanish data are concerned, Domínguez et al. (2002) remain agnostic as to whether the priming effects found between *loco/loca* reflect access to the common morpheme *loc-* or merely demonstrate how access to one word results in the activation of other words to which the accessed word is closely connected. In any event, this debate is not likely to be resolved in the foreseeable future. Therefore, what I would like to discuss is how a model of whole-word storage with orthographic and semantic links can explain the outcome of Allen and Badecker's study (1999). Domínguez et al. (2000:389) suggest that Allen and Badecker's experiments admit an interpretation in terms of organization among lexical items, but they stop short of detailing how this is possible.

My goal is to demonstrate how this is possible. I follow Bybee's (1988) framework quite closely. Consider Figure 4. The figure centers on three morphological families on whose members, *moros, moría*, and *moral,* Allen and Badecker (1999) centered their experiments. Elements contained within the same ellipse are all interconnected based on phonological, orthographic, and semantic similarity. The line connecting the first letter of the words *moral, muere, moras* and, *médicos* indicates a phonological/orthographic similarity. There are myriads of other connections that could be illustrated and other morphological relatives that could be included, but it is not possible to indi-

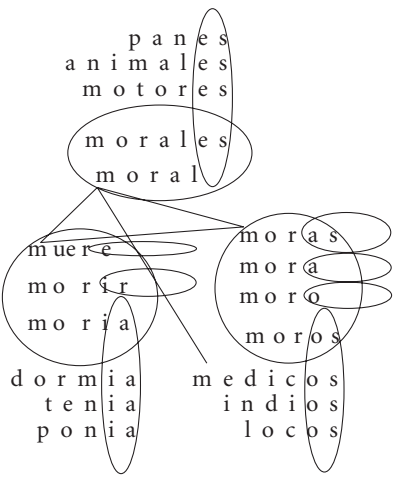

Figure 4. Patterns of interconnectedness between several words

cate all associations and all relatives without turning the figure into a tangled mass of lines and ellipses.

Note that while the members of the *moros* family are all connected based on *mor-* and the meaning they share, *moros* is also linked to other plural forms ending in *-os*. The *-as* of *moras* is similarly connected to other plurals ending in *-as* that are not exemplified in the figure for the sake of space. In like manner, the final *-ía* of *moría* is associated to other words that have the same imperfect tense and the same sequence of phonemes in the same position. According to Bybee,

> 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 identified. (1988: 127)

In other words, morphemes can be viewed as interconnected patterns that exist in two or more words that are both semantically and orthographically/phonologically similar. They exist as identifiable units even though their existence is not separate from the words that contain them.

As Allen and Badecker demonstrate, when *mor-os* is primed by *mor-ía* there is a great deal more inhibition that when the prime is *moral*. All three words share connections between the first three letters and none of the three share semantic characteristics. According to Allen and Badecker, this occurs because there is a morphological level, separate from the orthographic and semantic levels at which the stem of *mor-ía* (MOR-) and the stem of *mor-os* (MOR-) compete, while the stem of *moral* (MORAL) does not. Figure 4 shows how their findings can be explained in terms of orthographic and semantic overlap alone without the need to postulate separate morphological relationships.

Following the initial *mor-*, all of the members of the *moría* family have material such as *-e, -ir,* and *-ía* that have both orthographic and semantic connections to other words with these endings. In the same way, all of the members of the *moros* family have material such as *-as, -a, -o,* and *-os* that have both orthographic and semantic connections to other words with these endings. The *moral* family, however, is different. Following the initial *mor-* sequence of *moral*, for example, *-al* surely has orthographic links with word such a *peral* 'pear tree', and *cervical* 'cervical', but there is no corresponding semantic link. It is in this way that *moros* and *moría* have more in common than *moral.* Given the tight interconnectedness between the members of the *moría* family, some of which demonstrate less orthographic overlap than others (e.g., *muere/moría*),

it is not surprising that the prime *muere* will inhibit *moros* more than the prime *miran* that has no family members beginning with *mor-*.

I do not object to the assertion that *moros* and *moría* are both composed of the stem MOR- in contrast to *moral* whose stem is MORAL-, nor do I deny that there is such a thing as a morpheme. What I argue is that the units known as stems, morphemes, and affixes, as well as the boundaries between those units, may be defined in terms of phonological/orthographic and semantic associations between words that are stored as whole unsegmented entities. Morphology arises as the combination of the two and does not need to be represented separately.

This position is not without its difficulties, however. In a model of lexical connections, semantically transparent suffixes and prefixes should both function in a similar fashion. This is not always the case. In a study of English, Marslen-Wilson et al. (1994) found that semantically transparent prefixed words such as *unwind/rewind* and *unfasten/refasten* demonstrate a priming effect. The same does not occur between semantically transparent suffixed words such as *attractive/attraction* and *excitement/excitable*. It is tempting to explain the findings as due to a change in word category; the prefixed forms *unwind* and *rewind* are both verbs while in *attractive/attraction* the former is an adjective and the latter a noun. However, the authors also found that a base such as *punish* does prime for a suffixed form such as *punishment* even thought the base is a verb and the suffixed form a noun.

A model based on semantic links between lexical items is not adequate to capture these findings. According to Marslen-Wilson et al. (1994) there must be two different types of links. Those that hold between the suffixes of words that share the same stem must be inhibitory, but those that exist between prefixes are not. These effects are evidently language-specific since derivational prefixes in French demonstrated no priming effect while suffixes did (Segui & Zubizarreta 1995). It is unfortunate that there is currently no data on prefixes versus suffixes relevant to the Spanish language. Further research into morphological processing in Spanish will ultimately help determine which model is a more adequate representation of the role of morphology in Spanish.

7.5 Gender morphemes

Spanish nouns and adjectives ending in *-o* are masculine with a handful of exceptions. Those that end in *-a*, on the other hand, are generally feminine. This widespread pattern is responsible for the notion that *-a* and *-o* are morphemes. The question that arises in regards to visual word recognition is whether a word

such as *hermano* 'brother' is processed in terms of the stem *herman-* that it shares with *hermana* 'sister.' In this case, some hold that recognition would entail stripping the masculine morpheme and finding *herman-* in the mental lexicon. The other possibility is that words such as *hermano* are not parsed into morphemes, but are recognized as wholes.

Different measures of frequency offer a way of gaining insight into these competing hypotheses. If *hermano* is accessed as a whole word, its own frequency should influence recognition times. If, on the contrary, access is made to the stem *herman-*, the frequency of the stem, that is, the frequency of *hermano* plus the frequency of *hermana*, should be the deciding factor. Domínguez, Cuetos, and Segui (1999) performed a LDT in which they contrasted masculine and feminine forms with identical roots. Some feminine forms, such as *viuda* 'widow', are more frequent than their masculine counterparts, such as *viudo* 'widower'. These are called feminine dominant. In other pairs, such as *ciego/ciega* 'blind man/woman', the masculine form has a higher frequency of occurrence, and are referred to as masculine dominant.

In this experimental design, the stem frequency of the test items (i.e., frequency of *ciego* plus the frequency of *ciega*) was identical, while the frequency of the individual items (*ciego/ciega*) was manipulated. The experiment was administered to two groups of subjects in such a way that no subject responded to both the masculine and feminine form with the same stem. In the outcome, reaction times were significantly faster for masculine and feminine dominant forms. This indicates that recognition occurred on the basis of the individual words. If the stem frequency were the influencing factor, no differences would be expected between the time required to recognize *ciego* and *ciega* since both words have identical stem frequencies.

In a separate study, Domínguez et al. (1999) reversed the frequency relationships of the test items. They chose words with similar individual frequencies that differed in terms of their stem frequency. For example, *bello* 'beautiful' has an individual frequency of 58, while that of *culto* 'educated' is 59. However, the stem frequency of *bello* is 187, while *culto*'s stem frequency is 72. Under these conditions, no significant difference was found in the reaction times to the matched test words. Once again, individual frequency and not stem frequency appears to be responsible for the reaction times. The results of both experiments, when combined, demonstrate that the gender morpheme is not stripped from the stem in the process of word recognition. Instead, whole unparsed words are accessed from the mental lexicon.

7.6 Plural morphemes

The question that arises is whether the plural morpheme *-s* is processed in the same way as the gender morphemes. There is reason to believe that there may be differences. A word stripped of its gender morpheme yields a stem, not an extant word. In contrast, a word stripped of its plural morpheme remains a complete word. Sereno and Jongman (1997) conducted a series of experiments involving frequency effects with English plurals. In one study, they matched pairs of words with similar stem frequencies, but with widely differing plural and singular frequencies. For example, the stem frequency of *river* (freq. *river* + freq. *rivers*) is almost identical to the stem frequency of *window*. However, the plural form *windows* is much more frequent than the plural *rivers*. In other words, *window* is plural dominant while *river* is singular dominant. In an LDT, English speakers identified high-frequency plurals, such as *windows,* faster than low-frequency plurals such as *rivers*. This result is expected if both *rivers* and *windows* are accessed without parsing off the plural morpheme, but it would is difficult to account for if plural forms must be accessed via their stems.

Domínguez, Cuetos, and Segui (1999) carried out an experiment similar to Sereno and Jongman's, except that they used Spanish noun plurals and singulars. Their test words consisted of pairs of words with similar stem frequencies but whose plural and singular frequencies varied. They found that reaction times to singular words were dependent on the individual frequency of the singular form, not on the stem frequency. Not surprisingly, this suggests that recognition of singular nouns does not involve parsing them into morphemes, since singulars have no plural morpheme. However, reaction times to the plural forms the subjects saw were closely related to the stem frequency rather than the frequency of the individual plural word. In other words, Spanish plurals appear to be parsed.

However, a colleague and I (Eddington & Lestrade 2002) performed two LDTs similar to those of Domínguez et al. (1999) that yielded completely opposite results. In our experiment, words were equated on stem frequency and differed in terms of individual frequency. One group of subjects saw the plural forms of the test words and another group the singular forms of the same words. Plural forms with high individual frequencies were responded to more quickly than their lower frequency counterparts. Singular forms, on the other hand, showed no differential effects based on the individual frequency of the item.

The waters are further muddied by a later study by Domínguez, Cuetos, and Segui (2000). They report two additional studies relevant to plural process-

ing. In their 1999 study, test items were equated in terms of stem frequency, but had differing individual frequencies. In their 2000 experiment, they matched words with similar individual frequencies but varying stem frequencies. For example, the singulars *dama* and *dedo* have individual frequencies of 130 and 131 respectively, while their stem frequencies are 178 and 201. The plurals *botas* and *ratos* have individual frequencies of 55 and 57 respectively, while their stem frequencies are 64 and 286. The results of the 1999 study were that individual frequencies influenced reaction times to singular items, but not to plurals. However, when individual frequencies were equated, the stem frequency was found to affect reaction times to singulars as well as plurals. That is, *dedo* was recognized more quickly than *dama*. Plurals such as *ratos* were also recognized more quickly than plurals such as *botas*.

What is to be made of this mass of conflicting data? It may be tempting to conclude that there is something odd in the land of Spanish plurals. However, the data on English plurals is full of discrepancies as well (compare Sereno & Jongman 1997; Taft 1979). In an unpublished study, Harald Baayen and Robert Schreuder (personal communication) have also produced results on Dutch plurals that contradict their initial findings (Baayen, Dijkstra, & Schreuder 1997). Domínguez et al. (2000: 394–5) hypothesize three possible explanations for this messy state-of-affairs: 1) It may be explained in a model in which all lexical items are stored with massive numbers of interconnected units. Access to a singular form could trigger activation of its plural form, and vice versa, producing stem frequency effects. Under other conditions that are not yet understood, spreading activation may not occur which would yield frequency effects based on the individual item; 2) The effect of cumulative frequency may occur at the prelexical level before the word has been recognized; individual frequency, on the other hand, may become a factor only after the word has been identified; 3) A dual-route model may be at work in which words may be accessed both on the basis of their stem and on the basis of the entire unparsed word (Caramazza, Laudanna, Romani 1988; Schreuder and Baayen 1995).

7.7 Conclusions

The conflicting data from the plural studies on Spanish do not allow any generalizations to be made about how plurals and singulars are processed in visual word recognition. However, the evidence on the gender morphemes *-a* and *-o* is more straightforward. The morphemes are not stripped from the words during word recognition. Instead, the words are accessed as whole units.

One of the goals of research into morphological processing has been to evaluate whether a morphemic level exists that cannot be explained in terms of semantic or phonological/orthographic similarity. A number of experiments successfully demonstrate that morphology is not semantics, nor is it orthography. However, with a few exceptions, only a handful of studies accurately control for both semantic and orthographic/phonological overlap. Nevertheless, when this is done (e.g., Allen & Badecker 1999) morphology arises as a unique parameter in linguistic processing. In Section 7.4, I explain that morphology need not comprise a separate level of representation. Morphological relationships may emerge in the network of connections that exist between fully-formed lexical items (see Bybee 1988).

Chapter 8

Conclusions

What is the purpose of performing linguistic analyses? Does one carry out an analysis to show that a linguistic phenomenon can be made to fit into a formal mold? If so, that would make the practice of linguistics analogous to putting together some sort of complex jigsaw puzzle using linguistic materials. Such an analysis can be thought of as a description of a phenomenon, which is a worthy enterprise, however, taxonomy is not explanation.

When an empirical spin is put on the study of language, the kind of evidence that is relevant is automatically limited to measurable data produced by actual speakers. The nature of the research questions is also changed from questions about competence and formal mechanisms to questions of performance, since explanation in empirical linguistics ultimately relates to psychological processes.

In Chapter 1, I presented four reasons why the psychological relevance of many formal linguistic analyses is suspect. The first is that theoretical adequacy does not necessarily imply psychological significance. The fact that an analysis of a phenomenon can be carried out in a given formal framework does not in itself say anything about how speakers process that phenomenon in the course of comprehension and speech. Second, most formal approaches are non-empirical because they do not deal with tangible events that take place in time and space. This is problematic because such theories are able to escape potential falsification which is a requirement in empirical enterprises.

Third, formal analyses are established with little or no recourse to the speakers of the language via experimental psychology. Linguistic studies carried out as if language were separate from language speakers run the risk of being self-contained entities. If one's goal is to discover actual linguistic processes, the evidence must come from linguistic behavior and usage. Fourth, the limited base of evidence on which most analyses are founded is cause for skepticism. Language internal evidence such as phonotactic patterns and internal reconstruction has traditionally constituted the bulk of the evidence, but more external evidence from sources such as usage, corpus studies, and experiments is needed before claims about linguistics cognition can be made. Al-

though most formal analyses cannot be considered psychologically significant, they are viewed as the first step in arriving at psychologically valid analyses. Once linguistic analyses are based on empirical evidence that is obtained by psychological means, their psychological validity will be challenged less often.

What I hope to have achieved with this book is to give an overview of non-formal approaches to issues in Spanish phonology and morphology. In order to do so it was necessary to justify the use of experimental and quantitative evidence. In Chapter 2, experimental evidence was presented as a crucial part of an analysis that professes psychological significance. A strong and weak sense of psychological reality was defined and discussed in relationship to the traditional and experimental approaches to phonology. Several criticisms of experimental evidence were presented and contested, and a number of examples of experimentally acquired evidence were given that address the issue of the psychological reality of certain analyses. I presented several more examples in Chapter 3, whose focus is on the need for empirical evidence in linguistic analyses. In my view, too many linguistic processes are based on impressionistic observations rather than solid evidence. I presented a number of processes, such as depalatalization and uniform secondary stress, whose existence becomes questionable once a wider base of tangible evidence is considered.

One factor that is often overlooked in formal analyses is that of frequency. The important influence of phoneme cluster frequency, word frequency, and collocational frequency was discussed in Chapter 4. In Chapter 5, I defend the position that reference to past linguistic experience in the form of exemplars accounts for morphophonemic alternations in the formation of nominals and in gender assignment. An exemplar model is also able to account for dialectal variation. The utility of assuming an exemplar model is also evident in the data on diphthongization and stress assignment presented in Chapter 6.

In that chapter, I presented findings from different experimental paradigms that indicate that the syllable is an important processing unit in Spanish. Chapter 7 centers on morphological aspects of visual word recognition. The data suggest that there is a morphemic level of representation apart from the semantic, orthographic, and phonological levels. Gender morphemes appear to be stored as an integral part of the words that contain them, while the evidence relating to the status of the plural morpheme is contradictory.

The studies and experiments reported on in this book by no means constitute an exhaustive set. Nevertheless, many of the major issues that have occupied experimentalists and researchers with a quantitative bent are represented herein, along with many of the experimental methods currently utilized. I purposely included topics on which conflicting evidence exists. I believe that

questions of linguistic performance will ultimately be answered with empirical data. However, as in all sciences, lack of total consensus is a reality that must be dealt with.

Appendix

Experimental design, statistics, and research tools

9. Introduction

In the preceding pages, I have argued for the need for empirical data in linguistics. It is heartening to see a trend toward the use of more quantitative methods, however, I suspect that the reason that nonempirical studies continue to dominate many areas of linguistics is linked to the fact that few linguists receive explicit training in experimental design, data collection, and statistical analysis. Of course, this is less often the case in certain areas where one may find several field-specific treatises on these topics (e.g., language acquisition, Hatch & Farhady 1982; phonetics, Ladefoged 2003; corpus linguistics, Oakes 1998; sociolinguistics, Paolillo 2002). Often, the approaches and methods used in one field are not always directly applicable other areas of language study. Nevertheless, more and more linguistic specialties are producing their own works on methodology and statistics (e.g., syntax, Cowart 1997; historical linguistics, Embleton 1986; phonology, Beckman & Pierrehumbert forthcoming).

This chapter is meant to serve as a brief introduction to experimental design and statistics for those who are unfamiliar with them. Both of these issues are worthy of book-length discussions which are far beyond the scope of the present work. Gernsbacher (1994) discusses the issues in the different areas of psycholinguistics, and one can get a sense of the experimental paradigms that are important in each field from her work. I highly recommend Stanovich (1996) to those who are unfamiliar with experimental psychology. He defends the need for, and use of, tangible evidence in behavioral studies. Without inundating the reader with technical jargon or mathematical formulas, he explains why experiments and statistics are a crucial part of scientific research, and more specifically, how statistics are used in hypothesis testing.

My goal in this chapter is to give readers enough information about a few specific experimental paradigms so that they are able to better comprehend how some of the experiments reported on in previous chapters were carried out, and how the results were submitted to statistical analysis. In this way, read-

ers will be in a better position to think in experimental terms and may begin to design their own quantitative studies.

Before moving on to specific methods, I would like to suggest some other books for further reading at this point, rather than relegating those citations to an obscure footnote at the end. Anshen's (1978) 73 page book on statistics for linguistics is the ideal starting place for someone with no background in statistics. Its focus is mainly sociolinguistics. Butler's (1985) treatment is much more detailed and is written for a linguistic audience that is not overly comfortable with mathematics. Prideaux's work (1984) emphasizes experimental methods in linguistics rather than statistics and is highly recommended despite its age.

9.1 Correlation

A correlation is used to calculate the degree to which two measurable factors are related. For instance, there is a relationship or correlation between a person's height and his or her weight. The general tendency is that the taller a person is, the more the person will weigh. When both measurements increase (or both decrease) there is a positive correlation. A negative correlation involves one measurement decreasing as the other increases. For instance, in tests of mental abilities, as one's blood alcohol level increases, one's ability to perform physical and mental tasks decreases.

In order to perform a correlation, the data used must be numeric. There are two types of numeric data I will discuss to begin with. The first is ratio data. Examples of this would be formant frequencies, the duration of a stop, or a percentage of responses on a test item. With ratio data, the distance between values is constant and there is a natural zero. For example, the difference between 450 hertz and 500 is 50 hertz. This same 50 hertz distance is found between 100 and 150 hertz. On the hertz scale there is a natural 0 hertz, meaning lack of any periodicity.

Correlations may also be calculated on ordinal data as well. Ordinal data involve rankings. In a ranking, the time elapsed between the arrival of the first runner at the finish line and the second runner may be 0.5 seconds, while the difference between the second and third place runner may be .09 seconds. The differences in actual time are not relevant. Ordinal scales do not utilize the actual time differences, only the order.

9.1.1 *Examples of correlations*

One example of a ordinal correlation comes from two studies on diphthongization reported on briefly in Section 6.1 (Eddington 1996, 1998). Using different

Table 15. Order of likelihood of mid-vowels in neologisms

Suffix	Rank (1996)	Rank (1998)
-ero	1	2
-al	2	4
-(i)dad	3	5
-oso	4	1
-ista	5	3
-ísimo	6	8
-zuelo	7	7
-(c)ito	8	10
-azo	9	6
-(c)illo	10	9

Table 16. Example results from Albright et al. (2001)

Verb stem	Model Diph.	Subject Dipth.	Model Mid-vowel	Subject Mid-vowel
Lerr-	.50	.20	.50	.80
Gembl-	.47	.37	.53	.63
Bekt-	.13	.07	.87	.93

experimental paradigms in each study, I calculated how likely speakers were to prefer diphthongs or mid-vowels in producing Spanish neologisms that ended in one of ten suffixes. For example, is a person who puts *estiércol* 'manure' on a field an *estiercolero* or an *estercolero*? In order to compare the results of the two studies I ranked the suffixes in terms of how likely neologisms were to have a mid-vowel rather than a diphthong (Table 15).

Albright, Andrade, and Hayes (2001) performed a correlational analysis in their study of diphthongization as well. The data they used were ratio rather than ordinal. Subjects were asked to inflect 33 nonce verbs, such as *lerrar,* in a tense and person that could possibly contain a diphthong or not (*lierro* or *lerro).* They were interested in whether the phonetic context of the verbs was related to the subject's preferences. Their computer model predicted the probability of diphthongization for each nonce word based on the phonemes that surrounded the mid-vowel. The model's predictions were correlated with the percentage of subjects who chose each possible response. As Table 16 indicates, subjects produced forms such as *lierro* that contain diphthongs for the nonce stem *lerr-* 20% of the time and forms without diphthongs (e.g., *lerro*) 80% of the time. The model predicted a 50% probability for both possibilities.

A correlation considers pairs of values such as these, determines if there is a relationship between them, and calculates how strong that relationship is.

For ratio data, the most common statistical analysis is the Pearson product moment correlation which is expressed as a the correlation coefficient *r*. For ordinal data a Spearman rank order correlation is used whose correlation coefficient is ρ (rho). *R* and ρ can vary between 1 and –1. A perfect positive correlation yields *r* or ρ = 1 which means that when one measure moves up the other measure moves up in direct proportion to the other. A perfect negative correlation of *r* or ρ = –1 occurs when one measure moves one direction and the other moves proportionally in the opposite direction. When *r* or ρ = 0 there is no relationship between the scores on one measure and the scores on another. The study by Albright et al. yielded $r = .510$ while mine produced ρ = .758.

9.1.2 *Statistical significance*

The correlation coefficients for both of the studies are statistically significant, but what does that mean exactly? Significance may be thought of as the opposite of randomness. Assume that the subjects in Albright et al. used a roulette wheel to make their choice on the 33 test items. For each individual item, if the ball landed on black they would mark the first choice (e.g., *lierro*), and if it landed on red they would mark the second (e.g., *lerro*). We would assume that this procedure would result in a correlation coefficient close to 0. However, if the subjects followed this same routine thousands or millions of times, it is theoretically possible that their randomly generated choices would exactly match the model's predictions and yield a correlation of 1.00 on one go around. If the possibility is less than one in twenty that the predictions of the model and the subjects' choices could coincide by random selection (roulette wheel), the correlation is considered significant. This level of significance is called $p > .05$ meaning that the probability *p* is less that 1/20 (= .05) of getting the results by chance. The smaller the *p* value the more significant the results.

Now consider the coefficient obtained by Albright et al. and the one obtained in my experiments (.510 versus .758). Since mine is closer to 1.0 it is tempting to consider it more statistically significant, but that is not the case. My study had a significance of $p < .011$ while that of Albright et al. was much stronger at $p < .0025$. My study entailed comparing only ten pairs of numbers while Albright's involved 33.[49] In other words, when there are fewer test items the possibility of randomly achieving the same results is much higher than when there are more numbers to be correlated.

Consider a commercial that claims that four out of five dentists recommend toothpaste X. If only five dentists were actually consulted would you be impressed? Would you not be more motivated to buy it if 4,000 out of 5,000 dentists recommended toothpaste X, in spite of the fact that 4/5 and 4000/5000 are both 80%? In like manner, statistical formulas take into consideration factors such as the number of subjects, responses, and test items when calculating the statistical significance. This is important because it bears on how many subjects are needed in a study.

9.1.3 *Statistical software*

Statistics such as correlations may be calculated by hand, but there is really no need for tedious calculations given the widespread availability of computers and user-friendly software. There are numerous statistical packages available that calculate statistics such as correlations. Some are available without charge and may be downloaded from the internet. Some web sites offer on-line statistical calculators. Most universities have site licenses for commercial programs such as SPSS and SAS which may be installed on personal computers without a charge or with a minimal charge. These programs are recommended because of the large number of manuals and how-to books available (e.g., Kinnear 1999; Pavkov 2000). In addition, many universities sponsor statistics laboratories for their students and faculty. The advisors in the laboratories are usually well-versed in one or both of these commercial programs.

Step-by-step instructions on using a particular statistical package is not feasible in this chapter since each program is unique and any instructions provided here would soon be obsolete as newer versions of the programs become available. However, the format of the data is constant. For a rank order correlation there needs to be a column for each ranking with the individual test items in rows. Ratio data are structured in the same way. The data from the two studies would appear as in Table 17.

The data can be input directly into the statistics program or can be transferred from a spreadsheet or word processor. In fact, many spreadsheets can do

Table 17. Example data for correlational analysis

Spearman correlation			**Pearson correlation**		
Col. 1	**Col. 2**	**Col. 3**	**Col. 1**	**Col. 2**	**Col. 3**
-ero	1	2	*gembl-*	.47	.37
-al	2	4	*bekt-*	.13	.07
etc.			etc.		

correlations themselves. Inside of the statistics program one would select either a Spearman or a Pearson correlation, then specify that the data to be correlated appears in columns two and three.

The program will calculate a correlation coefficient and a level of significance. The level of significance may be calculated exactly ($p = .0025$) or the program may say that the results are significant at a certain level. If the results are significant at the .01 level, for example, this is reported as $p < .01$. The program may also give a two-tailed and a one-tailed level of significance. In general, the two-tailed level is used.[50] Albright et al. report the results of their statistical analysis in the established format: $r(31) = .510$, $p = .0025$. The *r* indicates that they performed a Pearson product moment correlation. The 31 in parentheses indicates the degrees of freedom; this is a number used by the program in its calculations and is related to the number of subjects or test items used. In a correlation, it is the number of pairs of scores that are compared (33) minus two. For this study, the correlation coefficient is .510 and the level of significance is .0025.

9.1.4 *Causation in correlations*

If a significant correlation is found, can one claim that one of the factors caused the other? For instance, there is a positive correlation between the amount of ice cream eaten and the number of drownings in the northeastern United States. Does this mean that eating ice cream makes it more likely that a person will drown? No, obviously there is a third factor that is responsible both for the ice cream sales and drownings – summer temperatures. While this may seem obvious, it is extremely common both in linguistic studies and in the popular media to claim that because a correlation has been found, so has a cause. Correlations demonstrate a relationship between factors, but not necessarily a causal relationship.

9.2 Chi square and multiple-choice experiments

Experiments involving multiple-choice responses are numerous. In Section 6.1, I discussed one such study (Eddington 1996) in which subjects were asked to decide which of two nonce words or neologisms sounded best to them. For example, is a person who collects *piedras* 'rock' a *piedrista* or a *pedrista*? In this case, only two choices were possible, but many studies include more than two choices. For example, Face's (2004) study, reported on in Section 6.3.1, compares the number of subjects who perceived nonce words ending in *-en, -an* and

-on to have final, penultimate, or final stress. This paradigm involves a more complex multidimensional chi square calculation that I will not expound on.

For the sake of simplicity, let us assume a test item with four choices (A, B, C, D) that 40 subjects have responded to. Completely random selection would result in ten responses being given to each choice, while something significant would obviously be occurring if all 40 subjects chose alternative B, for instance. Reality is usually something in between such as: A-15, B-9, C-11, D-5. The question is whether this distribution varies significantly from the random distribution. Chi square is an analysis that allow the significance of the distribution to be determined. It is one of the few statistics that is simple enough to be calculated by hand.

9.2.1 *Calculating chi square (χ^2)*

Chi square is used with nominal data, which contrasts with the ratio and ordinal data described in the previous section. Nominal data are basically frequency counts of the given variables within categories of the research group of interest (e.g., males versus females, Spanish versus English versus French speakers). In linguistic studies, the category of word final phonemes could include counts of the number of words ending in /d/ versus /l/ versus /n/. This is nominal data because there is no logical interval between /d/ and /n/ as there is between 55 decibels and 40 decibels. Of course, nominal data can be converted into ordinal data. That is, in terms of rank order more words may end in /n/, followed by /l/, and then /d/.

Returning to the multiple choice results (A-15, B-9, C-11, D-5) it is easy to calculate chi square by first determining the expected random distribution which would be ten responses to each choice. There were 15 actual responses to choice A. We first need to subtract the expected number of responses from the actual number, square that number, and divide it by the expected number:

$$\text{A} \quad \frac{(15-10)^2}{10} = 2.5$$

The next step is to repeat this process for all four of the responses and then sum the resulting numbers:

$$\text{A} \quad \frac{(15-10)^2}{10} = 2.5 \qquad \text{C} \quad \frac{(11-10)^2}{10} = 0.1$$

$$\text{B} \quad \frac{(9-10)^2}{10} = 0.1 \qquad \text{D} \quad \frac{(5-10)^2}{10} = 2.5$$

Since 2.5 + 0.1 + 0.1 + 2.5 equals 5.2, the chi square, which is represented as χ^2, is 5.2. The significance of this number is dependent on the degrees of freedom.

Table 18. Chi square table

df	p = 0.05	p = 0.01	p = 0.001
1	3.84	6.64	10.83
2	5.99	9.21	13.82
3	7.82	11.35	16.27
4	9.49	13.28	18.47
5	11.07	15.09	20.52

This is calculated as the number of choices minus one, therefore, there are 3 degrees of freedom or *df*. The statistical significance is determined by looking up the chi square and the degrees of freedom in a table such as Table 18.

Most statistics texts contain more extensive chi square tables as do many sites on the internet. At three degrees of freedom chi square must be 7.82 or higher in order to be significant at the lowest level of significance. Since 5.2 falls short of this we must conclude that the responses to the questionnaire are essentially random. The traditional way of reporting this is $\chi^2(3) = 5.2$, n.s., where 3 is the degrees of freedom and n.s. means 'not significant.' We know that the probability of this chi square being achieved by random draw from the hat is greater than .05, but we may want to know exactly what the probability is. A number of chi square calculators are available on the web. By inserting the degrees of freedom and the chi square into one of these calculators[51] the probability was found to be $p = 0.158$, which is much larger than the 0.05 level of significance.

9.3 Logistic regression with Varbrul

In sociolinguistics, the statistic of choice is logistic regression. For many years, a logistic regression program called Varbrul, which was specifically designed for linguistic analysis, has been freely available for download.[52] Sociolinguistic data, like experimental data, involve variation. The variation may be whether intervocalic /d/ is deleted or retained, or whether syllable-final /s/ is retained, aspirated, or deleted. The question that logistic regression tries to answer is what factors favor the use of one variant over another. In variationist studies, factors typically involve variables such as age, sex, social class, speech style, phonetic context etc. Some factors may favor deletion or retention, while other factors may disfavor it. Those that favor it may do so to varying degrees.

9.3.1 *Stress and syllable weight*

The best way to explain logistic regression as it is calculated with Varbul is with a concrete example. Waltermire (2004; Section 6.3.3) utilized Varbrul in his study of stress placement. In his study, subjects were asked to indicate where the stress falls on a series of written nonce words. Waltermire was interested in the role of syllable weight on the subjects' preferences. For example, the nonce word *bol.na.la* consists of a heavy first syllable followed by a light penultimate syllable and a light final syllable. When a subject indicated penultimate stress, this test response was encoded '(Phll.' The parenthesis indicates the beginning of a new line of data in Varbrul, 'P' indicates a penultimate response, 'h' says that the first syllable is heavy, and the two 'l's show that the final two syllables are light. The responses to all of the trisyllabic nonce words were encoded this way to be read by Varbrul.

In separate analyses Varbrul calculated the factors that favored final, penultimate, and final stress. The analysis for penultimate stress gave the results in Table 19 which are significant at $p < .05$.

Factor weights always range between one and zero. Weights of .550 and higher favor the factor, which in this case is penultimate stress. Factor weights below .450 disfavor penultimate stress, and hence favor antepenultimate or final stress. The effect of factors with weights between .450 and .550 is largely negligible. In this study, there are three factor groups: final, penultimate, and antepenultimate. Each factor group has two possible values: light and heavy. The range of a factor group is the highest factor weight minus the lowest. Factor groups with higher ranges affect stress assignment to a greater degree.

Waltermire's experiment demonstrates that the assignment of penultimate stress to the nonce words was most affected by the weight of the final syllable. Light final syllables strongly favor penultimate stress while heavy syllables strongly disfavor it. A heavy penultimate syllable slightly favors penultimate

Table 19. Waltermire's Varbrul results for penultimate stress

Syllable weights	Factor weights	Range
Final light syllable	.87	
Final heavy syllable	.13	74
Penultimate light syllable	.44	
Penultimate heavy syllable	.56	12
Antepenultimate light syllable	.54	
Antepenultimate heavy syllable	.46	8

stress while a light penultimate syllable slightly disfavors it. The influence of the weight of the antepenultimate syllable is almost absent.

9.3.2 *Documentation for Varbrul*

While Varbul is an invaluable tool for linguistic research it requires some time to master, and the steps required to perform an analysis are not always intuitive. Nevertheless, excellent documentation is available online; users of Apple computers should refer to Rand and Sankoff (1990)[53] while Robinson, Lawrence, and Tagliamonte (2001)[54] provide a manual for the Windows version. These manuals provide step-by-step instructions for the program, but say little about research design and interpreting the results. For this sort of information the reader is referred to Young and Bayley (1996) and Bayley and Young (forthcoming). The 1996 article includes a discussion of how to use an outdated version of Varbrul, but is still an important starting point for those not familiar with logistic regression since the information about data formatting and the interpretation of the results is still germane. The latter deals with the use of Varbul in sociolinguistics and the former for its use in studies of language acquisition; however, this does not detract from their usefulness as introductory guides. More advanced topics on Varbrul are covered by Paolillo (2002).

9.4 Analysis of variance and lexical decision tasks

9.4.1 *Lexical decision*

In a lexical decision task (LDT; Sections 6.2.2 and 7) subjects are presented words one at a time and must decide if a word is a real word or not by pressing either a 'yes' or 'no' key. The goal of the task is to determine if a previously presented word, the prime, can speed or slow the reaction time to a subsequently presented word called the target. Any effect is considered evidence that the two words are related in the mental lexicon.

LDTs can be varied along a number of parameters. The interval between the presentation of the prime and target can be manipulated by varying the number of intervening test items or by varying the number of milliseconds allowed to pass between the presentation of the prime and target (the stimulus onset asynchrony or SOA). If there is no interval between the prime and target the SOA represents the duration of the prime. In the masked priming paradigm, the prime is immediately preceded by a mask, such a series of hash marks, and immediately followed by the appearance of the target. The prime itself appears for a very short duration. These two elements, when combined, make the prime essentially invisible to the subjects' conscious awareness. The

cross-modal paradigm differs in that the primes are presented auditorily and the targets visually or vice versa.

9.4.2 *Analysis of variance (ANOVA)*

Domínguez, de Vega, and Cuetos (1997; Section 6.2.2) considered whether orthographic overlap and/or overlap in syllable structure influenced visual word recognition. They included test items such as those in Table 20.

The question that ANOVA is designed to answer is whether the means of one group are significantly different from those of another group. If not, the two groups must be considered subsets of the same group. More specifically, are the reaction time means for orthographically similar words (930 ms and 887 ms) significantly slower than the orthographically dissimilar items (872 ms and 871 ms)? On the other hand, did test words with overlapping syllable boundaries (930 ms and 871 ms) take longer to respond to than words with a differing syllable structure (887 ms and 872 ms)?

In general, orthographic overlap was found to slow reaction times, while syllabic overlap had no effect. However, two different factors were considered which raises the possibility that there could be an interaction between them. This turns out to be the case; orthographic similarity caused slower reaction times mainly when the syllable structure of the prime and target also overlapped (*nor.ma / nor.te*) in contrast to when the syllable structure was different (*no.ria / nor.te*). In other words, when considered alone syllable structure did not appear to be a significant factor. Only after the interaction between the factors was found to be significant could its influence be observed.

ANOVA is a calculation for interval or ratio data. In the case of reaction times, the data is ratio because there is a natural zero (0 ms). Interval data, such as dates and temperature, have no natural zero. The thing that is measured, which in this case is the reaction time, is called the dependent variable because its value depends on other factors, more specifically, it depends on the independent variables: syllabic overlap and orthographic overlap. The inde-

Table 20. Results of Domínguez et al. (1997) LDT experiment

Prime	Target	Relationship	Mean RT
nor.ma	*nor.te*	orthographic, syllabic	930 ms
no.ria	*nor.te*	orthographic, not syllabic	887 ms
sa.via	*nor.te*	not orthographic, not syllabic	872 ms
man.do	*nor.te*	not orthographic, syllabic	871 ms

Table 21. Hypothetical partial results from a LDT

Subject #	Prime	Target	Ortho. Overlap?	Syll. Overlap?	RT
1	*casco*	*caspa*	yes	yes	998
1	*noria*	*norte*	yes	no	916
1	*savia*	*cerco*	no	no	777
1	*mando*	*parte*	no	yes	612
2	*casco*	*caspa*	yes	yes	799
2	*noria*	*norte*	yes	no	971
2	*savia*	*cerco*	no	no	709
2	*mando*	*parte*	no	yes	801

pendent variables are the ones that are manipulated and whose influence on the dependent variable the experiment is designed to measure.

The computer program used to carry out the experiment (Section 9.4.3) gives the output of the experiment as a table of columns. Consider the hypothetical data in Table 21 keeping in mind that each subject responded to more than four test items.

The first row indicates that the first test subject saw *casco* as the prime and *caspa* as the target. The two words overlap in both orthography and syllable structure. Subject number one took 998 milliseconds to recognize that *caspa* is a legitimate Spanish word, which is the reaction time (RT). In this table, data about the nonwords is excluded. Notice that each subject sees each prime and target only once in the course of the experiment.

The data from the experiment may be uploaded into a statistics program such as SAS or SPSS from a word processor or spreadsheet. The statistical program would then be asked to perform an ANOVA with the data from the RT column as the dependent variable and the data from the two overlap columns as the independent variables. The program calculates three different statistics which are reported as values of *F*. It also computes a level of significance for each. The first two statistics indicate the significance that the reaction time was affected by orthographic or syllabic similarity beyond chance. The third is whether there is an interaction between the two independent variables.

ANOVAs are somewhat more complex than chi square and correlational analyses. There are many details about the statistical analysis and interpretation of ANOVAs that cannot be covered here. This outline should serve to introduce the major aspects of ANOVA as it relates to lexical decision tasks. The reader is referred to Butler (1985 chapter 10), Pavkov (2000), Woods, Fletcher, and Hughes (1986 Chapters 11–12) for in-depth treatments.

9.4.3 *LDT programs*

It should be apparent that conducting an LDT requires the use of computer software. Because it is a popular experimental paradigm numerous programs are available. Some programs are free of charge and may be downloaded from the internet; Psyscope[55] is written for Apple computers while DMDX[56] runs on the Windows platform. Documentation on their use is also available on the internet. These programs control such things as the order that the stimuli are presented, how long they appear for, the interval between stimulus items, etc. They also measure reaction times and give the results of the experiment in a format that may be incorporated into a spreadsheet, statistical package, or word processor.

9.5 Conducting experiments

There are a number of issues that arise in relationship to carrying out an experiment. The first is how many test subjects and test items are needed. It should be apparent from the above discussion that when more test subjects and test items are included the chances are better that small differences will be statistically significant. There is no simple answer to the question of how many test items and subjects are required; however, since most experiments are variations on previously conducted research, it is helpful to consult the model experiment for guidance in this regard.

In the past few decades, university administrations have taken an active role in overseeing any studies involving human or animal subjects. In part, this is due to a number of well-publicized experiments in psychology (e.g., Milgram 1974) whose methods raised ethical questions. Universities are also keenly interested in avoiding legal action or public outrage. As a result, any experiment with human subjects must be approved by an internal review board no matter how innocuous it may seem. The existence of this policy is common knowledge in fields such as the social sciences that traditionally carry out experiments. I mention this here because linguists who have not been introduced to experimental methods may be unaware of internal review policies.

9.6 Internal and external validity

Any experimentally acquired evidence needs to be evaluated in terms of its validity. Internal validity refers to the question of whether the experiment actually measures what it is designed to measure. The possibility always exists that an experimental outcome is the result of a factor other than the ones that are

initially assumed. For this reason, good experiments go to great lengths to minimize or eliminate extraneous factors. In lexical decision tasks, for instance, it has been established that word frequency and word length affect reaction times. An experiment that measures reaction time differences to groups of words with different morphological structure could find a significant difference and claim that it is due to morphology. However, if the word length and frequency of the words have not been equalized across groups, the results may be due to those factors rather than morphological differences. Therefore, the internal validity of the study would be suspect.

In order to obtain valid results it is often necessary not to divulge specifics about what an experiment is designed to measure. Knowledge of the purpose of the experiment could cause the subjects to react in a manner that they would not normally do under other conditions. In some instances, the expectations or biases of the experimenter him/herself may be consciously or unconsciously communicated to the subjects. In 1904, a horse named Clever Hans could purportedly perform mathematical calculations. It would give the answer to a problem by tapping its hoof on the ground. Later, it was demonstrated that the horse's trainer was doing the calculations and was unwittingly using body language to signal the horse to stop tapping once it had tapped the correct number of times. The best way to avoid this scenario is to use a double-blind method in which neither the subjects nor the person administering the test is aware of the purpose of the study.

The most common threat to the validity of an experiment is what is termed the Hawthorne effect. In an experimental setting, the subjects know that they are being tested, even though they may not know exactly what is being measured. Therefore, the setting itself may motivate them to perform in a way that they would not outside of the experimental setting. As already mentioned in Section 2.3.4, if converging results are obtained from several experiments that used differing methods, the results are less likely to be the product of the experimental situation itself.

This brings up the issue of external validity. The findings of an experiment are thought to be externally valid if they are not specific to the particular subjects who participated nor to the specific experimental task. On the one hand, researchers carefully design experiments in order to control for every foreseeable factor that could unfavorably influence the outcome. In doing so, however, the way that language is used during the experiment is less likely to reflect the way language is used in more natural circumstances. Therefore, where possible, experimental evidence should be compared with data from natural speech

such as that resulting from sociolinguistic interviews, observation, slips of the tongue, etc.

Another threat to external validity does not lie in the experimental setting, but in the subjects themselves. The possibility exists that they may not be representative of other speakers of the language. That is, the results may not be valid for the linguistic community as a whole. Political pollsters go to great lengths to insure that their findings are based on a sampling of people from many different ethnicities, occupations, regional affiliations, etc. More often than not, psycholinguists' subjects are students who have volunteered to participate. It is entirely possible that students of a certain age, or students who volunteer (as opposed to those who choose not to volunteer) process language in a manner different from other people.

At first, it may seem absurd to think that linguistic diversity could be compared to political diversity, but sociolinguistics has proven that diversity exists within one small community and even in one individual. Because of the issues of validity, the definitive answer to a scientific question is never answered by one study. It is only when a substantial body of supporting evidence is amassed that the workings of the human language faculty will become more transparent.

9.7 Tools for researching Spanish phonology and morphology

Choosing test items and controlling for factors such as word length and frequency are made much easier with the help of a number of research tools. One place to start is a raw list of over 90,000 Spanish words that may be downloaded from the internet.[57] This list does not include inflected forms, but is helpful for locating words with particular suffixes or phonological shapes, for example. Several frequency dictionaries are available which are based almost exclusively on written corpora. Juilland and Chang-Rodríguez (1964) was extracted from a small sample of Spanish. Data about the frequency of words in different written genres is included and the words are tagged according to part of speech. Tagging is important because in many studies it is important to know if '*casa,*' for instance, is a noun 'house' or a verb 's/he marries.' Alameda and Cuetos (1995) frequency count is based on a larger sample of about 5 million words. Although it is not tagged, it includes a count of the number of letters and the CV composition of each word.

The dictionary compiled by Sebastián-Gallés, Martín-Antonín, Cuetos-Vega, and Carreiras-Valiña (2000) is fully tagged and is available for a fee in CD-ROM format. It is based on a 5 million word corpus of written material.

The electronic format permits the user to perform intricate searches based on things such as narrative style, gender, tense, part of speech, etc. Lemma frequency is also encoded. (The lemma frequency of the noun *río*, for example, would be the sum total of the frequencies of *río* and *ríos*). This dictionary does have drawbacks, however. The search program has some technical flaws and there is no user support. In addition, close inspection of some of the searches I have performed indicates that many of the words have been mistagged. The forthcoming frequency dictionary by Davies should prove to be the standard reference frequency dictionary of Spanish. It is based on a corpus of 20 million words and will contain part of speech information and lemma frequencies.

In some studies, it is crucial to have information on units smaller or larger than the word. Carreiras, Álvarez, and de Vega (1993) provide a list of syllable frequencies, while Guirao and García-Jurado (1990) have compiled a list of the phoneme frequency of Latin American Spanish. Two corpora containing transcriptions of oral speech are available online. *The Child Language Data Exchange System* (CHILDES)[58] is a repository of transcribed interactions between caregivers and young children. *Corpus oral de referencia de la lengua española contemporánea*[59] includes adult conversations taken from many different genres (e.g., sportscasts, doctor's visits, informal conversations, etc.). Written corpora from Argentina and Chile are also available from the same site.[60]

One of the most versatile tools is *Corpus del español*, a 100 million word corpus compiled by Davies that includes materials from the 13th through 20th centuries. A user-friendly search engine to this database is accessible online.[61] Searches may be performed for single words, and more complex searches are possible based on lemma, part of speech, frequency, synonyms, sequences of several words, etc. The site includes tutorials and other documentation on how to best utilize it. The outcome of the search is given in the context of the sentence in which the material appears in the original document.

In Chapter Five and Section 6.3 I reported a number of computer simulations carried out using the Analogical Modeling (AM) algorithm. This program is available for UNIX, Apple, and Windows operating systems.[62] Parkinson (2002) provides instructions on how to use the AM program, while Lonsdale (2002) discusses how to structure the data files used by AM. Another computer program, the Tilburg Memory-based Learner (TiMBL) by Daelemans, Zavrel, van der Sloot, and van den Bosch (2001), has proven useful in modeling linguistic phenomena analogically as well, although I have not discussed it in this book. TiMBL runs on UNIX/LINUX machines. The program and user's manual may be downloaded from the internet.[63]

Notes

1. An earlier version of this chapter appeared as Eddington, David. 1996. The psychological status of phonological analyses. *Linguistica* 31, 17–37.

2. Schütze (2003), on the other hand, argues that grammaticality judgments are an important source of evidence for the reality of linguistic concepts, but have no special status in regards to other sources of evidence.

3. This is analogous to Dochery and Foulkes (2000) 'data-driven' versus 'theory-driven' approaches to linguistics.

4. In the classical movement of psychology, psychologists relied heavily on their own introspections. Later schools of psychology rejected this as methodologically flawed and subjective (Spence 1956: 4–15).

5. According to some definitions, informant judgments are an example of internal evidence.

6. It is regrettable for the field of linguistics that a well-known linguist such as Chomsky has been known to intentionally disregard evidence that contradicts his theories, and to be dishonest with the data in other ways (see Levine & Postal 2004).

7. This is not to say that formal theories have no useful application. For example, the formal notations set forth in Chomsky and Halle (1968) were employed by Otero (1971) to depict the phonological evolution of Spanish from Latin. Hartman (1981) has developed a computer program based on Otero's rules which derives Spanish words from Latin words. This program serves as a valuable research tool which can be used to test theories about diachronic Spanish linguistics. If it were not for the formal notations of classical generative phonology it is very unlikely that it would have been possible to develop this tool.

8. An earlier version of this chapter appeared as Eddington, David (1999). Role of experiments in phonological investigation. *Journal of Quantitative Linguistics, 6*, 14–28.

9. The distinction between adequacy in fact and adequacy in principle has been borrowed from Fought (1973: 157).

10. This study did receive a great deal of criticism (Chandler & Skousen 1997; Seidenberg & Hoeffner 1998). It is cited here to as an example of a study whose goal is strong reality.

11. García-Bellido offers a similar analysis (1986).

12. Bertinetto (1992) gives an example of a psycholinguistic experiment that produced different results when it was replicated.

13. Aspiration is a misnomer since s-aspiration would technically denote [s^h]. Nevertheless, I will use this term since it has become standardized in the literature on Spanish linguistics.

14. I am indebted to José Ignacio Hualde for this insight.

15. I was unable to verify the existence of this word. It may have been invented by the authors for the purposes of the experiment.

16. The phoneme is traditionally thought of as /d/ although there are good motivations for considering it to be /ð/ (Danesi 1982).

17. In many dialects, /θ/ does not belong to the phonemic inventory. Instead, these words are realized as /s/.

18. In many dialects, /x/ is realized as [h] and the case could be made that the phoneme is actually /h/.

19. It is interesting to note that /k/ > /s/ (e.g., *opaque* > *opacity*) was found to be productive in English (Pierrehumbert 2002).

20. The phoneme /ʎ/ is not universal in every dialect of Spanish. In many it is /y/ or /ỹ/.

21. It is somewhat odd that in the second edition of Whitley's monograph (2002), Kelm's study is cited, yet Whitley leaves his discussion of the four versus three levels of English and Spanish intonation intact.

22. He also recorded the speakers performing a role play in their non-native language, but these data are not important to the present discussion.

23. This one item is based on my personal observation.

24. The actual nonce words used in the study do not appear in the article, but were graciously provided by Hernán Emilio Pérez.

25. These words were found on www.corpusdelespanol.org, a 100 million word searchable corpus.

26. www.google.com

27. Vogel Sosa and MacFarlane's frequencies were based on a database of three million words, while Muller's was based on a 1.1 million word database.

28. In this study, the phonemic attributes of words are assumed to be the relevant variables. However, AM can also incorporate other variables such as sociolinguistic or syntactic variables (Skousen 1989: 97–100).

29. Actually a pointer is selected, but in order to simplify the explanation I have not discussed the role of pointers in the calculation.

30. Dual-gendered nouns such as *mar* 'sea', and *estudiante* 'male or female student' were not included, nor are nominalized adjectives such as *rojo* 'red'.

31. In Eddington (2002c), I performed other simulations using variable encodings that included more and less phonological information from each word, as well databases based on token and type frequency. The success rates from those simulations ranged from 93.6% to 96.4% correct.

32. Making predictions as if the word were unknown is tantamount to testing the productivity of the process. Evidence suggests that productivity is based on type frequency (Baayen & Lieber 1991; Bybee 1985, 1995; Wang & Derwing 1994).

33. It should be noted that relationships other than those between pairs of words are possible. One could incorporate variables representing a number of different verbal forms in order to predict the *yo* form, for example.

34. *anexar, confesar, dispersar, expresar, expulsar, impulsar, intrusar, precisar, profesar, progresar, regresar, revisar, televisar, tensar.*

35. *adoptar, afectar, cantar, conjuntar, desertar, desinfectar, detectar, difractar, editar, ejecutar, exceptar, excretar, exentar, infectar, insertar, intentar, inventar, inyectar, objetar, optar, proyectar, recolectar redactar, secretar, sujetar.*

36. Of course *escindir* would need to be marked as an exception.

37. In the context, the noun is clearly related to *rebelar* 'to rebel' and not to *revelar* 'to reveal.'

38. buscon.rae.es

39. www.corpusdelespanol.org

40. The only exception may be *cesión*. This word appears once in the 13th century, but the context it appears in, coupled with the orthographic uncertainty of the period, makes it possible that the word is actually *sesión*.

41. Prieto (1992) suggests that Bolivian Spanish may be of this type, as does Crowhurst's (1992) data on Sonoran Mexican Spanish. Based on my own informal observations, the Spanish of the Canary Islands may also have this type of diminutives.

42. The results reported here differ somewhat from those reported in Eddington (2002b). The reason for this is that in the previous database, a word such as *muerto* was divided into variables in this manner: mu/er/t/o. The current database divides words in such a way that the members of a diphthong are kept together: m/ue/r/t/o.

43. Words such as *grieto* 'crack' are not considered to have diphthongizing stems if none of their morphological relatives contains a non-diphthongizing stem such as **gret-*. *Tierra* 'earth', on the other hand, does have a diphthongizing stem, since it has relatives such as *terreno* 'terrain'.

44. This example is taken from Rapp 1992.

45. The authors used a measure of frequency called frequency of mention. This was obtained by asking Spanish speakers to write down words beginning with a particular syllable and counting the number of times a word was mentioned by the speakers.

46. Irregularly stressed words are those that are stressed on the antepenultimate syllable, or that have final stress and end in a vowel or *s*, or that have penult stress and end in a consonant other than *s*.

47. I am indebted to José Antonio Mompeán for administering the surveys.

48. The definition that entails that a morpheme must share both phonetic and semantic characteristics is not universal. Under an alternative definition which emphasizes semantics, *voy* and *fui* do have a common morpheme.

49. The difference is also due to the fact that Spearman and Pearson correlations are based on different kinds of data and different formulas, but this should not detract from the point being made.

50. See Butler (1985: 72-74) for a discussion of the difference between one- and two-tail levels of significance.

51. I used the one at: www.fourmilab.ch/rpkp/experiments/analysis/chiCalc.html when I wrote the text. Another is located at: faculty.vassar.edu/lowry/tabs.html#csq, and another at: www.graphpad.com/quickcalcs/chisquared1.cfm. The nature of the internet is such that these may or may not be available at the same site in the future. However, a search of the internet should yield a number of chi square calculators.

52. The MacIntosh version may be found at: www.crm.umontreal.ca/~sankoff/GoldVarb_Eng.html. The Windows version may be downloaded at: www.york.ac.uk/depts/lang/webstuff/goldvarb

53. www.crm.umontreal.ca/~sankoff/GoldVarbManual.Dir

54. www.york.ac.uk/depts/lang/webstuff/goldvarb/manual/manualOct2001.html

55. psyscope.psy.cmu.edu

56. www.u.arizona.edu/~kforster/dmdx/dmdx.htm

57. ftp.sil.org/pub/data/span-lex.zip. The file may be opened as a text (ASCII) file, although some computers will misinterpret the LEX extension as being a different format.

58. childes.psy.cmu.edu

59. ftp.lllf.uam.es/pub/corpus/oral

60. ftp.lllf.uam.es/pub/corpus

61. www.corpusdelespanol.org

62. humanities.byu.edu/am/amdownloads.html

63. ilk.kub.nl

References

Aha, David W., Dennis Kibler, & Marc K. Albert (1991). "Instance-based learning algorithms". *Machine Learning, 6*, 37–66.

Alameda, José Ramón, & Fernando Cuetos (1995). *Diccionario de frecuencias de las unidades lingüísticas del castellano*. Oviedo, Spain: University of Oviedo Press.

Albright, Adam, Argelia Andrade, & Bruce Hayes (2001). "Segmental environments of Spanish diphthongization". *UCLA Working Papers in Linguistics, 7,* 117–151.

Albright, Adam, & Bruce Hayes (1999). *An automated learner for phonology and morphology*. Manuscript, www.linguistics.ucla.edu/people/ hayes/learning/learner.pdf

Alegre, Maria, & Peter Gordon (1999). "Frequency effects and the representational status of regular inflections". *Journal of Memory and Language, 40*, 41–61.

Allen, Mark, & William Badecker (1999). "Stem homograph inhibition and stem allomorphy: Representing and processing inflected forms in a multilevel lexical system". *Journal of Memory and Language, 41*, 105–123.

Allen, Phillip, Michelle McNeal, & Donna Kvak (1992). "Perhaps the lexicon is coded as a function of word frequency". *Journal of Memory and Language, 31*, 826–844.

Alonso, Dámaso, Alonso Zamora Vicente, & María José Canellada (1950). "Vocales andaluzas: Contribución al estudio de la fonología peninsular". *Nueva Revista de Filología Hispánica, 4*, 209–230.

Alvar, Manuel, A. Llorente, & G. Salvador (1973). *Atlas lingüístico y etnográfico de Andalucía*, Vol. 6. Granada: Universidad de Granada, Concejo Superior de Investigaciones Científicas.

Alvar, Manuel, A. Llorente, T. Buesa, & Elena Alvar (1980). *Atlas lingüístico y etnográfico de Aragón, Navarra y Rioja*, Vol. 12. Paracuellos del Jamara, Madrid: Departamento de Geografía y Lingüística, Institución Fernando el Católico de la Excma. Diputación Provincial de Zaragoza.

Álvarez, Carlos J., Manuel Carreiras, & Manuel de Vega (1992). "Estudio ortográfico de la ortografía castellana (1): La frecuencia silábica". *Cognitiva, 4*, 75–105.

Álvarez, Carlos J., Manuel Carreiras, & Manuel de Vega (2000). "Syllable-frequency effect in visual word recognition: Evidence of sequential-type processing". *Psicológica, 21*, 341–374.

Álvarez, Carlos J., Manuel Carreiras, & Marcus Taft (2001). "Syllables and morphemes: Contrasting frequency effects in Spanish". *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27*, 545–555.

Alvord, Scott (2003). "The psychological unreality of quantity sensitivity in Spanish: Experimental evidence". *Southwest Journal of Linguistics, 22,* 1–12.

Ambadiang, Théophile (1996). "La formación de diminutivos en español: ¿Fonología o morfología?" *Lingüística Española Actual, 18*, 175–211.

Ambadiang, Théophile (1997). "Las bases morfológicas de la formación de diminutivos en español". *Verba, 24,* 99–132.

Ammer, Christine (1993). *Southpaws and Sunday punches and other sporting expressions*. New York: Dutton Books.

Anshen, Frank (1978). *Statistics for linguists*. Rowley, MA: Newbury House.

Aske, Jon (1990). "Disembodied rules versus patterns in the lexicon". In Kira Hall, Jean-Pierre Koenig, Michael Meacham, Sondra Reinman, & Laurel A. Sutton (Eds.), *Proceedings of the sixteenth annual meeting of the Berkeley Linguistic Society* (pp. 30–45). Berkeley: Berkeley Linguistic Society.

Baayen, R. Harald, Ton Dijkstra, & Robert Schreuder (1997). "Singulars and plurals in Dutch: Evidence for a parallel dual-route model". *Journal of Memory and Language, 37*, 94–117.

Baayen, R. Harald, & Rochelle Lieber (1991). "Productivity and English derivation: A corpus-based study". *Linguistics, 29,* 801–843.

Baayen, Harald, & Robert Schreuder (1999). "War and peace: Morphemes and full forms in a noninteractive activation parallel dual-route model". *Brain and Language, 68,* 27–32.

Bailey, Todd M. (1995). *Nonmetrical constraints on stress*. PhD diss., University of Minnesota.

Baker, Gordon P., & Peter Michael Stephan Hacker (1984). *Language, sense and nonsense: A critical investigation into modern theories of language*. Oxford: Blackwell.

Baker, William J. (1979). "On draining the conceptual swamp". In Gary D. Prideaux (Ed.), *Perspectives in experimental linguistics* (pp. 141–149). Amsterdam: John Benjamins.

Balota, David A. (1983). "Automatic semantic activation and episodic memory encoding". *Journal of Verbal Learning and Verbal Behavior, 22,* 88–104.

Bard, Ellen Gurman, Dan Robertson, & Antonella Corace (1996). "Magnitude estimation of linguistic acceptability. *Language, 72,* 32–68.

Bárkányi, Zsuzsanna (2002). "A fresh look at quantity sensitivity in Spanish". *Linguistics, 40*, 375–394.

Barlow, Michael, & Suzanne Kemmer (Eds.). (2000). *Usage-based models of language*. Stanford, CA: Center for the Study of Language and Information.

Barsalou, Lawrence W. (1999). "Perceptual symbol systems". *Behavioral and Brain Sciences, 22,* 577–609.

Bayley, Robert, & Richard Young (In press). "VARBRUL: A special case of logistic regression".

Becker, Curtis A. (1980). "Semantic context effects in visual words recognition: An analysis of semantic strategies". *Memory and Cognition, 8*, 493–512.

Beckman, Mary, & Janet Pierrehumbert (In press). *A laboratory course in phonology*. Basil Blackwell.

Benoît, Jacques, & Jorge A. Gurlekian (1992). "Étude comparée de quatre paramètres acoustiques des occlusives en espagnol de Buenos Aires et en français de Montréal". *Revue Québécoise de Linguistique, 22*, 257–272.

Bentin, Shlomo, & Laurie B. Feldman (1990). "The contribution of morphological and semantic relatedness to repetition priming at short and long lags: Evidence from Hebrew". *The Quarterly Journal of Experimental Psychology, 42A*, 693–711.

Benua, Laura (1995). "Identity effects in morphological truncation. Papers in Optimality Theory". *University of Massachusetts Occasional Papers in Linguistics, 18*, 77–136.

Bergen, John J. (1978). "A simplified approach for teaching the gender of Spanish nouns". *Hispania, 61*, 865–876.

Bertinetto, Pier Marco (1992). "The use and misuse of external evidence in phonology". In Wolfgang U. Dressler, Hans C. Luschützky, Oscar E. Pfeiffer, & John R. Rennison (Eds.), *Phonologica 1988* (pp. 33–47). Cambridge: Cambridge University Press.

Bertinetto, Pier Marco, Maddalena Agonigi, Lorenzo Cioni, María Luisa García Lecumberri, & Estibalitz González Parra (1998–1999). "An experimental investigation on syllable geometry in Spanish". *Quaderni del Laboratorio di Linguistica, 12–13,* 57–66.

Biber, Douglas (1998). *Corpus linguistics: Investigating language structure and use*. Cambridge: Cambridge University Press.

Black, Maria, & Shulamuth Chiat (1981). "Psycholinguistics without 'psychological reality.' " *Linguistics, 19,* 37–61.

Bod, Rens (1998). *Beyond grammar*. Stanford, CA: Center for the Study of Language and Information.

Botha, Rudolf P. (1971). *Methodological aspects of transformational generative phonology*. The Hague: Mouton.

Botha, Rudolf P. (1973). *The justification of linguistic hypotheses*. The Hague: Mouton.

Botha, Rudolf P. (1989). *Challenging Chomsky*. Oxford: Basil Blackwell.

Bradley, Diane (1980). "Lexical representation of derivational relation". In Mark Aronoff & Mary-Louis Keaton (Eds.), *Juncture* (pp. 37–55). Saratoga, CA: Anma Libri.

Bradley, Dianne C., Rosa M. Sánchez-Casas, & José E. García-Albea (1993). "The status of the syllable in the perception of Spanish and English". *Language and Cognitive Processes, 8,* 197–233.

Bresnan, Joan, & Ronald M. Kaplan (1982). "Introduction: Grammars as mental representations of language". In Joan Bresnan (Ed.), *The mental representation of grammatical relations* (pp. xvii–lii). Cambridge, MA: Massachusetts Institute of Technology Press.

Brisk, María Estela (1976). "The acquisition of Spanish gender by first grade Spanish-speaking children". In Gary D. Keller, Richard V. Teschner, & Silvia Viera (Eds.), *Bilingualism in the bicentennial and beyond* (pp. 143–160). New York: Bilingual Press.

Bromberg, Sylvain, & Morris Halle (2000). "The ontology of phonology (revised)". In Noel Burton-Roberts, Philip Carr, & Gerard Docherty (Eds.), *Phonological knowledge: Conceptual and empirical issues* (pp. 19–37). Oxford: Oxford University Press.

Brown, Esther (1999). *The posteriorization of labials in Spanish: A frequency account*. Manuscript, University of Colorado-Boulder.

Brown, R., & D. Mc Neill (1966). "The 'tip of the tongue' phenomenon". *Journal of Verbal Learning and Verbal Behavior, 5,* 325–337.

Bull, William E. (1965). *Spanish for teachers: Applied linguistics*. New York: Ronald Press.

Bunge, Mario (1980). *Epistemología*. Barcelona: Ariel.

Burton, Peter G. (1990). "A search for explanation of the brain and learning: Elements of the synchronic interface between psychology and neurophysiology". *Psychobiology, 18,* 119–161.

Burzio, Luigi (1996). "Surface constraints versus underlying representations". In J. Durand & B. Laks (Eds.), *Current trends in phonology: Models and methods* (pp. 97–112). Salford: European Studies Research Institute, University of Salford Publications.

Butler, Christopher (1985). *Statistics in linguistics*. Oxford: Basil Blackwell.

Butt, John, & Carmen Benjamin (1994). *A new reference grammar of modern Spanish*, 2nd ed. London: Edward Arnold.

Butterworth, B. (Ed.). (1983). *Lexical representation. Language production*, Vol. 2 (pp. 257–294). London: Academic Press.

Bybee, Joan (1985). *Morphology*. Amsterdam: John Benjamins.

Bybee, Joan (1988). "Morphology as lexical organization". In Michael Hammond & Michael Noonan (Eds.), *Theoretical approaches to morphology* (pp. 119–141). San Diego, CA: Academic Press.

Bybee, Joan L. (1994). "A view of phonology from a cognitive and functional perspective". *Cognitive Linguistics, 5,* 285–305.

Bybee, Joan (1995). "Regular morphology and the lexicon". *Language and Cognitive Processes, 10,* 425–455.

Bybee, Joan L. (1998). "The emergent lexicon". In M. Gruber, C. Derrick Higgins, K. S. Olson, & T. Wysocki (Eds.), *Proceedings of the Chicago Linguistic Society*, Vol. 34 (pp. 421–435). Chicago: Chicago Linguistic Society.

Bybee, Joan L. (2000). "The phonology of the lexicon: Evidence from lexical diffusion". In Michael Barlow & Suzanne Kemmer (Eds.), *Usage-based models of language* (pp. 65–85). Stanford, CA: Center for the Study of Language and Information.

Bybee, Joan (2001). *Phonology and language use*. Cambridge: Cambridge University Press.

Bybee, Joan L., & Dan I. Slobin (1982). "Rules and schemas in the development and use of the English past tense". *Language, 58,* 265–289.

Bybee, Joan L., & Carol Lynn Moder (1983). "Morphological classes as natural categories". *Language, 59,* 251–270.

Bybee, Joan, & Elly Pardo (1981). "On lexical and morphological conditioning of alternations: A nonce-probe experiment with Spanish verbs". *Linguistics, 19,* 937–968.

Campbell, Lyle (1979). "The quest for psychological reality: External evidence in phonology". In Eli Fischer-Jørgensen, Jørgen Rischel, & Nina Thorsen (Eds.), *Proceedings of The Ninth International Congress of Phonetic Sciences*, Vol. 2 (pp. 74–78). Copenhagen: University of Copenhagen.

Campbell, Lyle (1986). "Testing phonology in the field". In John J. Ohala & Jeri Jeager (Eds.), *Experimental Phonology* (pp. 163–173). Orlando: Harcourt Brace Jovanovich.

Caramazza, Alfonso, Alessandro Laudanna, & Cristina Romani (1988). "Lexical access and inflectional morphology". *Cognition, 28,* 297–332.

Carr, Philip (1990). *Linguistic realities: An autonomist metatheory for the generative enterprise*. Cambridge: Cambridge University Press.

Carr, Philip (2000). "Scientific realism, sociophonetic variation, and innate endowments in phonology". In Noel Burton-Roberts, Philip Carr, & Gerard Docherty (Eds.), *Phonological knowledge: Conceptual and empirical issues* (pp. 67–104). Oxford: Oxford University Press.

Carreira, María (1991). "The alternating diphthongs of Spanish: A paradox resolved". In Héctor Campos & Fernando Martínez-Gil (Eds.), *Current studies in Spanish linguistics* (pp. 407–445). Washington DC: Georgetown University Press.

Carreiras, Manuel, Carlos J. Alvarez, & Manuel de Vega (1993). "Syllable frequency and visual word recognition in Spanish". *Journal of Memory and Language, 32,* 766–780.

Carreiras, Manuel, & Manuel Perea (2002). "Masked priming effects with syllabic neighbors in a lexical decision task". *Journal of Experimental Psychology: Human Perception and Performance, 28,* 1228–1242.

Cassano, Paul Vincent (1972). "French influence in the Spanish of the River Plate". *Orbis, 21,* 174–182.

Cena, R. M. (1978). *When is a phonological generalization psychologically real?* Bloomington, IN: Indiana University Linguistics Club.

Chandler, Steve (2002). "Skousen's analogical approach as an exemplar-based model of categorization". In Royal Skousen, Deryle Lonsdale, & Dilworth B. Parkinson (Eds.), *Analogical modeling: An exemplar-based approach to language* (pp. 51–105). Amsterdam: John Benjamins.

Chandler, S., & R. Skousen (1997). *Analogical modeling and the English past tense: A reply to Jaeger et al. 1996.* Manuscript, humanities.byu.edu/am/jaeger.html

Chomsky, Noam (1964). *Current issues in linguistic theory*. The Hague: Mouton.

Chomsky, Noam (1972). *Language and mind*. New York: Harcourt Brace Jovanovich.

Chomsky, Noam (1976). "On the biological basis of language capacities". In R. W. Rieber (Ed.), *The neuropsychology of language* (pp. 1–24). New York: Plenum Press.

Chomsky, Noam (1980a). *Rules and representations*. New York: Columbia University Press.

Chomsky, Noam (1980b). "Rules and representations". *The Behavioral and Brain Sciences, 3,* 1–15, 42–61.

Chomsky, Noam (1981). *Lectures on government and binding*. Dordrecht: Foris.

Chomsky, Noam (1982). *The generative enterprise*. Dordrecht: Foris.

Chomsky, Noam (1986). *Knowledge of language*. New York: Praeger.

Chomsky, Noam (1995). "Language and nature". *Mind, 104,* 1–61.

Chomsky Noam, & Morris Halle (1968). *The sound pattern of English*. New York: Harper and Row.

Clegg, Jens (1997). *Native Spanish speaker intuition in gender assignment in English loan words*. Manuscript, Brigham Young University.

Clements, George N., & Samuel Jay Keyser (1983). *CV Phonology: A generative theory of the syllable*. Cambridge, MA: Massachusetts Institute of Technology Press.

Colina, Sonia (2003). "Diminutives in Spanish: A morpho-phonological account". *Southwest Journal of Linguistics, 22,* 45–88.

Costa, Albert, & Nuria Sebastian-Gallés (1998). "Abstract phonological structure in language production: Evidence from Spanish". *Journal of Experimental Psychology: Learning, Memory, and Cognition, 24,* 886–903.

Coste, J., & A. Redondo (1965). *Syntaxe de l'espagnol moderne*. Paris: Société d'Edition d'Enseignement Superieur.

Cottrell, Garrison W., & Kim Plunkett (1991). "Learning the past tense in a recurrent network: Acquiring the mapping from meaning to sounds". *Proceeding of the Thirteenth Annual Conference of the Cognitive Science Society* (pp. 328–333). Hillsdale, NJ: Erlbaum.

Cowart, Wayne (1997). *Experimental syntax: Applying objective methods to sentence judgements*. Thousand Oaks, CA: Sage.

Crespo, Luis A. (1949). "To become". *Hispania, 32,* 210–212.

Cressey, William W. (1978). *Spanish phonology and morphology: A generative view*. Washington, DC: Georgetown University Press.

Crowhurst, Megan J. (1992). "Diminutives and augmentatives in Mexican Spanish: A prosodic analysis". *Phonology, 9,* 221–253.

Cutler, Anne (1979). "The psychological reality of word formation and lexical stress rules". In Eli Fischer-Jørgensen, Jørgen Rischel, & Nina Thorsen (Eds.), *Proceedings of The Ninth International Congress of Phonetic Sciences*, Vol. 2 (pp. 79–85). Copenhagen: University of Copenhagen.

Daelemans, Walter, Steven Gillis, & Gert Durieux (1994). "Skousen's analogical modeling algorithm: A comparison with lazy learning". In Jones (Ed.), *Proceedings of the International Conference on New Methods in Language Processing* (pp. 1–7). Manchester: UMIST.

Daelemans, Walter, Jakub Zavrel, Ko van der Sloot, & Antal van den Bosch (2001). TiMBL: Tilburg memory-based learner, version 4.1 reference guide: Induction of linguistic knowledge technical report 01–04. Tilburg: ILK Research Group, Tilburg University.

Danesi, Marcel (1982). "The description of Spanish /b, d, g/ revisited". *Hispania, 65,* 252–258.

Dannenbring, Gary L., & Kevin Briand (1982). "Semantic priming and the word repetition effect in a lexical decision task". *Canadian Journal of Psychology, 36,* 435–444.

Daugherty, Kim, & Mark S. Seidenberg (1992). "Rules or connections? The past tense revised". *Proceedings of the Fourteenth Annual Conference of the Cognitive Science Society* (pp. 259–264). Hillsdale, NJ: Erlbaum.

Daugherty, Kim, & Mark S. Seidenberg (1994). "Beyond rules and exceptions: A connectionist approach to inflectional morphology". In Susan. D. Lima, Roberta. L. Corrigan, & Gregory. K. Iverson (Eds.), *The reality of linguistic rules* (pp. 353–388). Amsterdam: John Benjamins.

Davies, Mark (In press). *Frequency dictionary of Spanish*. Routledge.

Den Os, Els, & Rene Kager (1986). "Extrametricality and stress in Spanish and Italian". *Lingua, 69,* 23–48.

Derwing, Bruce L. (1973). *Transformational grammar as a theory of language acquisition*. London: Cambridge University Press.

Derwing, Bruce L. (1979). "Psycholinguistic evidence and linguistic theory". In Gary D. Prideaux (Ed.), *Perspectives in experimental linguistics* (pp. 113–138). Amsterdam: John Benjamins.

Derwing, Bruce L. (1980). "Against autonomous linguistics". In Thomas A. Perry (Ed.), *Evidence and argumentation in linguistics* (pp. 163–189). Berlin: Walter de Gruyter.

Derwing, Bruce L. (1983). "On drawing the right conclusions in psycholinguistics: Some critical areas of control in experimental and analytical practice". In Amy Dahlstrom et al. (Eds.), *Proceedings of the Ninth Annual Meeting of the Berkeley Linguistics Society* (pp. 64–73). Berkeley: Berkeley Linguistics Society.

Derwing, B. L., G. D. Prideaux, & W. J. Baker (1980). "Experimental linguistics in historical perspective". In Gary D. Prideaux, Bruce L. Derwing, & William J. Baker (Eds.), *Experimental linguistics* (pp. 1–13). Ghent: E. Story-Scientia.

Derwing, Bruce L., & Royal Skousen (1989). "Real-time morphology: Symbolic rules or analogical networks?" In Kira Hall, Michael Meacham, & Richard Shapiro (Eds.), *Proceedings of the Fifteenth Annual Meeting of the Berkeley Linguistic Society* (pp. 48–62). Berkeley, CA: Berkeley Linguistics Society.

Derwing, Bruce L., & Royal Skousen (1994). "Productivity and the English past tense: Testing Skousen's analogy model". In Susan D. Lima, Roberta L. Corrigan, & Gregory K. Iverson (Eds.), *The reality of linguistic rules* (pp. 193–218). Amsterdam: John Benjamins.

Derwing, Bruce L., & William J. Baker (1977). "The psychological basis for morphological rules". In John Macnamara (Ed.), *Language learning and thought* (pp. 85–110). New York: Academic Press.

Derwing, Bruce L., Yeo Boom Yon, & Sook Whan Cho (1993). "The organization of the Korean syllable: Experimental evidence". In Patricia M. Clancy (Ed.), *Japanese/Korean linguistics* (pp. 223–238). Stanford, CA: Center for the Study of Language and Information.

Díaz-Campos, Manuel (2000). "The phonetic manifestation of secondary stress in Spanish". In Héctor Campos, Elena Herberger, Alfonso Morales-Front, & Thomas J. Walsh (Eds.), *Hispanic linguistics at the turn of the millenium* (pp. 49–65). Somerville, MA: Cascadilla Press.

Diver, William (1979). "Phonology as human behavior". In Doris Aaronson & Robert W. Rieber (Eds.), *Psycholinguistic research: Implications and applications* (pp. 161–198). Hillsdale, NJ: Erlbaum.

Docherty, Gerard, & Paul Foulkes (2000). "Speaker, speech, and knowledge of sounds". In Noel Burton-Roberts, Philip Carr, & Gerard Docherty (Eds.), *Phonological knowledge: Conceptual and empirical issues* (pp. 105–130). Oxford: Oxford University Press.

Domínguez, Alberto, Fernando Cuetos, & Juan Segui (1999). "The processing of grammatical gender and number in Spanish". *Journal of Psycholinguistic Research, 28,* 485–498.

Domínguez, Alberto, Fernando Cuetos, & Juan Segui (2000). "Morphological processing in word recognition: A review with particular reference to Spanish data". *Psicológica, 21,* 375–401.

Domínguez, Alberto, Fernando Cuetos, & Juan Segui (2001). "The time-course of inflexional morphological priming". *Linguistics, 40,* 235–229.

Domínguez, Alberto, Manuel de Vega, & Fernando Cuetos (1997). "Lexical inhibition from syllabic units in Spanish visual word recognition". *Language and Cognitive Processes, 12,* 401–422.

Doughty, Catherine J., & Michael H. Long (Eds.). (2003). *The handbook of second language acquisition*. Malden, MA: Blackwell.

Drews, Etta, & Pienie Zwisterlood (1995). "Morphological and orthographic similarity in visual word recognition". *Journal of Experimental Psychology: Human Perception and Performance, 21,* 1098–1116.

Eberenz, Rolf (1985). "Aproximación estructural a los verbos de cambio en iberorromance". *Linguistique comparée et typologie des langues romanes* (pp. 460–475). Aix en Provence: Université de Provence.

Eddington, David (1991). "A phonotactic explanation of the *vos* and final -r variety *vosotros* imperatives". *Anuario de Estudios Filológicos, 14,* 125–30. Reprinted as *Pasar y sentaros:* Accounting for the alternative vosotros imperative. Círculo de la Lingüística Aplicada a la Comunicación. 12. www.ucm.es/info/circulo/no12

Eddington, David (1992). "Word-medial epenthesis in Spanish: A lexical phonological approach". *Southwest Journal of Linguistics, 11,* 14–28.

Eddington, David (1994). "The psychological validity of phonological generalizations: A priming experiment in Spanish". *Journal of Quantitative Linguistics, 1,* 132–143.

Eddington, David (1996). "Diphthongization in Spanish derivational morphology: An empirical investigation". *Hispanic Linguistics, 8,* 1–35.

Eddington, David (1998). "Spanish diphthongization as a non-derivational phenomenon". *Rivista di Linguistica, 10,* 335–354.

Eddington, David (1999). "On 'becoming' in Spanish: A corpus analysis of verbs expressing change-of-state". *Southwest Journal of Linguistics, 18,* 23–46.

Eddington, David (2000a). "Analogy and the dual-route model of morphology". *Lingua, 110,* 281–298.

Eddington, David (2000b). "Spanish stress assignment within the analogical modeling of language". *Language, 76,* 92–109.

Eddington, David (2001a). "Spanish epenthesis: formal and performance perspectives". *Studies in the Linguistics Sciences, 31,* 33–53.

Eddington, David (2001b). "Surface analogy and spelling rules in English vowel alternations". *Southwest Journal of Linguistics, 20,* 85–105.

Eddington, David (2002a). "Disambiguating Spanish change-of-state verbs". *Hispania, 85,* 921–929.

Eddington, David (2002b). "Spanish diminutive formation without rules or constraints". *Linguistics, 40,* 395–419.

Eddington, David (2002c). "Spanish gender assignment in an analogical framework". *Journal of Quantitative Linguistics, 9,* 49–75.

Eddington, David (2004a). "A computational approach to resolving certain issues in Spanish stress placement". In Timothy L. Face (Ed,), *Laboratory approaches to Spanish phonology* (pp. 95–115). Berlin: Mouton de Gruyter.

Eddington, David (2004b). "Issues in modeling language processing analogically". *Lingua, 114,* 849–871.

Eddington, David (In press). "An analogical simulation of Spanish nominals in *-ión*". In Gary Libben (Ed.), *Festschrift for Bruce L. Derwing.*

Eddington, David, & Patricia Lestrade (2002). "Are plurals derived or stored?" In James. F. Lee, Kimberly L. Geeslin, & J. Clancy Clements (Eds.), *Structure, meaning, and acquisition in Spanish* (pp. 269–284). Somerville, MA: Cascadilla Press.

Ellis, Nick C. (2002). "Frequency effects in language processing: A review with implications for theories of implicit and explicit language acquisition". *Studies in Second Language Acquisition, 24,* 143–188.

Elordieta, Gorka, & María M. Carreiras (1996). "An optimality theoretic analysis of Spanish diminutives". In Lise M. Dobrin, Kora Singer, & Lisa McNair (Eds.), *Proceedings of the Chicago Linguistic Society's 32nd meeting.* Chicago: Chicago Linguistic Society.

Embleton, Sheila M. (1986). *Statistics in historical linguistics*. Bochum: Brockmeyer.

Enríquez, Emilia V., Celia Casado, & Andrés Santos (1989). "La percepción del acento en español". *Lingüística Española Actual, 11,* 241–269.

Face, Timothy L. (2000). "The role of syllable weight in the perception of Spanish stress". In Héctor Campos, Elena Herberger, Alfonso Morales-Front, & Thomas J. Walsh (Eds.), *Hispanic linguistics at the turn of the millennium* (pp. 1–13). Somerville, MA: Cascadilla Press.

Face, Timothy L. (2003). "Where is stress? Synchronic and diachronic Spanish evidence. Theory, practice, and acquisition". In Paula Kempchinsky & Carlos-Eduardo Piñeros (Eds.), *Papers from the 6th Hispanic Linguistics Symposium and the 5th Conference on the Acquisition of Spanish and Portuguese* (pp. 21–39). Somerville, MA: Cascadilla Press.

Face, Timothy L. (2004). "Perceiving what isn't there: Non-acoustic cues for perceiving Spanish stress". In Timothy L. Face (Ed.), *Laboratory Approaches to Spanish Phonolgy* (pp. 117–141). Berlin: Mouton de Gruyter.

Face, Timothy, & Scott Alvord (2003). *Descriptive adequacy vs. psychological reality: The case of two restrictions on Spanish stress placement*. Manuscript, The University of Minnesota.

Feldman, Laurie B. (1992). "Morphological relationships revealed through the repetition priming task". In Michael Noonan, Pamela Downing, & Susan Lima (Eds.), *Linguistics and literacy* (pp. 239–254). Amsterdam: John Benjamins.

Feldman, Laurie, B. (1994). "Beyond orthography and phonology: Differences between inflections and derivations". *Journal of Memory and Language, 33,* 442–470.

Feldman, Laurie B., & J. Moskovljevic (1987). "Repetition priming is not purely episodic in origin". *Journal of Experimental Psychology: Learning, Memory, and Cognition, 13,* 573–581.

Fente, R. (1970). "Sobre los verbos de cambio o 'devenir' ". *Filología Moderna, 38,* 157–172.

Ferrand, Ludovic, & Juan Segui (1996). "Masked priming of word and picture naming: The role of syllabic units". *Journal of Memory and Language, 35,* 708–723.

Ferrand, Ludovic, & Juan Segui (1998). "The syllable's role in speech production: Are syllables chunks, schemas, or both?" *Psychonomic Bulletin and Review, 5,* 253–258.

Figueroa, Neysa (2000). "An acoustic and perceptual study of vowels preceding deleted post-nuclear /s/ in Puerto Rican Spanish". In Héctor Campos, Elena Herberger, Alfonso Morales-Front, & Thomas J. Walsh (Eds.), *Hispanic linguistics at the turn of the millennium* (pp. 66–79). Sommerville, MA: Cascadilla Press.

Fischler, Ira (1977a). "Semantic facilitation without expectancy in a lexical decision task". *Journal of Experimental Psychology: Human Perception and Performance, 3,* 18–26.

Fischler, Ira (1977b). "Semantic facilitation without association in a lexical decision task". *Memory and Cognition, 5,* 335–339.

Forster, Kenneth I. (1987). "Form-priming with masked primes: The best match hypothesis". In Max Coltheart (Ed.), *Attention and performance 12: The psychology of reading* (pp. 127–146). Hove: Erlbaum.

Forster, K. I., Davis, C., Schnoknecht, C., & Carter, R. (1987). "Masked priming with graphemically related forms: Repetition or partial activation?" *Quarterly Journal of Experimental Psychology, 39A,* 211–251.

Fougeron, Cécile, & Donca Steriade (1997). "Does deletion of French schwa lead to neutralization of lexical distinctions?" In G. Kokkinakis, N. Fakotakis, & E. Dermatas (Eds.), *Proceedings of the 5th European Conference on Speech Communication and Technology*, Vol. 7 (pp. 943–937). Rhodes: University of Patras.

Fought, John (1973). "On the adequacy of recent phonological theories and practices". *Annual Review of Anthropology, 2,* 153–164.

Fowler, Carol A., Shirley E. Napps, & Laurie Feldman (1985). "Relations among regular and irregular morphologically related words in the lexicon as revealed by repetition priming". *Memory and Cognition, 13,* 241–255.

Fowler, Carol A., George Wolford, Ronald Slade, & Louis Tassinary (1981). "Lexical access with and without awareness". *Journal of Experimental Psychology: General, 110,* 341–362.

Frisch, Stefan (1996). *Similarity and frequency in phonology*. PhD Dissertation, Northwestern University.

Frisch, Stefan A., Nathan R. Large, Bushra Zawaydeh, & David B. Pisoni (2001). "Emergent phonotactic generalizations in English and Arabic". In Joan Bybee & Paul Hooper (Eds.), *Frequency and the emergence of linguistic structure* (pp. 159–179). Amsterdam: John Benjamins.

Fromkin, Victoria A. (1975). "When does a test test a hypothesis, or, what counts as evidence?" In David Cohen & Jessica R. Wirth (Eds.), *Testing linguistic hypotheses* (pp. 43–64). Washington, DC: Hemisphere.

Fromkin, Victoria A. (1980). "The psychological reality of phonological descriptions". In Eli Fischer-Jørgensen & Nina Thorsen (Eds.), *Proceedings of the Ninth International Congress of Phonetic Sciences*, Vol. 3 (pp. 195–202). Copenhagen: University of Copenhagen.

García-Albea, José E., Rosa M. Sánchez-Casas, & José M. Igoa (1998). "The contribution of word form and meaning to language processing in Spanish: Some evidence from monolingual and bilingual studies". *Syntax and Semantics, 31,* 183–209.

García-Bellido, Paloma (1986). "Lexical diphthongization and high-mid alternations in Spanish: An autosegmental account". *Linguistic Analysis, 16,* 61–92.

Geeraerts, Dirk (1989). "Introduction: Prospects and problems of prototype theory". *Linguistics, 27,* 587–612.

Gernsbacher, Morton Ann (1994). *Handbook of psycholinguistics*. San Diego: Academic Press.

Goldinger, Stephen D. (1997). "Words and voices: Perception and production in an episodic lexicon". In Keith Johnson & John W. Mullennix (Eds.), *Talker variability in speech processing* (pp. 33–65). San Diego, CA: Academic Press.

Goldsmith, John (1993). "Harmonic phonology". In John Goldsmith (Ed.), *The last phonological rule* (pp. 21–60). Chicago: University of Chicago Press.

Goyvaerts, Didier L. (1978). "From triumphant euphoria to bitter disappointment?" *Communication and Cognition, 11*, 287–307.

Grainger, Jonathan, Pascale Cole, & Juan Segui (1991). "Masked morphological priming in visual words recognition". *Journal of Memory and Language, 30,* 370–384.

Greene, Judith (1974). *Psycholinguistics*. New York: Viking Press.

Guirao, Migeulina, & María A. García Jurado (1990). "Frequency of occurrence of phonemes in American Spanish". *Revue Québecoise de Linguistique, 19,* 135–150.
Hacker, P. M. S (1990). "Chomsky's problems". *Language and Communication, 10,* 127–148.
Hale, Mark, & Charles Reiss (2000). "Phonology as cognition". In Noel Burton-Roberts, Philip Carr, & Gerard Docherty (Eds.), *Phonological knowledge: Conceptual and empirical issues* (pp. 161–184). Oxford: Oxford University Press.
Hall, Robert A. Jr. (1987). *Linguistics and pseudo-linguistics*. Amsterdam: John Benjamins; Reprinted from 1981/82. Review of Newmeyer "Linguistic theory in America". *Forum Linguisticum, 6*, 177–88, and 1973, *Journal of English Linguistics, 7,* 21–42.
Halle, Morris (1973). "Prolegomena to a theory of word formation". *Linguistic Inquiry, 4,* 3–16.
Halle, Morris (1977). "Tenseness, vowel shift, and the phonology of the back vowels in Modern English". *Linguistic Inquiry, 8,* 611–625.
Halle, Morris, James W. Harris, & Jean-Roger Vergnaud (1991). "A reexamination of the stress erasure convention and Spanish stress". *Linguistic Inquiry, 22,* 141–159.
Halle, Morris, & K. P. Mohanan (1985). "Segmental phonology of Modern English". *Linguistic Inquiry, 16,* 57–116.
Hammond, Robert M. (1978). "An experimental verification of the phonemic status of open and closed vowels in Caribbean Spanish". In Humberto López Morales (Ed.), *Corrientes actuales en la dialectología del caribe hispánico* (pp. 93–125). Río Piedras: Editorial Universitaria de la Universidad de Puerto Rico.
Harmon, Gilbert (1980). "Two quibbles about analyticity and psychological reality". *Behavioral and Brain Sciences, 3,* 21–22.
Harris, James W. (1969). *Spanish phonology*. Cambridge, MA: Massachusetts Institute of Technology Press.
Harris, James W. (1977). "Remarks on diphthongization in Spanish". *Lingua, 41,* 261–305.
Harris, James W. (1978). "Two theories of nonautomatic morphophonological alternations: Evidence from Spanish". *Language, 54,* 41–60.
Harris, James W. (1983). Syllable structure and stress in Spanish: A nonlinear analysis. Cambridge MA: Massachusetts Institute of Technology Press.
Harris, James W. (1985a). "Spanish diphthongization and stress: A paradox resolved". *Phonology Yearbook, 2,* 31–45.
Harris, James W. (1985b). "Spanish word markers". In Frank H. Nuessel Jr. (Ed.), *Current issues in Hispanic phonology and morphology*. Bloomington, IN: Indiana Linguistics Club.
Harris, James W. (1987). "Epenthesis processes in Spanish". In Carol Neidle & Rafael A. Núñez-Cedeño (Eds.), *Studies in Romance languages* (pp. 107–122). Dordrecht: Foris.
Harris, James W. (1989a). "How different is verb stress in Spanish". *Probus, 1,* 241–258.
Harris, James W. (1989b). "Our present understanding of Spanish syllable structure". In Peter C. Bjarkman & Robert M. Hammond (Eds.), *American Spanish: Theoretical and applied perspectives* (pp. 151–169). Washington, DC: Georgetown University Press.
Harris, James W. (1991a). "The exponence of gender in Spanish". *Linguistic Inquiry, 22,* 27–62.

Harris, James W. (1991b). "With respect to metrical constituents in Spanish". In Héctor Campos & Fernando Martínez-Gil (Eds.), *Current studies in Spanish linguistics* (pp. 447–473). Washington, DC: Georgetown University Press.

Harris, James W. (1992). *Spanish stress: The extrametricality issue*. Bloomington, IN: Indiana University Linguistics Club.

Harris, James W. (1993). "Integrity of prosodic constituents and the domain of syllabification rules in Spanish and Catalan". In Ken Hale & Samuel J. Keyser (Eds.), *The view from building 20: Essays in linguistics in honor of Sylvain Bromberger* (pp. 177–193). Cambridge, MA: Massachusetts Institute of Technology Press.

Harris, James (1994). "The OCP, prosodic morphology and Sonoran Spanish diminutives: A reply to Crowhurst". *Phonology, 11,* 179–190.

Hartman, Steven Lee (1981). "A universal alphabet for experiments in comparative phonology". *Computers and the Humanities, 15,* 75–82.

Hatch, Evelyn, & Hossein Farhady (1982). *Research design and statistics for applied linguistics*. New York: Newbury House.

Hay, Jennifer, Janet Pierrehumbert, & Mary Beckman (2003). "Speech perception, well-formedness, and the statistics of the lexicon". In John Local, Richard Ogden, & Rosalind Temple (Eds.), *Papers in laboratory phonology VI* (pp. 58–74). Cambridge: Cambridge University Press.

Henderson, Leslie, Julie Wallis, & Denise Knight (1984). "Morpheme structure and lexical access". In H. Bouma & D. G. Bouwhuis (Eds.), *Attention and Performance X* (pp. 211–226). Hillsdale, NJ: Lawrence Erlbaum.

Higginbotham, James (1991). "Remarks on the metaphysics of linguistics". *Linguistics and Philosophy, 14,* 555–566.

Hillinger, Michael L. (1980). "Priming effects with phonemically similar words: The encoding-bias hypothesis reconsidered". *Memory and Cognition, 8,* 115–123.

Hochberg, Judith (1988). "Learning Spanish stress: Developmental and theoretical perspectives". *Language, 64,* 683–706.

Honsa, Vladimir (1965). "The phonemic systems of Argentinean Spanish". *Hispania, 48,* 275–283.

Hooper, Joan B. (1976). *An introduction to natural generative phonology*. New York: Academic Press.

Hooper, Joan B. (1978). "Rule morphologization in Natural Generative Phonology". In Robert A. Fox, Michael W. La Galy, & Bill J. Darden (Eds.), *Papers from the parasession on Natural Phonology* (pp. 160–170). Chicago: Chicago Linguistic Society.

Hooper, Joan B. (1981). "The empirical determination of phonological representations". In Terry Meyers, John Laver, & John M. Anderson (Eds.), *The cognitive representation of speech* (pp. 347–357). Amsterdam: North Holland.

Hooper, Joan B., & Tracy Terrell (1976). "Stress assignment in Spanish: A natural generative analysis". *Glossa, 10,* 64–110.

Hualde, José Ignacio (1991). "On Spanish syllabification". In Héctor Campos & Fernando Martínez-Gil (Eds.), *Current studies in Spanish linguistics* (pp. 475–493). Washington, DC: Georgetown University Press.

Hualde, José Ignacio (1999). "Patterns in the lexicon: Hiatus with unstressed high vowels in Spanish". In Javier Gutiérrez Rexach & Fernando Martínez-Gil (Eds.), *Advances in Hispanic linguistics*, Vol. 1 (pp. 182–197). Somerville, MA: Cascadilla Press.

Hualde, José Ignacio, & Mónica Prieto (2002). "On the diphthong/hiatus constrast in Spanish: Some experimental results". *Linguistics, 40,* 217–234.

Hume, Elizabeth (2001). "Metathesis: Formal and functional considerations". In Elizabeth Hume, Norval Smith, & Jeroen van de Weijer (Eds.), *Surface syllable structure and segment sequencing* (pp. 1–25). Leiden: Holland Institute of Generative Linguistics.

Hume, Elizabeth (2004). "The indeterminacy/attestation model of metathesis". *Language, 80,* 203–237.

Hutchinson, Larry G. (1974). "Grammar as theory". In David Cohen (Ed.), *Explaining linguistic phenomenon* (pp. 43–73). New York: John Wiley.

Itkonen, Esa (1976). *Linguistics and empiricalness: Answers to criticisms*. Helsinki: University of Helsinki.

Itkonen, Esa (1978a). *Grammatical theory and metascience*. Amsterdam: John Benjamins.

Itkonen, Esa (1978b). "Linguistics: Nonempirical and empirical". In Wolfgang U. Dressler & Wolfgang Meid (Eds.), *Proceedings of the 12th International Congress of Linguists* (pp. 157–158). Innsbruck: University Innsbruck.

Itkonen, Esa (1983). *Causality in linguistic theory*. London: Croom Helm.

Jackendoff, Ray (1975). "Morphological and semantic regularities in the lexicon". *Language, 51,* 639–671.

Jaeggli, Osvaldo A. (1980). "Spanish diminutives". In Frank Nuessel Jr. (Ed.), *Contemporary studies in Romance languages: Eighth Annual Linguistics Symposium on Romance Languages* (pp. 145–158). Bloomington, IN: Indiana University Linguistics Club.

Jeager, Jeri J. (1980). "Testing the psychological reality of phonemes". *Language and Speech, 23,* 233–253.

Jeager, Jeri J. (1984). "Assessing the psychological status of the vowel shift rule". *Journal of Psycholinguistic Research, 13,* 13–36.

Jeager, Jeri J. (1986). "On the acquisition of abstract representations for English vowels". *Phonology Yearbook, 3,* 71–97.

Jeager, Jeri J., Alan H. Lockwood, David L. Kemmerer, Robert D. Van Valin Jr., Brian W. Murphy, & Hanif G. Khalak (1996). "A positron emission tomographic study of regular and irregular verb morphology in English". *Language, 72,* 451–497.

Jiménez-Gonález, Juan E., & María del Rosario Ortiz-González (1994). "Phonological awareness in learning literacy". *Intellectica, 1,* 155–181.

Juilland, Alphonse, & Eugenio Chang-Rodríguez (1964). *Frequency dictionary of Spanish words*. The Hague: Mouton.

Kac, Michael (1974). "Autonomous linguistics and psycholinguistics". *Minnesota Working Papers in Linguistics and Philosophy of Language, 2,* 42–47.

Kac, Michael (1978). *Corepresentation of grammatical structure.* Minneapolis and London: University of Minnesota Press and Croom Helm.

Kac, Michael (1980). "In defense of autonomous linguistics". *Linguistics, 50,* 243–248.

Kager, René (1999). *Optimality Theory*. Cambridge: Cambridge University Press.

Katz, Jerrold J. (1981). *Language and other abstract objects*. Totowa, NJ: Rowman and Littlefield.

Katz, Jerrold J. (1985). "An outline of platonist grammar". In Jerrold J. Katz (Ed.), *The philosophy of linguistics* (pp. 172–203). Oxford: Oxford University Press.

Katz, Jerrold J., & Paul M. Postal (1991). "Realism vs. conceptualism in linguistics". *Linguistics and Philosophy, 14,* 515–554.

Kelm, Orlando R. (1995). "Acoustic measurement of Spanish and English pitch contours: Native and non-native speakers". *Hispanic Linguistics, 6/7,* 435–448.

Kempley, S. T., & John Morton (1982). "The effects of priming with regularly and irregularly related words in auditory word recognition". *British Journal of Psychology, 73,* 441–445.

Kenstowicz, M. (1996). "'Base-identity constraints and uniform exponence: Alternatives to cyclicity". In J. Durand & B. Laks (Eds.), *Current trends in phonology: Models and methods* (pp. 363–393). Salford: European Studies Research Institute, University of Salford Publications.

Kenstowicz, Michael, & Charles Kisseberth (1979). *Generative phonology*. San Diego, CA: Academic Press.

Kinnear, Paul R., & Colin D. Gray (1999). *SPSS for Windows made simple*. Hove: Psychology Press.

Kiparsky, Paul (1968). "Linguistic universals and linguistic change". In Emmon Bach & Richard Harms (Eds.), *Universals in linguistic theory* (pp. 171–204). New York: Holt, Rinehart, and Winston.

Kiparsky, Paul (1973). "Elsewhere in phonology". In Stephen Anderson & Paul Kiparsky (Eds.), *A Festschrift for Morris Halle* (pp. 93–106). New York: Holt, Rinehart, and Winston.

Kiparsky, Paul (1975). "What are phonological theories about?" In David Cohen & Jessica R. Wirth (Eds.), *Testing linguistic hypotheses* (pp. 187–209). Washington, DC: Hemisphere.

Kiparsky, Paul, & Lise Menn (1977)." On the acquisition of phonology". In John Macnamara (Ed.), *Language learning and thought* (pp. 47–78). New York: Academic Press.

Klein, Philip W. (1983). "Spanish gender morphology". *Papers in Romance, 5,* 57–64.

Klein, Philip W. (1989). "Spanish 'gender' vowels and lexical representation". *Hispanic Linguistics, 3,* 147–162.

Kolers, Paul A., & Henry L. Roediger (1984). "Procedures of mind". *Journal of Verbal Learning and Verbal Behavior, 23,* 425–449.

Labov, William (1994). *Principles of linguistic change*. Oxford: Blackwell.

Ladefoged, Peter (2003). *Phonetic data analysis: An introduction to fieldwork and instrumental techniques*. Malden, MA: Blackwell.

Lakoff, George (1993). "Cognitive phonology". In John Goldsmith (Ed.), *The last phonological rule* (pp. 117–145). Chicago: University of Chicago Press.

Lass, Norman J. (Ed.). (1996). *Principles of experimental phonetics*. Saint Louis: Mosby.

Lass, Roger (1976a). *English phonology and phonological theory*. London: Cambridge University Press.

Lass, Roger (1976b). "On generative taxonomy and whether formalisms 'explain.' " *Studia Linguistica, 30,* 139–154.

Lass, Roger (1984). *Phonology: An introduction to basic concepts*. Cambridge: Cambridge University Press.

Lathrop, Thomas A. (1984). *Curso de gramática histórica española*, trans. by Juan Gutiérrez & Ana Blas. Barcelona: Ariel.

Laudanna, Alessandro, William Badecker, & Alfonso Caramazza (1992). "Priming homographic stems". *Journal of Memory and Language, 28,* 531–546.

Levine, Robert D., & Paul M. Postal (2004). "A corrupted linguistics". In Peter Collier & David Horowitz (Eds.), *The anti-Chomsky reader* (pp. 203–231). San Francisco: Encounter Books.

Linell, Per (1976). "Is linguistics an empirical science?" *Studia Linguistica, 30,* 77–94.

Lipski, John M. (1994). *Latin American Spanish*. New York: Longman.

Lipski, John M. (1997). "Spanish word stress: The interaction of moras and minimality". In Fernando Martínez-Gil & Alfonso Morales-Front (Eds.), *Issues in the phonology and morphology of the major Iberian languages* (pp. 559–593). Washington, DC: Georgetown University Press.

Llisterri, Joaquim, & Dolors Poch (1986). "Análisis acústico del timbre vocálico en las realizaciones normativas del plural en andaluz oriental". Paper presented at the XVI Simposio de la Sociedad Española de Lingüística.

Lloyd, Paul (1987). *From Latin to Spanish*. Philadelphia: American Philosophical Society.

Logan, Gordon D. (1988). "Toward and instance theory of automatization". *Psychological Review, 95,* 492–527.

López Morales, Humberto (1989). *Sociolingüística*. Madrid: Gredos.

Love, Nigel (1992). Review of "Linguistic realities: An autonomous metatheory for the generative enterprise", by Philip Carr. *Language and Communication, 12,* 79–92.

Lukatela, Georgije, Claudia Carello, & M. T. Turvey (1987). "Lexical representation of regular and irregular inflected nouns". *Language and Cognitive Processes, 2,* 1–17.

Lukatela, Georgije, B. Gligorijevic, Aleksandar Kostic, & M. T. Turvey (1980). "Representation of inflected nouns in the internal lexicon". *Memory and Cognition, 8,* 415–423.

Lust, Barbara C., & Claire Foley (Eds.). (2004). *First language acquisition: The essential readings*. Malden, MA: Blackwell.

MacKay, Donald G. (1982). "The problems of flexibility, fluency, and speed-accuracy trade-off in skilled behavior". *Psychological Review, 89,* 60–94.

Manelis, Leon, & David A. Tharp (1977). "The processing of affixed words". *Memory and Cognition, 5,* 690–695.

Marcel, Anthony J. (1983). "Conscious and unconscious perception: Experiments on visual masking and word recognition". *Cognitive Psychology, 15,* 197–237.

Marcos Marín, Francisco, director (1992). Corpus oral de referencia de la lengua española contemporánea. Textual corpus, Universidad Autónoma de Madrid.

Marcus, Gary F., Ursula Brinkmann, Harald Clahsen, Richard Wiese, Andreas Woest, & Steven Pinker (1993). "German inflection: The exception that proves the rule". *Occasional Paper No. 47*. Center for Cognitive Science, Massachusetts Institute of Technology.

Marslen-Wilson, William, Lorraine Komisarjevsky Tyler, Rachelle Waksler, & Lianne Older (1994). "Morphology and meaning in the English mental lexicon". *Psychological Review, 101,* 3–33.

Martin, Randi C., & Cary R. Jensen (1988). "Phonological priming in the lexical decision task: A failure to replicate". *Memory and Cognition, 16,* 505–521.

Martínez Celdrán, Eugenio (1991). "Duración y tensión en las oclusivas no iniciales del español: Un estudio perceptivo". *Revista Argentina de Lingüística, 7*, 51–71.

Martínez Melgar, Antonia (1986). Estudio experimental sobre un muestreo de vocalismo andaluz. *Estudios de Fonética Experimental, 2,* 195–248.

Martínez Melgar, Antonia (1994). "El vocalismo del andaluz oriental". *Estudios de fonética experimental, 6,* 13–64.

Matthews, Robert J. (1991). "Psychological reality of grammars". In Asa Kasher (Ed.), *The chomskyan turn* (pp. 182–199). Cambridge, MA: Basil Blackwell.

McCarthy, J. (1995). *Extensions of faithfulness: Rotuman revisited.* Manuscript, University of Massachusetts, Amherst. ROA 110, ruccs.rutgers.edu/roa.html

McCarthy, J., & A. Prince (1994a). *The emergence of the unmarked.* Manuscript, University of Massachusetts, Amherst. ROA 13, ruccs.rutgers.edu/roa.html

McCarthy, J., & A. Prince (1994b). *An overview of prosodic morphology.* Manuscript, University of Massachusetts, Amherst. ROA 59, ruccs.rutgers.edu/roa.html

McClelland, James L. (1988). "Connectionist models and psychological evidence". *Journal of Memory and Language, 27,* 107–123.

Medin, Douglas L., & Marguerite M. Schaffer (1978). "Context theory of classification learning". *Psychological Review, 85,* 207–238.

Milgram, Stanley (1974). *Obedience to authority: An experimental view.* New York: Harper and Row.

Miranda, Inés (1999). *An optimality theoretic analysis of Nicaraguan Spanish diminutivization.* PhD Dissertation, University of Washington.

Mohanan, K. P. (1986). *The theory of lexical phonology.* Dordrecht: D. Reidel.

Mohanan, K. P. (1997). *Posting to Linguist List discussion group*, Vol. 8–575. April 23, 1997. www.linguistlist.org

Monsell, Stephen (1985). "Repetition and the lexicon". In Andrew W. Ellis (Ed.), *Progress in the psychology of language*, Vol. 2 (pp. 147–195). London: Erlbaum.

Morgan, Terrell Alan (1984). *Consonant-glide-vowel alternations in Spanish: A case study in syllabic and lexical phonology.* PhD Dissertation, University of Texas at Austin.

Morin, Regina (2002). "Coronal and velar softening in Spanish: Theoretical, historical and empirical evidence of lexicalization". *Southwest Journal of Lingusitics, 21,* 137–163.

Morin, Yves-Charles (1988). "Morphological conditioning in phonologically transparent processes: Evidence from the evolution of vowel reduction in Vinzelles Occitan". *Canadian Journal of Linguistics, 33,* 431–442.

Moskowitz, Breyne Arlene (1973). In Timothy E. Moore (Ed.), *Cognitive development and the acquisition of language* (pp. 223–260). New York: Academic Press.

Muller, Bethany B. A. (2002). *Effects of collocational frequency and transitional probabilities on language processing: Collocations involving Spanish 'de.'* Manuscript, University of New Mexico.

Murrell, Graham A., & John Morton (1974). "Word recognition and morphemic structure". *Journal of Experimental Psychology, 102,* 963–968.

Myerson, Rosemarie Farkas (1976). "Children's knowledge of selected aspects of 'sound pattern of English.' " In Robin N. Campbell & Philip T. Smith (Eds.), *Recent advances in the psychology of language* (pp. 377–420). New York: Plenum Press.

Napps, Shirley E. (1989). "Morphemic relationships in the lexicon: Are they distinct from semantic and formal relationships?" *Memory and Cognition, 17,* 729–739.

Napps, Shirley E., & Carol A. Fowler (1987). "Formal relationships among words and the organization of the mental lexicon". *Journal of Psycholinguistic Research, 16,* 257–272.
Natalicio, Diana (1983). "Native speaker intuitions as a basis for determining noun gender rules in Spanish". *Southwest Journal of Linguistics, 6,* 49–55.
Navarro Tomás, Tomás (1957). *Manual de pronunciación española*. Madrid: Concejo Superior de Investigaciones Científicas.
Navarro Tomás, Tomás (1966). *El español en Puerto Rico*. Río Piedras: Editorial Universitaria de la Universidad de Puerto Rico.
Nosofsky, Robert M. (1988). "Exemplar-based accounts of relations between classification, recognition, and typicality". *Journal of Experimental Psychology: Learning, Memory, and Cognition, 14,* 700–708.
Nosofsky, Robert M. (1990). "Relations between exemplar similarity and likelihood models of classification". *Journal of Mathematical Psychology. 34,* 393–418.
Nosofsky, Robert M., & Thomas J. Palmeri (1997). "An exemplar-based random walk model of speeded classification". *Psychological Review, 104,* 266–300.
Núñez-Cedeño, Rafael A. (1985). "On the three-tiered syllabic theory and its implications for Spanish". In Larry D. King & Catherine A. Maley (Eds.), *Selected papers from the XIIIth Linguistic Symposium on Romance Languages* (pp. 261–285). Amsterdam: John Benjamins.
Núñez-Cedeño, Rafael A. (1986). "Teoría de la organización silábica e implicaciones para el análisis del español caribeño". In Rafael A. Núñez-Cedeño, Iraset Páez Urdaneta, & Jorge M. Guitart (Eds.), *Estudios sobre la fonología del español del Caribe* (pp. 75–94). Caracas: La Casa de Bello.
Núñez-Cedeño, Rafael A. (1993). *Morfología de la sufijación española*. Santo Domingo: Universidad Nacional Pedro Henríquez Ureña.
Núñez-Cedeño, Rafael A., & Alfonso Morales-Front (1999). *Fonología generativa contemporánea de la lengua española*. Washington, DC: Georgetown University Press.
Oakes, Michael P. (1998). *Statistics for corpus linguistics*. Edinburgh: Edinburgh University Press.
Ohala, John J. (1974). "Experimental historical phonology". In J. M. Anderson & C. Jones (Eds.), *Historical linguistics II* (pp. 353–389). Amsterdam: North Holland.
Ohala, John J. (1990). "There is no interface between phonology and phonetics: A personal view". *Journal of Phonetics, 18,* 153–171.
Ohala, John J. (2003). "Phonology based on testable, real-world, not exclusively linguistic, primitives". *Paper presented at the annual meeting of the Chicago Linguistic Society*.
Olshewsky, Thomas (1985). "Chomsky's realism". *Proteus, 2,* 19–25.
Otero, Carlos-Peregrín (1971). *Evolución y revolución en Romance*. Barcelona: Seix Barral.
Pallier, Christophe, Núria Sebastián-Gallés, Teodora Felguera, Anne Christophe, & Jacques Mehler (1993). "Attentional allocation within the syllabic structure of spoken words". *Journal of Memory and Language, 32,* 373–389.
Palmeri, Thomas J., Stephen D. Goldinger, & David B. Pisoni (1993). "Episodic encoding of voice attributes and recognition memory for spoken words". *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19*, 309–328.
Paolillo, John C. (2002). *Analyzing linguistic variation: Statistical models and methods*. Stanford: Center for the Study of Language and Information.

Parkinson, Dilworth (2002). "Running the Perl/C version of the analogical modeling program". In Royal Skousen, Deryle Lonsdale, & Dilworth B. Parkinson (Eds.), *Analogical modeling: An exemplar-based approach to language* (pp. 365–383). Amsterdam: John Benjamins.

Pavkov, Thomas W. (2000). *Ready, set, go!: A student guide to SPSS 9.0 for Windows*. Mountain View, CA: Mayfield Publications.

Pensado, Carmen (1997). "On the Spanish depalatalization of /ɲ/ and /ʎ/ in rhymes". In Fernando Martínez-Gil & Alfonso Morales-Front (Eds.), *Issues in the phonology and morphology of the major Iberian languages* (pp. 595–618). Washington, DC: Georgetown University Press.

Pérez, Hernán Emilio (1998). "Incidencia de dos rasgos acústicos en la percepción de la correlación /p-t-k/ vs. /b-d-g/". *Revista de Lingüística Teórica y Aplicada, 36,* 113–125.

Pérez-Pereira, Miguel (1991). "The acquisition of gender: What Spanish children tell us". *Journal of Child Language, 18,* 571–590.

Pierrehumbert, Janet (2001). "Exemplar dynamics: Word frequency, lenition, and contrast". In Joan Bybee & Paul Hooper (Eds.), *Frequency and the emergence of linguistic structure* (pp. 137–158). Amsterdam: John Benjamins.

Pierrehumbert, Janet (2002). *An unnatural process*. Manuscript, Northwestern University.

Pierrehumbert, Janet, Mary E. Beckman, & D. R. Ladd (2000). "Conceptual foundations of phonology as a laboratory science". In Noel Burton-Roberts, Philip Carr, & Gerard Docherty (Eds.), *Phonological knowledge: Conceptual and empirical issues* (pp. 273–303). Oxford: Oxford University Press.

Pinker, Steven (1991). "Rules of language". *Science, 253,* 530–534.

Pinker, Steven, & Alan Prince (1988). "On language and connectionism: Analysis of a parallel distributed processing model of language acquisition". *Cognition, 28,* 73–193.

Pinker, Steven, & Alan Prince (1994). "Regular and irregular morphology and the psychological status of rules of grammar". In Susan D. Lima, Roberta L. Corrigan, & Gregory K. Iverson (Eds.), *The reality of linguistic rules* (pp. 321–351). Amsterdam: John Benjamins.

Pisoni, David (1997). "Some thoughts on 'normalization' in speech perception". In Keith Johnson & John W. Mullennix (Eds.), *Talker variability in speech processing* (pp. 9–32). San Diego, CA: Academic Press.

Plunkett, Kim, & Virginia Marchman (1991). "U-shaped learning and frequency effects in a multi-layered perception". *Cognition, 38,* 43–102.

Poplack, Shana, Alicia Pousada, & David Sankoff (1982). "Un estudio comparativo de la asignación de género a préstamos nominales". In Orlando Alba (Eds.), *El español de caribe* (pp. 239–269). Santiago, Dominican Republic: Universidad Católica Madre y Maestro.

Popper, Karl R. (1968). *The logic of scientific discovery*. 2nd ed. New York: Harper and Row; original English edition 1959, London: Hutchinson.

Prado, Marcial (1982). "El género en español y la teoría de la marcadez". *Hispania, 65,* 258–266.

Prasada, Sandeep, & Steven Pinker (1993). "Generalisation of regular and irregular morphological patterns". *Language and Cognitive Processes, 8,* 1–56.

Prideaux, Gary (1971). "On the notion 'linguistically significant generalization.'" *Lingua, 26,* 337–347.
Prideaux, Gary (1980). "Rejection of autonomous linguistics". *Linguistics, 50,* 245–247.
Prideaux, Gary D. (1984). *Psycholinguistics: The experimental study of language*. London: Croom Helm.
Prieto, Pilar (1992). "Morphophonology of the Spanish diminutive formation: A case for prosodic sensitivity". *Hispanic Linguistics, 5,* 169–205.
Prieto, Pilar, & Jan van Santen (1996). "Secondary stress in Spanish: Some experimental evidence". In Claudia Parodi, Carlos Quícoli, Mario Saltarelli, & María Luisa Zubizarreta (Eds.), *Aspects of Romance linguistics: Selected papers from the Linguistics Symposium on Romance Languages* (pp. 337–356). Washington, DC: Georgetown University Press.
Quilis, Antonio (1971). "Caracterización fonética del acento español". *Travaux de Linguistique et de Littérature, 9,* 53–72.
Quilis, Antonio (1981). *Fonética acústica de la lengua española*. Madrid: Gredos.
Quilis, Antonio (1993). *Tratado de fonología y fonética españolas*. Madrid: Gredos.
Rainer, Franz (2003). "Studying restrictions on patterns of word formation by means of the Internet". *Italian Journal of Linguistics, 15,* 131–140.
Rand, David, & David Sankoff (1990). *GoldVarb: A variable rule application for MacIntosh*. Montréal: Centre de recherches mathématiques, Université de Montréal.
Rapp, Brenda C. (1992). "The nature of sublexical orthographic organization: The bigram trough hypothesis examined". *Journal of Memory and Language, 31,* 33–53.
Rastle, Kathleen, Matt H. Davis, William D. Marslen-Wilson, & Lorraine K. Tyler (2000). "Morphological and semantic effects in visual word recognition: A time-course study". *Language and Cognitive Processes, 15,* 507–537.
Real Academia Española (1985). *Esbozo de una nueva gramática de la lengua española*. Madrid: Espasa-Calpe.
Real Academia Española (1995). *Diccionario de la lengua española, CD-Rom Version*. Madrid: Espasa Calpe.
Reiter-Boomershine, Amanda (2004). "Perceiving and processing dialectal variation in Spanish: An exemplar theory approach". *Paper presented at the 8th Hispanic Linguistics Symposium*. Minneapolis, MN.
Riesbeck, Chris K., & Roger S. Schank (1989). *Inside case-based reasoning*. Hillsdale, NJ: Erlbaum.
Ringen, Jon D. (1975). "Linguistic facts: A study of the empirical scientific status of transformational generative grammars". In David Cohen & Jessica R. Wirth (Eds.), *Testing linguistic hypotheses* (pp. 1–42). Washington, DC: Hemisphere.
Rischel, Jørgen (1978). "Some general remarks on realism in current phonological work". In Kirsten Gregersen, Hans Basbøll, & Jacob Mey (Eds.), *Papers from The Fourth Scandinavian Conference of Linguistics* (pp. 419–431). Odense: Odense University Press.
Robinson, John, Helen Lawrence, & Sali Tagliamonte (2001). *GoldVarb 2001: A multivariate analysis application for Windows*.
Roca, Iggy (1986). "Secondary stress and metrical rhythm". *Phonology Yearbook, 3,* 341–370.
Roca, Iggy (1988). "Theoretical implications of Spanish word stress". *Linguistic Inquiry, 19,* 393–423.

Roca, Iggy (1990). "Morphology and verbal stress in Spanish". *Probus, 2,* 321–350.

Roca, Iggy (1991). "Stress and syllables in Spanish". In Héctor Campos & Fernando Martínez-Gil (Eds.), *Current studies in Spanish linguistics* (pp. 599–635). Washington, DC: Georgetown University Press.

Roca, Iggy (1997). "On the role of accent in stress systems: Spanish evidence". In Fernando Martínez-Gil, & Alfonso Morales-Front (Eds.), *Issues in the phonology and morphology of the major Iberian languages* (pp. 619–664). Washington, DC: Georgetown University Press.

Rosenblat, Angel (1952). "Género de los sustantivos en -e y en consonante". *Estudios dedicados a Menéndez-Pidal*, Vol. 3 (pp. 159–202). Madrid: Concejo Superior de Investigaciones Científicas.

Rueckl, Jay G. (1990). "Similarity effects in word and pseudoword repetition priming". *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16,* 374–391.

Rumelhart, David E., & James L. McClelland (1986). "On learning the past tense of English verbs". In James L. Mc Clelland, David E. Rumelhart, & the PDP Research Group (Eds.), *Parallel distributed processing*, Vol. 2 (pp. 216–271). Cambridge, MA: Massachusetts Institute of Technology Press.

Rusch, Debbie, Marcela Domínguez, & Lucía Caycedo Garner (1996). *Fuentes: Conversación y gramática.* Boston: Houghton Mifflin.

Saltarelli, Mario (1997). "Stress in Spanish and Latin: Where morphology meets prosody". In Fernando Martínez-Gil & Alfonso Morales-Front (Eds.), *Issues in the phonology and morphology of the major Iberian languages* (pp. 665–694). Washington, DC: Georgetown University Press.

Salvador, Gregorio (1977). "Unidades fonológicas vocálicas en andaluz oriental". *Revista Española de Lingüística, 7,* 1–23.

Sampson, Geoffrey (2001). *Empirical linguistics*. London: Continuum.

Sanders, Benjamin P. (1998). "The Eastern Andalusian vowel system: Form and structure". *Rivista di Linguistica, 10,* 109–135.

Saporta, Sol, & Heles Contreras (1962). *A phonological grammar of Spanish*. Seattle: University of Washington Press.

Scarborough, Don L., Charles Cortese, & Hollis S. Scarborough (1977). "Frequency and repetition effects in lexical memory". *Journal of Experimental Psychology, 3,* 1–17.

Schiller, Niels O. (1998). "The effect of visually masked syllable primes in the naming latencies of words and pictures". *Journal of Memory and Language, 39,* 484–507.

Schiller, Niels O. (1999). "Masked syllable priming of English nouns". *Brain and Language, 68,* 300–305.

Schlenck, Klaus-Jürgen (1988). "The analysis of the English velar nasal in lexical phonology: An experimental study". *Journal of English Linguistics, 21,* 29–38.

Schnitzer, Marc L. (1999). "La 's' realenga: Estudio empírico del silabeo en el español puertorriqueño con sujetos analfabetos". In Amparo Morales, Julia Cardona, Humberto López Morales, & Eduardo Forastieri (Eds.), *Estudios de lingüística hispánica: Homenaje a María Vaquero* (pp. 588–599). Río Piedras: Editorial de la Universidad de Puerto Rico.

Schreuder, Robert, & Harald Baayen (1995). "Modeling morphological processing". In Laurie B. Feldman (Ed.), *Morphological aspects of language processing* (pp. 131–154). Hillsdale, NJ: Erlbaum.

Schreuder, Robert, Marjon Grendel, Nanda Poulisse, Ardi Roelofs, & Marlies van de Voort (1990). "Lexical processing, morphological complexity, and reading". In D. A. Balota, G. B. Flores d'Arcais, & K. Rayner (Eds.), *Comprehension processes in reading* (pp. 125–141). Hillsdale, NJ: Erlbaum.

Schütze, Carson T. (1996). *The empirical base of linguistics: Grammaticality judgments and linguistic methodology*. Chicago: University of Chicago Press.

Schütze, Carson T. (2003). "Linguistic theory and empirical evidence: Clarifying some misconceptions". *Paper presented and the 25th Annual Meeting of the German Society for Linguistics*, Munich.

Sebastián-Gallés, Núria, Emmanuel Dupoux, Juan Seguí, & Jacques Mehler (1992). "Contrasting syllabic effects in Catalan and Spanish". *Journal of Memory and Language, 31,* 18–32.

Sebastián Gallés, Núria, M. Antònia Martín Antonín, Fernando Cuetos Vega, & Manuel F. Carreiras Valiña (2000). *LEXESP: Léxico informatizado del español*. Barcelona: Biblioteca Universitària.

Segui, Juan, Emmanuel Dupoux, & Jacques Mehler (1990). "The role of the syllable in speech segmentation, phoneme identification, and lexical access". In Gerry T. M. Altmann (Ed.), *Cognitive models of speech processing* 2(pp. 63–280). Cambridge, MA: Massachusetts Institute of Technology Press.

Segui, Juan, & María Luisa Zubizarreta (1985). "Mental representation of morphologically complex words and lexical access". *Linguistics, 23,* 759–774.

Seidenberg, Mark S. (1987). "Sublexical structures in visual word recognition: Access units or orthographic redundancy?" In Max Coltheart (Ed.), *Attention and performance XII: The psychology of reading* (pp. 245–263). Hillsdale, NJ: Erlbaum.

Seidenberg, Mark S. (1989). "Reading complex words". In Greg N. Carlson & Michael K. Tannenhaus (Eds.), *Linguistic structure in language processing* (pp. 53–105). Dordrecht: Kluwer.

Seidenberg, Mark S. (1992). "Connectionism without tears". In Steven Davis (Ed.), *Connectionism: Theory and practice* (pp. 84–122). New York: Oxford University Press.

Seidenberg, Mark, & Maggie Bruck (1990). "Consistency effects in the generation of past tense morphology". *Paper presented a the 31st meeting of the American Psychonomic Society*, New Orleans.

Seidenberg, M. S., & J. H. Hoeffner (1998). "Evaluating behavioral and neuroimaging data on past tense processing". *Language, 74,* 104–122.

Seidenberg, Mark, & James L. McClelland (1989). "A distributed developmental model of word recognition and naming". *Psychological Review, 96,* 523–568.

Sereno, Joan A. (1991). "Graphemic, associative, and syntactic priming effects at a brief stimulus onset asynchrony in lexical decision and naming". *Journal of Experimental Psychology: Learning, Memory, and Cognition, 17,* 459–477.

Sereno, Joan A., & Allard Jongman (1997). "Processing of English inflectional morphology". *Memory and Cognition, 25,* 425–437.

Shanks, David R. (1995). *The psychology of associative learning*. Cambridge: Cambridge University Press.

Sheidlower, Jesse (2002). "Confronting 'NOO'kyuh-luhr' proliferation". *New York Times, Oct. 13*, section 4, page 2, column 1.

Skousen, Royal (1975). *Substantive evidence in phonology.* The Hague: Mouton.
Skousen, Royal (1989). *Analogical modeling of language*. Dordrecht: Kluwer.
Skousen, Royal (1992). *Analogy and structure*. Dordrecht: Kluwer.
Skousen, Royal (1995). "Analogy: A non-rule alternative to neural networks". *Rivista di Linguistica, 7,* 213–232.
Slowiaczek, Louisa M., & D. B. Pisoni (1986). "Effects of phonological similarity on priming in auditory lexical decision". *Memory and Cognition, 14,* 230–237.
Smead, Robert N. (2000). "On the assignment of gender to Chicano anglicisms: Processes and results". *Bilingual Review, 25,* 277–297.
Smith, Neils (1999). *Chomsky: Ideas and ideals*. Cambridge: Cambridge University Press.
Spence, Kenneth W. (1956). *Behavior theory and conditioning*. New Haven, CN: Yale University Press.
Stanners, Robert F., James J. Neiser, William P. Hernon, & Roger Hall (1979). "Memory representation for morphologically related words". *Journal of Verbal Learning and Verbal Behavior, 18,* 399–412.
Stanovich, Keith E. (1996). *How to think straight about psychology*, 4th ed. New York: Harper Collins.
St. Clair, Robert (1971). "Diphthongization in Spanish". *Papers in Linguistics, 4,* 421–432.
Steinberg, Danny D. (1975). "Chomsky: From formalism to mentalism and psychological invalidity". *Glossa, 9,* 218–252.
Steinberg, Danny D., & Robert K. Krohn (1975). "The psychological validity of Chomsky and Halle's vowel shift rule". In E. F. K. Koerner (Ed.), *The transformational-generative paradigm* (pp. 233–259). Amsterdam: John Benjamins.
Stemberger, Joseph Paul (1985). "An interactive activation model of language production". In Andrew W. Ellis (Ed.), *Progress in the psychology of language*, Vol. 1 (pp. 143–186). London: Erlbaum.
Stemberger, Joseph Paul (1994). "Rule-less morphology at the phonology-lexicon interface". In Susan D. Lima, Roberta L. Corrigan, & Gregory K. Iverson (Eds.), *The reality of linguistic rules* (pp. 147–169). Amsterdam: John Benjamins.
Stemberger, Joseph Paul (1996). "The scope of the theory: Where does 'beyond' lie?" In Lisa McNair, Kora Singer, Lise M. Dobrin, & Michelle M. Aucoin (Eds.), *Proceedings of the parasession of the Chicago Linguistic Society's 32nd meeting* (pp. 139–164). Chicago: Chicago Linguistic Society.
Steriade, Donca (1997). *Lexical conservatism and its analysis*. Manuscript, University of California, Los Angeles.
Steriade, Donca (1999). In Linguistic Society of Korea (Eds.), *Lexical conservatism. Linguistics in the morning calm*, Vol. 4 (pp. 157–180). Hanshin: Linguistic Society of Korea.
Stockwell, Robert T., J. Donald Bowen, & Ismael Silva-Fuenzalida (1956). "Spanish juncture and intonation". *Language, 32,* 641–665.
Stockwell, Robert T., & J. Donald Bowen (1965). *The sounds of English and Spanish*. Chicago: University of Chicago Press.

Stolz, Jennifer A., & Laurie B. Feldman (1995). "The role of orthographic and semantic transparency of the base morpheme in morphological processing". In Laurie B. Feldman (Ed.), *Morphological aspects of language processing* (pp. 109–129). Hillsdale, NJ: Erlbaum.

Taft, Marcus (1979). "Recognition of affixed words and the word frequency effect". *Memory and Cognition, 7,* 263–272.

Taft, Marcus, & Kenneth I. Forster (1975). "Lexical storage and retrieval of prefixed words". *Journal of Verbal Learning and Verbal Behavior, 14,* 638–647.

Taylor, John R. (1995). *Linguistic categorization: Prototypes in linguistic theory*, 2nd ed. Oxford: Clarendon Press.

Teschner, Richard V. (1983). "Spanish gender revisited: -Z words as illustrating the need for expanded phonological and morphological analysis". *Hispania, 66,* 252–256.

Teschner, Richard V., & William M. Russell (1984). "The gender patterns of Spanish nouns: An inverse dictionary-based analysis". *Hispanic Linguistics, 1,* 115–132.

Tobin, Yishai (1997). *Phonology as human behavior*. Durham, NC: Duke University Press.

Treiman, Rebecca (1986). "The division between onsets and rimes in English syllables". *Journal of Memory and Language, 25,* 476–491.

Vogel Sosa, Anna, & James MacFarlane (2002). "Evidence for frequency-based constituents in the mental lexicon: Collocations involving the word 'of.' " *Brain and Language, 83,* 227–236.

Waltermire, Mark (2004). "The effect of syllable weight on the determination of spoken stress in Spanish". In Timothy L. Face (Ed.), *Laboratory approaches to Spanish phonology* (pp. 171–191). Berlin: Mouton de Gruyter.

Wang, H. Samuel, & Bruce L. Derwing (1986). "More on English vowel shift: The back vowel question". *Phonology Yearbook, 3,* 99–116.

Wang, H. Samuel, & Bruce L. Derwing (1994). "Some vowel schemas in three English morphological classes: Experimental evidence". In Matthew Y. Chen & Ovid J. L. Tzeng (Eds.), *In honor of William S-Y. Wang: Interdisciplinary studies on language and language change* (pp. 561–575). Taipei: Pyramid Press.

Wheeler, Cathy J. (1980). "On the relationship between phonology and psychology". *Papers in Linguistics, 13,* 51–100.

Whitley, Stanley (1976). "Stress in Spanish: Two approaches". *Lingua, 39,* 301–332.

Whitley, M. Stanley (1986). *Spanish/English contrasts*. Washington, DC: Georgetown University Press.

Whitley, M. Stanley (2002). *Spanish/English contrasts*, 2nd ed. Washington, DC: Georgetown University Press.

Woods, Anthony, Paul Fletcher, & Arthur Hughes (1986). *Statistics in language studies*. Cambridge: Cambridge University Press.

Yngve, Victor H. (1986). *Linguistics as a Science*. Bloomington and Indianapolis: Indiana University Press.

Yngve, Victor H. (1996). *From grammar to science: New foundations for general linguistics*. Amsterdam: John Benjamins.

Yngve, Victor H. (2000). "The depth hypothesis and the new hard science linguistics". In Milena Kovacic, Milojevic Sheppard, Silvana Orel-Kos, & Janez Oresnik (Eds.), *Exploring language from different perspectives* (pp. 191–202). Ljubljana: Filozofska fakulteta Univerze v Ljubljana.

Yoon, Yeo Bom, & Bruce L. Derwing (2001). "A language without a rhyme: Syllable structure experiments in Korean". *Canadian Journal of Linguistics, 46,* 187–237.

Young, Richard, & Robert Bayley (1996). "VARBRUL analysis for second language acquisition". In Robert Bayley & Dennis R. Preston (Eds.), *Second language acquisition and linguistics variation* (pp. 253–306). Amsterdam: John Benjamins.

Zamora, Juan Clemente (1975). "Morfología bilingüe: La asignación de género a los préstamos". *Bilingual Review, 2,* 239–247.

Zamora Vicente, Alonso (1974). *Dialectología española*, 2nd ed. Madrid: Gredos.

Zamora Munné, Juan C., & Eduardo Béjar (1987). "El género de los préstamos". *Revista española de lingüística, 17,* 131–137.

Zimmer, Karl E. (1969). "Psychological correlates of some Turkish morpheme structure conditions". *Language, 45,* 309–321.

Zwicky, Arnold M. (1975). "The strategy of generative phonology". In Wolfgang U. Dressler & F. V. Marks (Eds.), *Phonologica 1972* (pp. 151–165). Salzburg: Wilhelm Fink.

Zwicky, Arnold M., Peter H. Salus, Robert I. Binnick, & Anthony L. Vanek (1971). *Studies out in left field: Defamatory essays presented to James D. Mc Cawley on the occasion of his 33rd or 34th birthday*. Edmonton: Linguistic Research Inc.

Name index

Subject index

In the series STUDIES IN FUNCTIONAL AND STRUCTURAL LINGUISTICS
the following volumes have been published thus far or are scheduled for publication:

43 **ANDREWS, Edna and Yishai TOBIN (eds.):** Toward a Calculus of Meaning. Studies in markedness, distinctive features and deixis. 1996. xxviii, 432 pp.

44 **JESSEN, Michael:** Phonetics and Phonology of Tense and Lax Obstruents in German. 1999. xx, 394 pp.

45 **MARTÍN-VIDE, Carlos (ed.):** Mathematical and Computational Analysis of Natural Language. Selected papers from the 2nd International Conference on Mathematical Linguistics (ICML '96), Tarragona, 1996. 1998. xviii, 391 pp.

46 **ŠTEKAUER, Pavol:** An Onomasiological Theory of English Word-Formation. 1998. x, 192 pp.

47 **MARTÍN-VIDE, Carlos (ed.):** Issues in Mathematical Linguistics. Workshop on Mathematical Linguistics, State College, PA, April 1998. 1999. xii, 214 pp.

48 **REID, Wallis, Ricardo OTHEGUY and Nancy STERN (eds.):** Signal, Meaning, and Message. Perspectives on sign-based linguistics. 2002. xxii, 413 pp.

49 **HLADKÝ, Josef (ed.):** Language and Function. To the memory of Jan Firbas. 2003. x, 339 pp.

50 **DUŠKOVÁ, Libuše (ed.):** Dictionary of the Prague School of Linguistics. Translated from the French, German and Czech sources. Author: Josef Vachek †. In collaboration with Josef Dubský. Translated by Aleš Klégr, Pavlína Šaldová, Markéta Malá, Jan Čermák, Libuše Dušková. 2003. x, 216 pp.

51 **CONTINI-MORAVA, Ellen, Robert S. KIRSNER and Betsy RODRÍGUEZ-BACHILLER (eds.):** Cognitive and Communicative Approaches to Linguistic Analysis. 2004. viii, 388 pp.

52 **GORLACH, Marina:** Phrasal Constructions and Resultativeness in English. A sign-oriented analysis. 2004. x, 150 pp.

53 **EDDINGTON, David:** Spanish Phonology and Morphology. Experimental and quantitative perspectives. 2004. xvi, 197 pp.