

PATTERNS OF SOUND, PATTERNS IN MIND:
PHONOLOGICAL REGULARITIES IN SPEECH PRODUCTION

by
Matthew Andrew Goldrick

A dissertation submitted to Johns Hopkins University in conformity with the
requirements for the degree of Doctor of Philosophy.

Baltimore, Maryland, October 2002

©Matthew Andrew Goldrick 2002
All rights reserved

ABSTRACT

Linguistic research has documented a wide range of regularities in phonological structure. Within languages, certain sound combinations are ill-formed, never appearing in words of the language; across languages, certain phonological structures are preferred to others. Less well understood are the types of regularities that are encoded by the spoken production system. To explore this question, section 1 describes three theories regarding the types of regularities that are encoded. These theories are: one, the Instance-Based theory—gradient regularities based on within-language token frequency of segmental and supra-segmental structures are encoded; two, the Lexical Distribution theory—gradient regularities based on within-language type frequency of segmental and supra-segmental structures are encoded; and three, the Markedness theory—categorical regularities based on cross-linguistic and within-language markedness of sub-segmental, segmental, and supra-segmental structures are encoded.

Building on previous research, a framework for spoken production processing is described in section 2. The three theories are situated within this general framework. Section 3 then reviews previous research regarding the types of regularities that are encoded. These studies suggest that categorical within-language phonological regularities are encoded by the spoken production system, but fail to distinguish between the three theories.

Section 4 reports the results of two experimental studies designed to contrast the predictions of the three theories. These two experiments are the first to demonstrate that sub-segmental regularities must be encoded by the spoken production system. Experiment 1 uses an implicit learning paradigm. As predicted by the Markedness theory, participants in this experiment are sensitive to sub-segmental regularities. Furthermore, gradient regularities are encoded, supporting the predictions of the Instance-Based and Lexical Distribution theories. Experiment 2 examines biases in speech errors. The biases conform to the regularities of the Markedness theory, but exhibit gradient effects. These results support a theory incorporating elements of all three theories (i.e., gradient as well sub-segmental regularities are encoded).

Section 5 discusses the implications of the results presented in section 2, 3, and 4 for the computational mechanisms implementing phonological processing. Future work to extend this research is outlined, including an extension to existing computational theories that may account for the full range of results.

Advisors: Brenda Rapp

Paul Smolensky

ACKNOWLEDGEMENTS

This thesis is dedicated to the memory of my grandfather, Charles DeGeorge, who started my career in science by: one, teaching me to navigate by the stars; and two, buying me my first computer (a TRS-80).

I would like to thank those that kept me sane during my graduate education, and have made my life so much fun. Most of all, I thank Chris Gioia Goldrick for being a fantastic wife, friend, and party companion. I thank my friends for keeping me grounded in the “real” world outside of the lab. Much love to: everyone in Baltimore, especially the graduate students in Cognitive Science; the Crew; and all the freaks in Raleigh, Charlotte, D.C., New York, Boston, Chicago, Urbana, and Seattle. I’d also like to thank my family for their love and support. Special thanks to my parents, Ralph and Tina Goldrick, for the emotional and financial support for these many years.

I’d like to thank the faculty and students of the Department of Cognitive Science for teaching me so much over the past nine years. Thanks to Luigi Burzio, William Badecker, and Michael McCloskey for asking hard questions. Special thanks to Lisa Davidson and Colin Wilson for helping me navigate the peculiar world of sound structure.

As a sophomore, I took “Foundations of Cognitive Science” with Paul Smolensky. It opened up an entire world of questions and concepts that I had never considered before.

Amazingly, every course and every project with him has involved a similar moment—a view of a world of ideas outside the minutiae of any one particular project or discipline. I thank him for always reminding me of why cognitive science just might work.

From the first day I met her, Brenda Rapp has been the ideal teacher and mentor. If I've done one thing right during my education, it was choosing to continue to work with her in graduate school. Our collaboration has not only been intellectually productive but a pleasure. She has my eternal gratitude. With respect to this work in particular, I'd like to thank her for spending many hours puzzling over the relationship of psychology and linguistics with me.

For making this dissertation possible, I'd like to thank my committee: Dana Boatman, Sanjeev Khudanpur, Brenda Rapp, Amy Shelton, and Paul Smolensky, as well as William Badecker and Argye Hillis. For helpful comments on portions of this work, I'd like to thank audiences at Johns Hopkins University, Brown University, Massachusetts Institute of Technology, the HUMDRUM workshop at the University of Massachusetts Amherst, and the Eighth Conference on Laboratory Phonology (especially Donca Steriade for her thoughtful commentary). Thanks to Paul Smolensky and especially Brenda Rapp for providing detailed comments on earlier drafts. Last, but not least, I'd like to thank the several dozen individuals that suffered through boring tongue twister tasks to provide me with data.

TABLE OF CONTENTS

ABSTRACT	II
ACKNOWLEDGEMENTS.....	IV
TABLE OF CONTENTS.....	VI
LIST OF TABLES.....	VIII
LIST OF FIGURES	X
SECTION 1: INTRODUCTION.....	1
<i>The encoding of phonological regularities</i>	3
<i>Types of phonological regularities</i>	5
<i>Contrasting theories of regularities</i>	13
<i>Summary: Conceptual framework</i>	20
SECTION 2: A FRAMEWORK FOR SPOKEN PRODUCTION PROCESSING.....	22
<i>Cognitive processes</i>	22
<i>An architecture for spoken production</i>	31
<i>Internal structure of the spoken form component</i>	38
<i>Internal structure of the phonological sub-component</i>	49
<i>Summary: Spoken production processing</i>	66
SECTION 3: PREVIOUS STUDIES OF THE ENCODING OF PHONOLOGICAL REGULARITIES..	67
<i>Evidence supporting encoding of regularities</i>	68

<i>Null results regarding encoding of phonological regularities</i>	91
<i>Anti-regularity effects</i>	94
<i>Summary: Previous research on the encoding of phonological regularities</i>	97
SECTION 4: EXPERIMENTAL INVESTIGATIONS	98
<i>Experiment 1: Implicit learning of phonological regularities</i>	100
<i>Experiment 2: Biases in speech errors</i>	128
<i>General Discussion: Experimental investigations</i>	148
SECTION 5: GENERAL DISCUSSION	157
<i>Mechanisms of phonological processing</i>	157
<i>Future directions</i>	172
<i>Conclusion: Patterns of sound, patterns in mind</i>	173
APPENDIX A: CHANCE RATES OF VIOLATING TARGET SYLLABLE POSITION,	
EXPERIMENT 1	175
APPENDIX B: DETERMINATION OF REGULARITIES FOR EACH THEORY, EXPERIMENT 2.	178
APPENDIX C: PREPARATION OF MATERIALS, EXPERIMENT 2	180
REFERENCES	184
CURRICULUM VITAE	216

LIST OF TABLES

<u>Table 1.</u> Theoretical proposals regarding what types of regularities are encoded.....	98
<u>Table 2.</u> Restricted control and unrestricted consonants, Experiment 1.	103
<u>Table 3.</u> Percentage of errors violating target syllable position for each consonant pair, Experiment 1.	116
<u>Table 4.</u> Percentage of errors violating target syllable position for each condition, Experiment 1.	117
<u>Table 5.</u> Distributional similarity of consonants, Experiment 1.	122
<u>Table 6.</u> Cross-linguistic regularities, Experiment 2.	131
<u>Table 7.</u> Test pairs, Experiment 2.	132
<u>Table 8.</u> Control pairs, Experiment 2.	133
<u>Table 9.</u> Performance on control pairs, Experiment 2.	137
<u>Table 10.</u> Performance on test pairs, Experiment 2.	139
<u>Table 11.</u> Sum frequency of segment classes, control and test pairs.	150
<u>Table 12.</u> Difference in sum frequency of segment classes for control and test pairs...	152

Table A1. Mean statistics for control pairs, Experiment 2..... 181

Table A2. Mean statistics for test pairs, Experiment 2. 182

Table A3. Spelling of vowels, Experiment 2 stimuli. 183

LIST OF FIGURES

<u>Figure 1.</u> Architecture of the spoken production system.....	33
<u>Figure 2.</u> Performance on all segment pairs, Experiment 1.....	110
<u>Figure 3.</u> Comparison of /f/ and /s/ across conditions, Experiment 1.....	112
<u>Figure 4.</u> Comparison of /f/ across conditions, Experiment 1.....	113

SECTION 1: INTRODUCTION

Our science, our mathematics, our languages are all patterns of patterns.

(Johnson, 1996: 323)

Phonological descriptions characterize words (i.e., lexical items) at an abstract level of form. One phonological description of the word “king” is /kɪŋ/. This description specifies, in part, that the word is composed of: one, an obstruction of the vocal tract near the back of the mouth without vocal cord vibration (/k/); followed by an open vocal tract with the tongue body high and front in the mouth, with vocal cord vibration (/ɪ/); and ending with an obstruction of the vocal tract at the back of the mouth, with vocal cord vibration and the velum (fleshy structure between the oral and nasal cavities) lowered (/ŋ/). This description is abstract, in part because it does not detail the precise physical realization of a particular utterance. For example, it does not specify that the /k/ closure occurs at point X in the back of the mouth, but merely specifies a region of the oral cavity in which the closure should occur. The description is at the level of form because it concerns itself with distinctions that do not necessarily signify changes in meaning. “King” was decomposed into 3 parts that provide no clue as to the meaning of the whole¹. For example, “corn” starts with the same sound (phonologically speaking) as “king,” yet “corn” has nothing to do with medieval political structures.

¹ This is not to say that sounds never convey meaning; for example, onomatopoeic words like “buzz” sound like the thing they denote.

As linguistic research has shown, languages exhibit patterns of preferences (regularities) at the phonological level; certain phonological structures are more well-formed than others. Within a language, we see that concepts are not paired with any possible sequence of sounds; rather, there are constraints on the types of sound strings that are used. For example, in English, lexical items beginning with the segment /ŋ/ are dispreferred or ill-formed, even though the segment may occur at the end of words (as it does in “king”). This regularity is specific to English; for example, Vietnamese allows word-initial /ŋ/ (e.g., “Nguyen”). Not only are there phonological regularities within languages, but there are also regularities across languages. For example, across many languages we find evidence of a regular relationship between two places of articulation, coronal and dorsal. Coronal articulations involve the tongue tip and front of the mouth (e.g., the first sound in “top”) and dorsal articulations involve the body of the tongue and back of the mouth (e.g., the first and last sounds in “king”). First, languages tend to have more coronal phonemes than dorsal phonemes². In French, for example, there are 9 coronal phonemes but only 2 dorsal ones (Paradis & Prunet, 1991a). Second, languages tend to restrict the occurrence of dorsal phonemes (but not coronal phonemes) in certain environments. For example, Finnish has both coronal and dorsal stops at the beginning of words, but no dorsal stops are found at the end of words (Yip, 1991). Converging

² Traditionally, a phoneme is defined as the minimal phonological unit that can be used to signal a contrast in meaning (Anderson, 1985).

lines of evidence like these suggest that across languages, the coronal place of articulation is preferred to that of dorsal (see the papers in Paradis & Prunet, 1991b, for a review).

The encoding of phonological regularities

At many levels of description, then, phonological structure exhibits regularities. This dissertation is concerned with the relationship between these regularities and spoken production processes. The following questions are examined: does the spoken production system encode these regularities? What types of regularities are encoded?

What do I mean by “encoded”? I assume that the cognitive system is made up of a set of interacting processes; a subset of these support our ability to produce language. This set of processes encodes language by reflecting aspects of its structure; in other words, there is a correspondence between the structure of language and the structure of the cognitive system (Palmer, 1978). It is important to distinguish between aspects of language structure that are directly encoded in the structure of the cognitive system versus those that are only indirectly encoded. Indirectly encoded structure is present in the input and output of the system but is not reflected by the system’s internal structure (see Cummins, 1986, for discussion of a similar notion of “inexplicit information”).

For example, suppose our cognitive system encoded phonological structure by maintaining a “list” of the phonological forms of all the words in our language. In this

system, phonological processing consists of retrieving the phonological form of each word in a sentence and concatenating them together (this is clearly inadequate to explain our actual behavior). Such a cognitive system directly encodes the phonological form of each word; it does not directly encode phonological regularities across words. However, even though these regularities are not directly encoded, they can still be observed in the outputs of the cognitive system. If more words begin with /t/ than with /k/, this regularity would be reflected in the distribution of outputs. In this sense, the cognitive system indirectly encodes regularities. Its outputs reflect regularities, even though they are not reflected by the internal structure of the cognitive system.

In contrast, suppose the “list” of phonological forms was organized such that the words with the more frequent initial segments (e.g., /t/-initial words like “top”) were listed before words with less frequent initial segments (e.g., /k/-initial words like “cop”). If retrieval of words required cycling through the entire list, the cognitive system would require less time to access words with frequent initial segments. In this system, generalizations across words are directly encoded. A phonological regularity (initial segment frequency) directly corresponds with some aspect of a cognitive processing mechanism (the mechanism searching for stored phonological word forms uses segment frequency to guide its search). This thesis is concerned with identifying the types of regularities that are directly encoded by the cognitive system.

Within the cognitive system, I focus specifically on spoken production processes, as distinct from the processes involved in perception, encoding of meaning, etc. Underlying this question is the assumption that there is a division of labor within the cognitive system. Cognitive processes, or components of the cognitive system, each encode some subset of the knowledge possessed by the whole of the system (the total knowledge of the system being a function of the component processes as well as the interactions among them). Within this system, then, we can ask: what (perhaps proper) subset of regularities encoded by the cognitive system are encoded by speech production processes in particular?

Types of phonological regularities

Before asking what particular regularities are encoded by the spoken production system, we must consider what types of regularities could be encoded. A phonological regularity is a characterization of phonological well-formedness. Regularities distinguish structures that are ill-formed (e.g., absent or infrequent in a language) from those that are well-formed (e.g., present or highly frequent). For example, a within-language regularity in English categorizes words beginning with /ŋ/ as ill-formed (or irregular), while words beginning with /n/ are classified as well-formed (or regular). I will define regularities in terms of three features:

- Scope: How widespread is the regularity? Here, I'll distinguish two types of regularities. If a regularity characterizes well-formedness within a particular language, it is a within-language regularity. If it characterizes well-formedness across all human languages, it is a cross-linguistic regularity.
- Scale: If a regularity characterizes patterns using a continuous scale, it is a gradient regularity. For example, word-finally in English, /k/ has a relative frequency of .028, while /g/ has a relative frequency of .003. According to this regularity, /k/ is not only more well-formed than /g/, but it is 9.3 times as well-formed as /g/. In contrast, a categorical regularity only distinguishes well-formed and ill-formed structures. For example, in English, there are no words that begin with /ŋ/. This regularity as a categorical distinction: words that begin with /ŋ/ are absolutely ill-formed relative to words that begin with other segments in the English inventory (e.g., /n/).
- Granularity: At what level of phonological structure is the regularity found? I'll distinguish three levels of structure: sub-segmental, segmental, or supra-segmental. These are defined in the next subsection. Note that regularities may be stated within a particular structural context (hereafter, the contextual restriction of the regularity). For example, regularities may characterize the well-formedness of a particular sub-segmental unit (e.g., voicing) within the context of a particular

supra-segmental unit (e.g., codas: in German, voiceless, but not voiced stops, are found in coda) . Some regularities are not restricted to a particular context; they are generalizations across all contexts.

Levels of phonological structure

Sub-segmental level

The most basic phonological distinctions are found at the sub-segmental level. There are two basic proposals for how these distinctions are best characterized. One is distinctive feature theory (Chomsky & Halle, 1968). This proposal is based on the observation that human languages make use of a limited number of articulatory³ dimensions to contrast sounds. A distinctive feature is a categorical specification of a speech sound along one of these dimensions. For example, many languages distinguish consonant sounds in terms of presence or absence of vocal fold vibration; this theory uses the feature [voice] to specify whether vibration is present [+voice] or absent [-voice]. A contrasting proposal is that these distinctions are best characterized in terms of gestures (Browman & Goldstein, 1989, 1992). Gestures can be distinguished from features in at least two ways: one, gestures can express degrees along an articulatory dimension (not

³ Prior to Chomsky & Halle, theorists made use of acoustically based features (Jakobson, 1941/1968; Trubetzkoy, 1939/1969). This idea has been resurrected in some recent work (Coleman, 1998; Harris & Lindsey, 2000). Here I maintain the articulatory basis of distinctive features. Also note that it has been proposed that features (e.g., Sagey, 1986) and gestures (e.g., Browman & Goldstein, 1989) can be organized into more complex units that are larger than the most basic sub-segmental distinctions but smaller than segments (see below for discussion of segments). The role of such units within spoken production processing has not been widely investigated; I therefore omit discussion of such distinctions.

just the categories + or –); two, gestures can specify degrees of duration (whereas a feature is merely present or absent). Gestures are therefore more directly related to the physical realization of sounds in particular utterances, whereas features are more abstract.

Both approaches assume that the most basic phonological level specifies variation along a limited number of articulatory dimensions. I remain neutral with respect to the featural and gestural proposals; none of the studies reviewed or presented here will distinguish between the predictions of these theories.

Segmental level

The term “segment” has been used in a number of different ways. I use the term to refer to a phonological level that characterizes spoken forms as a series of units (segments) that serve to anchor other levels of structure. For example, a segment can serve as an anchor for a number of sub-segmental units, grouping them into a single unit. In Autosegmental Phonology, the segmental level is known under a variety of names: the CV (for Consonant Vowel), skeletal, timing (Goldsmith, 1990), and root node level (Archangeli & Pulleyblank, 1994).

The role of such a level is less clear in gestural theory (i.e., Articulatory Phonology; Browman & Goldstein, 1989, 1992). This theory proposes that gestures can enter into coordination relationships with one another, establishing larger units. Specifically, it is possible for segment-sized coordination relationships to arise (although other size

coordination relationships are possible). This has been noted by researchers in this general framework, who have proposed that “gestures cohere in bundles corresponding, roughly, to traditional segmental descriptions” (Saltzman & Munhall, 1989: 365; see also Byrd, 1996).

As with the feature/gesture distinction, I remain neutral as to whether Autosegmental notions of segments or Articulatory Phonological notions of coordination relationships are more appropriate characterizations of the segmental level, and assume that the results reported here are similarly neutral. I use the term “segment” to refer to a phonological level that organizes sub-segmental units into groups.

Supra-segmental level

Many different supra-segmental distinctions have been proposed. Most of them are based on the notion of syllable (Kenstowicz, 1994)⁴. A syllable is organized around the peak, a unit composed of one or two segments (e.g., in English, usually a vowel). The peak is surrounded, potentially on both sides, by margin segments. For example, the word “supplant” has two syllables (at least in slow speech). The first syllable has as its peak a schwa. Preceding this peak is the margin segment /s/. The second syllable has as its peak the vowel /æ/. The pre-peak margin segments are /pl/, and the post-peak margin

⁴ Articulatory Phonology subsumes the notion of syllabic structure under the more general concept of gestural coordination (the only difference between segment- and syllable-sized coordinations being that the span of coordination structures is larger; Browman & Goldstein, 1992). As before, I assume that the results presented below do not distinguish between more traditional notions of syllables and large-scale gestural coordination relationships.

segments are /nt/. Within a language, peak and margin positions tend to be filled by different classes of segments. For example, in English, peak positions tend to be filled by vowels and margins by consonants. But the distinction between peaks and margins does not reduce to the sub-segmental distinctions between segments (e.g., vowels and consonants). Considering English again, nasal and liquid segments can occur in both peak and in margin (e.g., /n/ in margin — “knot”; /n/ in peak — “button”).

The syllable itself has internal structure (beyond peak and margin). The peak and post-peak margin are referred to as the rime, while the pre-peak margin is referred to as the onset. The onset-rime distinction is based on linguistic research showing that the distribution of stress can be influenced by the content of the peak and the post-peak margin (i.e., the rime) but not by the content of the pre-peak margin (i.e., the onset). Within the rime, the post-peak margin is referred to as the coda.

A second type of supra-segmental structure is metrical. Syllables are organized into prosodic groupings, where one syllable has greater prominence than other. This is usually referred to as stress.

Summary: Phonological levels

There are three basic levels of phonological structure, constituting a loose hierarchy⁵.

1. Sub-segmental level: The most basic level of phonological representation, expressing a limited number of distinctions related to articulatory dimensions.
2. Segmental level: A level that organizes the sub-segmental level into groups.
3. Supra-segmental level: A level that organizes the segmental level into groups and expresses prosodic structure.

A sample of phonological regularities

In this section, I review some examples of phonological regularities to illustrate the three features that define a regularity.

Example 1: German word-final devoicing

In Standard German (hereafter, German), there are no words within the native lexicon that end in voiced stops. Sound sequences like *[hand]⁶ are absolutely ill-formed relative to sequences like [hant] ‘hand.’ Using the three features above, we can define this regularity as:

- Scope: Within-language. This regularity characterizes well-formedness within German.

⁵ It is important to note that “hierarchical” does not imply that lower levels of structure are grouped exclusively into higher levels of structure. For example, a sub-segmental unit can be shared across two segments, and a segment can be shared across two supra-segmental units. In other words, the “groups” defined by higher level structure can overlap (see Goldsmith, 1990, for discussion).

⁶ ‘*’ denotes a form that is extremely unlikely to be a form within the language (i.e., an ungrammatical or phonotactically illegal form).

- Scale: Categorical. Words with final voiceless stops are regular (well-formed); those with final voiced stops are irregular.
- Granularity: Sub-segmental. This regularity characterizes well-formedness of a particular sub-segmental unit (voicing).
 - Contextual restriction: Supra-segmental. The regularity is specific to the end of words.

Example 2: English word-final voiced vs. voiceless stops

Considering the distribution of word-final segments in other languages, we can also find gradient regularities. In English, word-final voiceless stops are more well-formed than voiced stops, but voiced stops are not absolutely ill-formed. In the CELEX lexical database (Baayen, Piepenbrock, & Gulikers, 1995), voiceless stops have a type frequency of 14.4%, while voiced stops have a type frequency of 11.3%. This regularity can be defined as:

- Scope: Within-language. This regularity characterizes well-formedness in English.
- Scale: Gradient. Words with final voiceless stops are well-formed relative to those with final voiced stops (although final voiced stops are allowed).
- Granularity: Sub-segmental. This regularity concerns the distribution of a particular sub-segmental unit (voicing).

- Contextual restriction: Supra-segmental. This regularity is specific to a particular word position in English.

Example 3: Cross-linguistic distribution of onset vs. coda

There are regularities not only within languages but also across languages. Cross-linguistically, syllables with codas are ill-formed relative to syllables with onsets.

Languages that have syllables with codas always have syllables with onsets (Bell, 1971; but see Breen & Pensalfini, 1999); furthermore, there are many languages that have syllables with onsets but do not have syllables with codas.

- Scope: Cross-linguistic. This regularity characterizes well-formedness across human languages.
- Scale: Categorical. Syllables with codas but not onsets are absolutely ill-formed relative to syllables with onsets.
- Granularity. Supra-segmental. This regularity concerns the well-formedness of two syllabic constituents, onset and coda.
- Contextual restriction: None. This regularity is true across all contexts.

Contrasting theories of regularities

Different types of regularities often make similar well-formedness distinctions. For example, with respect to scope, structures that are irregular within a particular language are also irregular across languages. In German and English, we find that (in coda)

voiceless stops are preferred to voiced stops (examples 1 and 2). The same (categorical) regularity is found cross-linguistically (Lombardi, 1995; Maddieson, 1984). Similar correlations have been reported in multiple languages for a number of different cross-linguistic regularities (Berg, 1998; Frisch, 1996, 2000; Greenberg, 1966; Trubetzkoy, 1939/1969; Zipf, 1935). This complicates the determination of which particular regularities are encoded by the language processing system. Suppose that the spoken production system prefers voiceless stops to voiced stops. Is this due to the encoding of a within-language or cross-linguistic regularity? Since the two regularities are correlated, this observation alone cannot distinguish between the two alternatives.

Below, I describe three proposals that make specific claims about the types of phonological regularities are encoded. Although these proposals exhibit some overlap, they also make certain distinct predictions. By contrasting these theories, we can gain some insight into the particular region of regularity space encoded by the spoken production system.

Instance-Based Theory

According to this theory, the spoken production system encodes regularities based on token frequency—the number of instances of a phonological structure in running speech. Structures that occur infrequently in running speech are ill-formed relative to those that occur frequently. It is based on a processing theory of speech perception (PARSYN:

Luce, Goldinger, Auer, & Vitevitch, 2000). Regularities associated with two different types of structures are encoded. First, regularities based on the (log-weighted) token frequency of segments in particular linear positions in the word are encoded. In other words, there are regularities based on the frequency of segments in first position (e.g., /t/ in “top”), and another set of regularities for segments in second position (e.g., /t/ in “stop”). Regularities based on the transitional probability of all possible pairs of segments (based on token frequency) are also encoded. For example, the regularity of the sequence /tɑ/ (as in “top”), will be based on the forward probability of /t/ followed by /ɑ/ as well as the backward probability of /t/ preceding /ɑ/.

With respect to each feature of regularities, then, this theory makes certain claims regarding what types of regularities are encoded:

- Scope: Within-language. Well-formedness is based on token frequency within a language.
- Scale: Categorical (presence versus absence in frequency counts) as well as gradient (more or less frequent).
- Granularity: Segmental (token frequency of segments) and supra-segmental (transitional probabilities of segments).

Lexical Distribution Theory

According to this theory, the spoken production system encodes regularities based on type frequency—the number of lexical items that contain a given phonological structure. Structures that occur in few lexical items are ill-formed relative to those that occur in many lexical items. Coleman & Pierrehumbert (1997; see also Frisch, Large, & Pisoni, 2000) formulate a Stochastic Phonological Grammar (SPG) that embodies this assumption. In the SPG, regularity of phonological structures is defined by the relative type frequency of syllable constituents in particular contexts/positions. Syllable constituents are defined in terms of segments (e.g., the onset of “top” /t/ is distinct from the onset of “stop” /st/). These constituents are then distinguished in terms of: one, position of the syllable in the word (initial, medial, final⁷); two, stress value of the syllable; and three, whether the constituent is an onset or rime. For example, some regularities characterize the well-formedness of onsets of word-initial stressed syllables (e.g., /t/ in “top”). Another set of regularities characterizes the well-formedness of onsets of word-initial unstressed syllables (e.g., /t/ in “topography”).

⁷ Coleman & Pierrehumbert (1997) describe only mono- and disyllabic words, distinguishing the following syllable positions: initial, non-final (e.g., /kæt/ in “Kathmandu”); final, non-initial (e.g., /kæt/ in “meerkat”); and simultaneously initial and final (e.g., /kæt/ in “cat”). Frisch et al. (2000) extend this to multi-syllabic words and only distinguish initial, medial and final positions. The predictions of these variants are, in fact, rather similar (see section 4).

With respect to each feature, this theory claims that the following types of regularities are encoded by the spoken production system:

- Scope: Within-language. Well-formedness is based on by type frequency within a language.
- Scale: Categorical (presence versus absence in frequency counts) as well as gradient (more or less frequent).
- Granularity: Segmental (the units that make up syllable constituents) and supra-segmental (type frequency of syllable constituents in certain prosodic/word environments).

Markedness Theory

This theory is based on a distillation of a number of different linguistic theories built around the notion of markedness (Battistella, 1996; Chomsky & Halle, 1968, chapter 9; Greenberg, 1966, 1978; Jakobson, 1939/1984; Kean, 1975/1980; Trubetzkoy, 1939/1969; Prince & Smolensky, 1993). The notion of “markedness” is based on cross-linguistic regularities. To discover such regularities, linguists use converging evidence from a variety of cross-linguistic generalizations. One set of generalizations concerns typological implications—whether the presence of some sound structure in a language implies the presence of some other sound structure (Greenberg, 1966; Maddieson, 1984). An example of such an implication is example 3 above; the presence of codas in a

language implies the presence of onsets. Another set of generalizations concerns defective distributions — whether different languages tend to ban some sound structure but not another in a particular environment (Battistella, 1996; Yip, 1991). For example, in many languages, coronal sounds are found in certain positions where dorsal sounds are not (Yip, 1991). The expectation is that converging evidence from a wide variety of sources (e.g., absence of some sound structure in many languages as well as severe restrictions on the structure in many others) will provide the best picture of the true cross-linguistic regularities. Linguistic research has found these cross-linguistic regularities at all the levels of structure discussed above (sub-segmental, segmental, and supra-segmental).

Markedness theory defines these regularities using a dichotomy between phonological structures. Marked structures are cross-linguistically ill-formed, while unmarked structures are well-formed⁸. Within-language regularities are then defined in terms of markedness. Particular languages either respect the markedness distinction or are neutral. For example, a cross-linguistic regularity is that onsets are preferred to codas (example 5 above). Particular languages can either respect this regularity (allowing onsets, but not codas), or remain neutral (allowing both onsets and codas). Crucially,

⁸ Note that this requires a certain abstraction over the (sometimes gradient) cross-linguistic data.

according to markedness theory, no language can respect the opposite pattern (allowing codas, but not onsets).

With respect to each feature, the markedness theory claims the following types of regularities are encoded:

- Scope: Cross-linguistic (as defined by markedness) as well as within-language (whether the language respects a markedness distinction or remains neutral).
- Scale: Categorical. Markedness is categorical; for a given pair of structures, one is marked/ill-formed, and the other unmarked/well-formed⁹.
- Granularity: Sub-segmental, segmental, and supra-segmental.

Contrasts between these theories

These 3 theories claim certain types of regularities are encoded by the spoken production system. Here, I briefly note major distinctions between these theories with respect to each feature defining regularities.

- Scope: Do spoken production processes encode cross-linguistic regularities?
— Instance-Based: No; Lexical Distribution: No; Markedness: Yes.
- Scale: Do spoken production processes encode gradient regularities?
— Instance-Based: Yes; Lexical Distribution: Yes; Markedness: No.

⁹ Some theories have allowed for ordinal markedness distinctions, stating an ordered (but not continuous) well-formedness relationship between structures (e.g., Prince & Smolensky's (1993) constraint HNUC, which distinguishes the relative markedness of several sonority categories).

- Granularity: Do spoken production processes encode sub-segmental regularities?
 - Instance-Based: No; Lexical Distribution: No; Markedness: Yes.

It is important to note that these theories constitute claims about what types of phonological regularities are encoded; they do not constitute claims about how these regularities are encoded. Recall the hypothetical cognitive system outlined above, where segmental frequency is reflected by the order in which the system searches for stored phonological word forms. With respect to phonological regularities, there are two separate (but related) properties of this system: one, it encodes a phonological regularity based on segmental frequency; two, it encodes this regularity in a mechanism that searches for stored phonological word forms. The theories outlined above constitute claims about the former properties (i.e., what types of regularities are encoded?). This dissertation focuses on these claims.

Summary: Conceptual framework

At the phonological level, human language exhibits a rich set of regularities. These regularities can be defined in terms of 3 features: scope, scale, and granularity. This thesis aims to explore what types of regularities are encoded by the speech production system. To examine this question, I have defined 3 theoretical positions that claim different types of regularities are encoded.

The second section lays out the foundation for an investigation of these theories by characterizing an architecture for the spoken production system. This architecture specifies the cognitive processes involved in producing speech (and their relation to other cognitive processes). Particular attention will be devoted to those cognitive processes that represent and manipulate phonological structure (as these processes will be involved in encoding phonological regularities). This architecture forms the basis for interpretation of behavioral results reviewed in the third section. In that section, I discuss previous studies that have examined the types of regularities encoded by the speech production system. These studies support the idea that the production system encodes regularities, but do little to resolve the question of what particular types of regularities are encoded. This problem is addressed in the fourth section, where I present experimental studies that examine certain contrasts between the theories reviewed above. The final section discusses implications of these results for theories of phonological processing and discusses future extensions of this work.

SECTION 2:

A FRAMEWORK FOR SPOKEN PRODUCTION PROCESSING

Psychological experiments measure the performance of individuals while they perform certain tasks. To evaluate and design experiments, we must have a theory of what underlies their performance in these tasks. This section presents such a framework, within which we can evaluate contrasting proposals concerning the types of phonological regularities encoded by the production system. The first sub-section lays out the basic assumptions about cognitive processes: the building blocks of the processing architecture. The processing architecture is then developed, at an increasing level of detail, concluding with a discussion of phonological processing in spoken production.

Cognitive processes

I assume that our ability to produce spoken language is supported by a number of different cognitive processes. A cognitive process computes a function (here, a “processing function”) that maps elements in the domain of the function (“input patterns”) to elements in the range (“output patterns”). For example, suppose there is a lexical semantics processing component. It computes a processing function that maps input patterns encoding the meaning of lexical concepts (e.g., [gasoline engine, multiple passengers, land transportation, road], [furry, domesticated, feline]) onto output patterns encoding particular lexical items (e.g., [BUS], [CAT]). Processing functions are

implemented by a mechanical process that transforms an input representation into an output representation (Palmer & Kimchi, 1986). These representations realize the input and output patterns by associating distinct patterns with distinct representational states¹⁰.

The process is “mechanical” because the production of an output representation is automatic, governed purely by the structure¹¹ of the input representation and the internal organization of the process. It is important to note that there is a distinction between a processing function and its implementation. Following this distinction, I will first discuss the properties of processing functions, and then discuss how such functions could be implemented.

Processing functions specify a structured, probabilistic relationship between input and output patterns. For the moment, I limit the discussion to processes that take a single pattern as input and produce a single pattern as output. Although only one pattern will be output at any particular time, processing functions are flexible; they specifying a mapping

¹⁰ This does not entail that all distinct representational states map onto distinct patterns. For example, suppose the output representation of the lexical semantics process consists of two units A and B, and that these two units encode two output patterns. Suppose the output pattern [BUS] corresponds to output representational state [A=1, B=0]. Similarly, [CAT] corresponds to output representational state [A=0, B=1]. All patterns correspond to distinct representational states. However, we can generate another distinct representational state that corresponds to no pattern. For example, [A=.5, B=.5] is different from both states listed above, but it does not correspond to any single pattern.

¹¹ The mechanical nature of the processing system raises some rather difficult questions. As mechanical processes, cognitive processes operate purely on the structure of the representations. For example, your pocket calculator does not “know” what the content of the key marked “2” is. If you popped off the label for the key and replaced it with “5”, the calculator’s behavior would not change; if you entered “5+1”, it would output “3” not “6.” This has struck many theorists as an undesirable property of a theory of cognition. Addressing this issue is outside the bounds of this dissertation; here I will assume that a mechanical explanation is appropriate for studying cognition. For discussion of the problem of content and its relation to cognitive science (in particular cognitive psychology), see Cummins (1991) and Palmer (1978) and references therein.

from an input to multiple similar outputs (where similarity is a function of the structure of the input and/or output representations). These multiple outputs express the range of outputs that could be observed across many different responses. I refer to the probabilities associated with these mappings as “output probabilities.” Under ideal circumstances, a cognitive process computes the highest probability mapping. Under less than ideal circumstances (e.g., damage to the processing mechanisms, a reduction in processing resources), the cognitive process will not completely fail; instead, it will produce (across multiple responses) a range of similar outputs in proportion to their output probabilities. This allows the processing function to gracefully degrade, continuing to occasionally produce the “correct” (i.e., highest probability) mapping, as well as outputs that are similar to the target.

To illustrate the concept of processing function, consider the lexical semantics example. For the input pattern [gasoline engine, multiple passengers, land transportation, road], the processing function might assign 85% output probability to the output pattern [BUS] (which matches all the semantic features of the input), 10% to [TRAIN] (sharing 3 out of 4 features), and 5% to [PLANE] (sharing 2 out of 4 features). Under ideal circumstances, this cognitive process will map the input to [BUS] (i.e., compute the highest probability mapping). Following damage, the process will exhibit graceful

degradation—[BUS] will be the most likely output, followed by [TRAIN] and lastly by [PLANE].

Cognitive processes take time to generate an output for a given input (here, “processing time”). I assume that output probabilities and processing times are inversely related; that is, high probability mappings will take less time to generate than low probability mappings. For example, following the lexical semantics example, suppose the features [gasoline engine, one to two passengers, land transportation, can be self-propelled] map to [MOPED] with 55% probability. Due to the difference in absolute probabilities, it would take more time for the lexical semantics process to generate [MOPED] (output probability 55%) as compared to [BUS] (output probability 85%).

Based on these assumptions, I will examine error and reaction time data to uncover the structure of cognitive processes. First, the probabilities associated with errors provide us with an estimate of output probabilities, and thus a characterization of the processing function. For example, the occurrence of errors reflecting degrees of semantic similarity (e.g., target “bus” produced as “train” more often than as “plane”), would support the lexical semantics processing function characterized above. Second, reaction times can be used to estimate processing time, providing us with another window into the structure of the processing function. For example, if experiments showed that producing words with highly semantically related neighbors (e.g., target “moped” is closely related to

“motorcycle”) is slower than producing words without highly related neighbors (e.g., “bus”), we would have evidence for a processing function based on semantic similarity.

I also assume that if a cognitive process has recently produced some output, the output probabilities associated with that pattern will be temporarily increased. This assumption allows us to couple priming methodologies with error and reaction time studies. If priming is based on the “normal” output probability (i.e., it is a function of the output probability assigned by the processing function), we can compare two primed representations on error and reaction time measures using the same logic as above. For example, recall that two alternatives for [BUS] were [TRAIN] (output probability 10%) and [PLANE] (output probability 5%). If priming doubles the output probability, then priming [TRAIN] will increase its probability to 20% and priming [PLANE] will increase its probability to 10%. Since the output probability relationship is preserved, error and reaction time data will show the same (qualitative) relationship under priming as they would in the absence of priming. The difference is that the quantitative relationship is magnified; this increases the power of our observations (e.g., more errors will be produced).

Mapping to multiple patterns

The above discussion has assumed that processes generate a representation corresponding to a single output pattern. I assume that representations containing

multiple output patterns are also possible. In these representations, the degree to which each pattern is present is related to its output probability. For example, the lexical semantics process could produce a representation consisting of 85% [BUS], 10% [TRAIN], and 5% [PLANE]. Note that a distributional representation is not an unstructured collection of patterns; it not only specifies that the patterns for “bus,” “train,” and “plane” are present in the representation, but that “bus” is the most dominant pattern. Subsequent processes will be sensitive to this dominance, ensuring that the ultimate output will be appropriate to a single output (e.g., “bus”).

The relationship between the distribution over output patterns and the output probabilities of these patterns allows a similar interpretation of error and reaction time data, as well as the use of priming methodologies. With respect to error data, I assume that under ideal conditions the probability of a particular distribution over output patterns will be maximally similar to the distribution of output probabilities. Under less than ideal operating conditions, the probability of distributions over output patterns will be based on similarity to the distribution of output probabilities. Within the lexical semantics example, the most probable distributions will be those in which [BUS] dominates the representation. Less probable distributions will be ones in which [TRAIN] dominates the representation, and even less probable will be ones in which [PLANE] dominates the representation. Thus, the likelihood with which a given error is produced (i.e., a given

output pattern dominates the representation) is related to the output probability of the corresponding pattern¹².

Similar assumptions guide the interpretation of reaction time and priming data. With respect to processing time, I assume that distributions that are composed of a single output pattern take less time to generate than those composed of multiple output patterns. Thus, the greater the absolute output probability of a pattern, the more the output representation will be dominated by a single pattern, and the less time it will take to generate that output. Finally, with respect to priming, I assume that the increase in output probabilities for primed patterns will result in an increased presence in the output distribution.

Although distributional and non-distributional representations have a similar relationship to psychological data (e.g., reaction times, errors), they do not have an identical relationship. For example, distributional representations allow complex information to be relayed to other processes. In the lexical semantics example, a distributional representation could allow other processes to be sensitive to semantic neighbors of the highest probability output pattern.

¹² Proximity to the output probability distribution may not be the only factor that determines the probability of different distributions. For example, processes may require that representations be relatively unambiguous—for example, requiring that one single pattern make up 60% of the output. This requirement is commonly implemented in spreading activation frameworks (see below) through lateral inhibition mechanisms, which force units to “compete” and ensure that if one unit is active the others units are inactive.

Implementing processing functions

A commonly used framework in psychological theory for implementing processing functions is that of the spreading activation network¹³. In a spreading activation network, input and output patterns are instantiated as patterns of activity over sets of simple processing units. Sets of connections allow activation values to pass between these units. Connections are designed to realize output probabilities by distributing activation values over output units in proportion to output probability (i.e., the higher the probability, the higher the activation value). The output of the network is determined by relative activation value; the pattern that is output by the network will be the one corresponding to the most active units.

To illustrate the operation of a spreading activation network, we could implement the lexical semantics process in the following way. The network would consist of a set of input units corresponding to semantic features, and a set of output units corresponding to lexical items. Connections between these units would create an activation distribution over the output units corresponding to the output probability distribution. For example, suppose the input is 1.0 units of activation on each of the following input units: [gasoline engine], [multiple passengers], [land transportation], [road]. If this is the input, the connections with the output units would be designed to activate the following output

¹³ These networks are part of a larger family of parallel activation or connectionist networks that assume computation involves simple processing units, distributed patterns of activation over the units, and numerical connection weights (Rumelhart, Hinton, & McClelland, 1986).

units, with these values: [BUS] .85 units, [TRAIN] .10, [PLANE] .05. [BUS] has the highest activity level, so it is the output of the network.

These networks can be used to simulate various aspects of performance, implementing the processing assumptions outlined above. First, if the procedure for selecting the output pattern is based on the difference in activation values (e.g., Luce choice rule), the “reaction time” of the network will be related to the absolute output probabilities (the greater the output probability, the lower the activation of units corresponding to competing patterns). Second, priming can be simulated by allowing the activation of network outputs to persist over time. This will increase the activation level of previous outputs of the system, and thus increase the output probability of patterns. To ensure that priming is temporary, the persistent activation must decay over time. Finally, errors can be generated by adding Gaussian noise (with a mean of zero) to the activation values of output units. This alters the variability of activation values, but not their means—forcing errors to respect the output probabilities. Taking the [BUS] example, if errors are caused solely by a constant increase in the variability of activation values, [TRAIN] will be more likely to overtake [BUS] than [PLANE]. Since [TRAIN] has a higher mean activation value, lower variability will be required to produce it as an error.

The spreading activation network illustrates one way to implement these cognitive functions; alternative processing mechanisms could also be used, as long they respect the properties of cognitive functions outlined above. The claims of this dissertation regarding the processing system are stated at the functional level (i.e., processing functions), not at the level of mechanisms implementing these functions (i.e., spreading activation networks).

An architecture for spoken production

Processing components

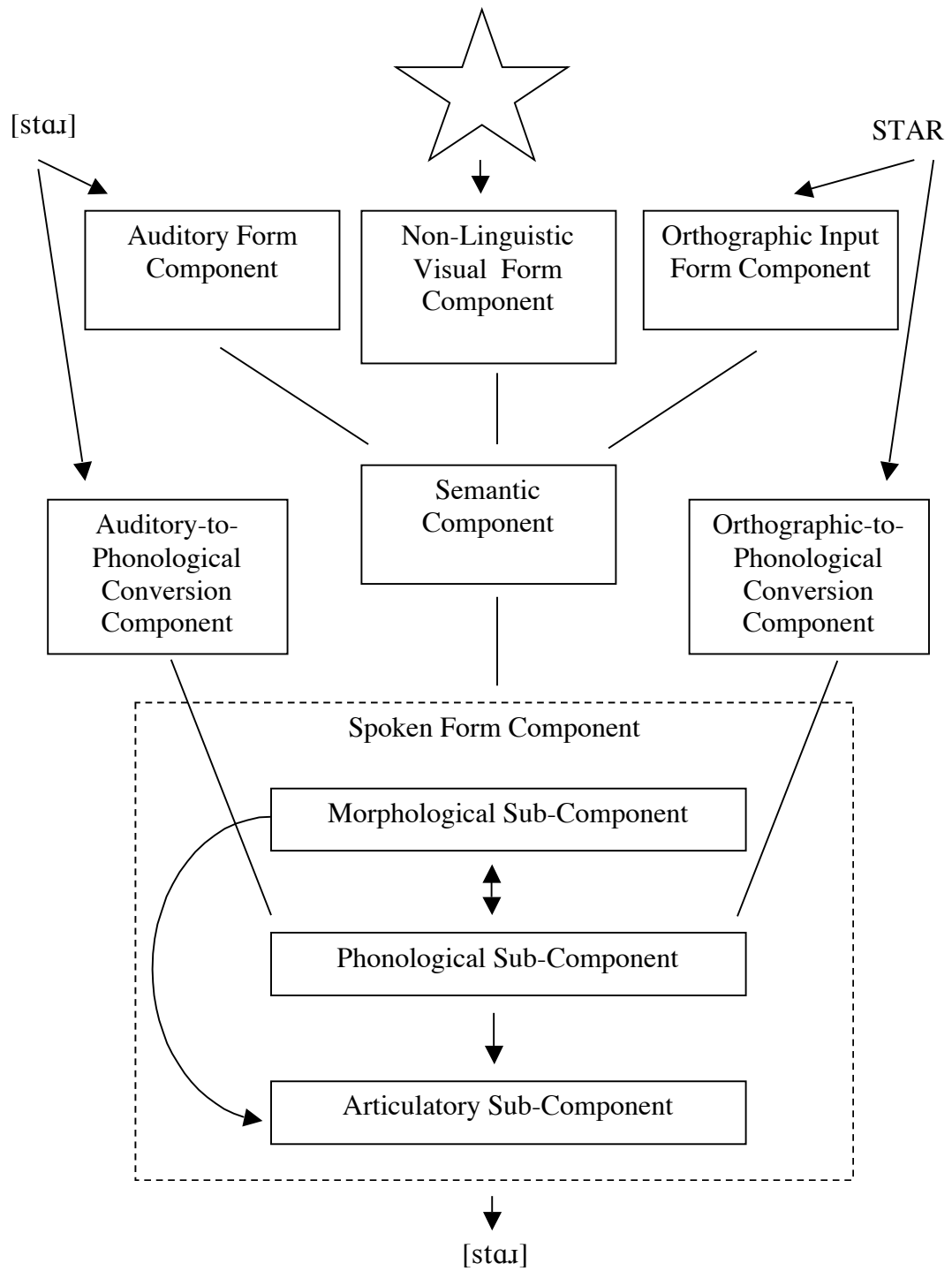
Processing components are groupings of different cognitive processes and representations. They are defined by two properties: one, the type of information (e.g., orthographic vs. phonological) that is manipulated and represented by the processes that make up the component; and two, the functional independence of processes within the component from processes manipulating other types of information¹⁴. Two sets of processes are functionally independent if the operation of one set of processes does not necessarily require the operation or use of the other set of processes. This does not mean that the two sets of processes do not interact in some fashion, but rather that such interaction is not necessary for the operation of either set of processes. Note that at this

¹⁴ The concept of processing component is similar to the concept of “modules” (Fodor, 1983). The main difference between processing components and some definitions of modules is that processes in independent components can exhibit a substantial degree of interaction. (This definition may be more in line with Fodor’s definition of modules, as opposed to other others’ interpretations of his definition; see Coltheart, 1999, for discussion.)

level of description, the particular processes and representations that make up each component are not specified.

I define 7 basic components that may be used to support performance in spoken language production tasks (e.g., picture naming, repetition, oral reading, etc.). The interaction of these components is shown in Figure 1 below. Boxes denote groups of processes (components or sub-components); arcs denote interaction between processes. Bi-directional arcs indicate feedback from one set of processes to another (see below for discussion). For unlabeled arcs, I have remained neutral as to whether feedback is present.

Figure 1. Architecture of the spoken production system.



The first 5 of these components are defined as follows:

- Auditory form component: Processes and representations that take as input representations of sound information in the external environment and identify the lexical items corresponding to that sound information.
- Orthographic input form component: Processes and representations that take as input representations of linguistic visual information (e.g., letters) that occurs in the environment and identify the lexical items corresponding to that visual information.
- Non-linguistic visual form component: Processes and representations that take as input representations of visual information (other than orthographic information) that occurs in the external environment and identify the lexical items corresponding to that visual information.
- Semantic component: Processes and representations that manipulate meaning information.
- Spoken form component: Processes and representations that retrieve sound information corresponding to lexical items and manipulate sound information to support spoken behavior (i.e., speech). This component is the focus of this dissertation; its internal structure will be discussed below.

The independence of these processing components is supported by studies showing that processing within one component does not require processing by another component (see Hillis, 2001, and Rapp, Folk, & Tainturier, 2001, for reviews). I briefly review the evidence regarding the independence of the auditory and spoken form components.

Independence of auditory form and spoken form components

The nature of the relationship between auditory form and spoken form components has been quite controversial (for a review, see Martin & Saffran, 2002). Some theorists claim that one functional component processes sound information for both input and output, while others claim that independent components exist. Theories proposing separate components can accommodate much of the evidence suggesting common processes by allowing interaction between input and output processes. It is less clear how theories proposing a single component can accommodate data supporting independence. Theoretical debates have therefore focused on this evidence.

Martin & Saffran (2002) review several different types of evidence supporting independent processes. The strongest evidence comes from dissociations between input and output processing. First, several studies describe individuals with deficits to the spoken form component (output processing) in the context of an intact auditory form component (input processing). Critics claim that the apparent dissociation in performance arises because input tasks are less demanding on the processing system than

output tasks¹⁵. In fact, close inspection of some of the reported dissociations reveals that input processing is often (but not always) impaired with respect to more difficult tasks (a similar critique has been made of studies of neurologically-intact participants). In an attempt to rule out these “processing asymmetry” explanations, other studies have documented the converse dissociation: impaired auditory form processing with intact spoken form processing. Even though many of these cases are controversial, there are some cases that suggest that spoken form processing can be selectively impaired. Hillis (2001) reviews the case of Dr. O. Dr. O lost the ability to understand spoken words. In spite of this deficit, he could understand written words (suggesting an intact semantic component), as well as repeat and discriminate spoken words (suggesting intact peripheral auditory processing). Cases such as this undermine the “processing asymmetry” account, support the existence of independent input and output components.

Lexical and non-lexical chains of processes

Note that the auditory and orthographic input form components process information that is specific to lexical items; that is, they use auditory or orthographic information to access stored representation of lexical items. These representations are then used to access semantic processes, which in turn access morphological processes in the spoken

¹⁵ See Smolensky (1996) for a similar proposal regarding the development of phonological grammars (where receptive grammatical abilities are acquired before productive grammatical abilities).

form component (see Hillis, 2001, and Rapp et al., 2001, for reviews supporting this general architecture).

The chain of processes described above is based around access of information specific to lexical items. I assume that there is also a non-lexical chain of processes. These processes manipulate and represent both words and non-words, without reference to their lexical status (see Hanley, Kay, & Edwards, 2002, for evidence supporting a non-lexical chain for auditory information, and Rapp et al., 2001, for evidence supporting a similar set of processes for orthographic information). An important link in this non-lexical chain are the auditory- and orthographic-to-phonological conversion components, defined here:

- Auditory-to-phonological conversion component: Processes that take as input representations of sound information in the external environment and relay this information to phonological processes.
- Orthographic-to-phonological conversion component: Processes that take as input representations of visual linguistic information in the external environment and relay this information to phonological processes.

Internal structure of the spoken form component

I have characterized a processing component as being made up of processes and representations that manipulate one type of information. Within this fairly broad characterization of information (e.g., auditory vs. orthographic), a processing component makes finer-grained distinctions. The next section will characterize the sub-components that make up the spoken form component. A sub-component is defined using the same criteria as a processing component; a sub-component, however, has a more specific domain of information for representation and processing. I will present evidence below supporting three distinct sub-components within the spoken form component:

- Morphological sub-component: Processes and representations that take as input semantic representations and identify the morphological units corresponding to that representation of meaning.
- Phonological sub-component: Processes and representations that transform input from the morphological sub-component and the auditory/orthographic conversion processes into a phonological representation. This component encodes phonological regularities.
- Articulatory sub-component: Processes and representations that take as input a phonological representation and produce the sequence of articulatory gestures corresponding to that representation.

I first review evidence supporting a distinction between these sub-components. Subsequently, I will discuss how these sub-components interact, focusing on the effects the morphological and articulatory sub-components have on the phonological sub-component.

Evidence for the independence of the morphological sub-component

A morpheme is the smallest meaningful unit of language. Morphemes can be free-standing elements (e.g., words like “dog”) or may only be found in the presence of other morphemes (e.g., the /s/ or /z/ at the end of “cats” or “dogs”). Evidence that the spoken production system distinguishes between morphological constituents comes from the productivity of morphological processes. For example, we can readily inflect novel words using existing morphemes: if “blinch” is a verb, I may have “blinched” yesterday. Evidence from spontaneous speech errors also supports a distinction between morphological constituents. In some errors, morphological elements are moved independently of one another. For example, in the exchange error “It pays to wait” → “It waits to pay” (where → means “produced as”), two verb roots are exchanged while the /s/ inflection remains in place (Garrett, 1984). This is not a simple sound exchange error; the /s/ inflection surfaces as the correct allomorph [s] following “wait” (as opposed to [z], the allomorph associated with “pay”).

But is morphological information represented at the level of form? Given that morphemes are associated with meanings, one may be tempted to associate such representations not with the spoken form component, but with the semantic component. Evidence that words are represented in terms of their morphological constituents at the level of form comes from individuals that make morphological errors in the context of intact semantic processing (see Allen & Badecker, 2001, for a review). For example, Allen & Badecker review the case of SJD. She exhibited normal comprehension, but made many morphological errors in spoken production¹⁶. Such a result would be uninteresting if SJD's morphological errors were spoken form errors that just happened to result in morphologically related forms. This possibility is excluded by a comparison of SJD's performance on affixed words and unaffixed homophones (e.g., bowled/bold, links/lynx). SJD performs much worse on affixed words (50% error rate vs. 20% on the unaffixed homophones). Furthermore, while SJD produces many morphological errors on affixed targets (e.g., bowled → bowling), she does not produce any pseudo-morphological errors on unaffixed homophones (e.g., no errors like bold → *bowling).

¹⁶ SJD also made similar errors in written production, suggesting that some morphological processes may be shared across both modalities. Crucial to the discussion here is the claim that morphological processes are distinct from the semantic component and interact with other sub-components of the spoken form component.

This suggests that the spoken form component represents and processes the distinction between different morphological constituents¹⁷ independent of other form distinctions.

Evidence for the independence of the articulatory sub-component

Spoken form processes must eventually result in the production of speech acts: that is, movements of the articulators to produce sound. The articulatory sub-component generates this information on the basis of phonological representations. Drawing a distinction between these representations and processes and the morphological sub-component has not been controversial; what has been less clear is how to distinguish the abstract representation of form (i.e., the phonological sub-component) from this physical realization (i.e., the articulatory sub-component; for a recent overview, of the debate in linguistics, see the papers in Burton-Roberts, Carr, & Docherty, 2000).

Experimental evidence from speech production tasks supports a distinction between articulatory and phonological representation. First, when the physical apparatus of speech production is disrupted (e.g., through insertion a bite block, or mechanical disturbance of lip position), the production system rapidly adapts itself to compensate (for reviews, see MacNeilage, 1981; Saltzman & Munhall, 1989). This compensation is driven by the attempt to produce the correct (phonological) target. For example, in a

¹⁷ Evidence suggest that not all words are represented in terms of the morphological constituents at the level of form. Case studies of individuals with morphological deficits at the level of form (such as SJD) have shown that their morphological errors are largely limited to productive morphological structures (Allen & Badecker, 2001). Similar results have been reported for experimentally elicited errors in unimpaired participants (Stemberger & MacWhinney, 1988).

study reviewed by Saltzman & Munhall (1989), participants produced word-initial /pi/ sequences. Occasionally, a participant's lower lip was unexpectedly pulled down just prior to making the lip closure for the /p/; when this occurred, the onset of voicing (laryngeal abduction) was delayed so that a voiceless stop would still be produced. The voicing gesture was dynamically updated to compensate for the disturbance of the lip. These compensatory effects suggest that underlying the real-time physical realization of speech there is an invariant speech plan. If no such plan was present, speakers would not be able to recover from significant disruption to articulatory processing. I identify phonological structures with these underlying plans that serve as input to the articulatory sub-component.

Additional evidence suggests that articulatory processes can be selectively disrupted, supporting a separation between these two processes. MacKay (1987; see also MacKay & MacDonald, 1984) argues that many individuals that suffer from intrinsic stuttering have a specific disruption to the articulatory sub-component. This is based on three observations. First, many of these individuals do not stutter during inner speech (which presumably involves phonological, but not articulatory processing). This does not appear to be attributable solely to social factors, as many stutterers report stuttering when speaking aloud to themselves. Second, stuttering rates increase as a function of the number of muscles involved in articulation, a factor that would most obviously effect

articulatory processing. Finally, many intrinsic stutterers appear to have difficulty controlling muscles used in articulation even when these muscles are used in non-linguistic tasks, again suggesting a problem in coordination and manipulation of articulators. These data suggest that the articulatory sub-component is selectively disrupted in many individuals suffering from intrinsic stuttering, supporting its independence from the phonological sub-component.

Evidence for the independence of the phonological sub-component

The phonological sub-component receives input from the morphological sub-component (as well as conversion processes) and uses this input to generate a phonological representation; this representation serves as input to the articulatory sub-component. Below, I will discuss some of the specific representations that are contained within this sub-component. One piece of evidence that this sub-component is independent of the articulatory and morphological sub-components comes from a case of an acquired language deficit (Goldrick, Rapp, & Smolensky, 1999). BON had difficulties in spoken production as the result of a left-hemisphere stroke. Her impairment could not be attributed to perceptual and/or semantic difficulties, suggesting an impairment to the spoken form component. The articulatory sub-component did not appear to be damaged, as she could correctly manipulate and control her articulators (i.e., she did not exhibit dysarthria). Further analysis of her performance supports a phonological, as opposed to

morphological, locus of her deficit. Virtually all of her errors were phonologically related word and nonwords. Her errors did not exhibit effects of any morphological variables (e.g., lexical frequency or neighborhood density¹⁸). Furthermore, her performance was significantly influenced by phonological variables (e.g., syllable position of segments; segment frequency). This suggests a deficit specifically to the phonological sub-component, independent of deficits to morphological and/or articulatory sub-components.

Interactions between spoken form sub-components

The preceding sections have motivated the independence of the morphological and articulatory sub-components from the phonological sub-component. The next few sections characterize the interactions of these sub-components, focussing on the effect each sub-component has on the phonological sub-component.

Interaction of morphological and phonological sub-components

Morphological representations serve as input to the phonological sub-component. Within the phonological sub-component, one set of processes (phonological retrieval processes) specify the phonological form of morphemes. The lexical bias effect shows that this retrieval process is interactive; the outcome of phonological retrieval is shaped by the morphological sub-component. Phonological errors include morphologically and semantically unrelated lexical items (e.g., “cat” → “hat”) as well as nonwords (e.g., “cat”

¹⁸ The number of words phonologically related to the target (see, e.g., Vitevitch, 2002).

→ “zat”). Some studies have indicated that these errors exhibit a lexical bias; they result in lexical items more often than predicted by chance. This has been reported in studies of spontaneous speech errors (English: Dell & Reich, 1981; Harley, 1984; Stemberger, 1985; but see Garrett, 1976; null result in Spanish: del Viso, Igoa, & García-Albea, 1991) as well as in speech errors in aphasic individuals (English: Best, 1996; Gagnon, Schwartz, Martin, Dell, & Saffran 1997; but see Nickels & Howard, 1995). The finding has also been observed in phonological error elicitation studies—experimental paradigms that induce speech errors (English; SLIPs paradigm: Baars, Motley, & MacKay, 1975; Dell, 1986, 1990; Humphreys, 2002; tongue twisters: Hay, Pierrehumbert, Beckman, & West, 1999; Wilshire & McCarthy, 1996; see section 3 for description of paradigms)¹⁹.

The higher-than-expected rate of word outcomes suggests that the morphological sub-component not only serves as input to the phonological sub-component, but also shapes its output. This pattern can be accounted for if the output of the phonological retrieval process is “fed back” to the morphological sub-component. The morphological sub-component can use this feedback to adjust its input to the phonological sub-component and shape its output. For example, suppose that [CAT] is the input to the phonological retrieval process. If the retrieval process is disrupted, word /hæt/ and nonword /zæt/ may

¹⁹ It has also been reported in a meta-linguistic speech production task (Carter & Bradshaw, 1984; Fowler, 1987). However, it is unclear if effects in this task derive from the phonological sub-component (e.g., Carter & Bradshaw, 1984, report effects of orthographic similarity in performance). A second source of evidence, not discussed here, is the presence of phonological neighborhood effects in speech production (Vitevitch, 2002).

both be present in the output. If this output is fed back to the morphological sub-component, it can adjust its output, increasing the presence of [HAT] in the input to the retrieval process. This modified input will increase the likelihood that [HAT] will be produced as an error. Crucially, no such support will be provided to nonword outcomes like /zæt/.

Note that this interaction does not eliminate any distinction between the morphological and phonological sub-components. I have proposed that phonological retrieval processes feed back to the morphological sub-component. I assume that there are other processes within the phonological sub-component that do not exhibit this high degree of interaction with the morphological sub-component. This limitation on interaction between the two sub-components allows for the two processes to be functionally independent (accounting for the patterns of selective impairment—such as BON—reviewed above).

Interaction of phonological and articulatory sub-components

Phonological representations serve as input to the articulatory sub-component, guiding the execution of articulation. During processing, the flow of information is unidirectional; the articulatory sub-component does not feed back to the phonological sub-component. This claim is based on two observations. First, as discussed above, the articulatory sub-component exhibits compensatory effects that derive from an underlying

invariant speech plan (i.e., a phonological representation). The fact that such a representation is “invariant” means that it is insensitive to the demands of the articulatory sub-component. If feedback was present, we would expect that compensation would be limited; eventually, disruption of the articulatory sub-component would feed back and disrupt the phonological sub-component (i.e., the disruption would alter the phonological target). This does not appear to be the case. MacKay’s (1987) observations regarding intrinsic stuttering support a similar conclusion. Unlike normal speech errors (many of which arise in the phonological sub-component), stuttering errors cannot be voluntarily “corrected;” they disrupt speech processing for extended periods of time. If feedback from the articulatory sub-component was present, it would seem reasonable to expect that alternative phonological structures would be automatically generated to circumvent articulatory disruptions (instead, alternatives must be consciously generated). Overall, it appears that during spoken production the articulatory sub-component must do everything it can to respond to the demands of the phonological sub-component, while the phonological sub-component is completely unresponsive to articulatory demands.

Although this relationship appears to hold during on-line speech production, there is abundant evidence that articulatory demands constrain and shape phonological regularities (Archangeli & Pulleyblank, 1994; Boersma, 1998; Hayes, 1999; Lindblom, 2000; but see Anderson, 1981). This may be due to the mechanisms by which

regularities are acquired (as proposed by Hayes, 1999). Although the exact mechanisms are unclear, it is clear that feedback in the adult spoken production system does not cause such effects.

Interaction of morphological and articulatory sub-components

I briefly note that there is some evidence that the morphological and articulatory sub-components interact directly. Jurafsky, Bell, & Girand (in press) present evidence that homophones (morphologically distinct items that have identical phonological representations; see Dell, 1990, for discussion) are produced with distinct articulations. Since a morphological distinction can influence articulatory processing in the absence of phonological distinctions, I assume that the morphological and articulatory sub-components directly interact.

Summary: Architecture for spoken production

The preceding sections have laid out the basic architecture of the spoken production system, focussing on the role of phonological processes. I have described how these processes interact with perceptual and semantic processing components, as well as their interaction with other sub-components within the spoken form component. Laying out this architecture allows us to interpret experiment results. Specifically, it provides a framework for generating alternative accounts of the source of experimental results; by

ruling these out, we can infer that the effects arise within the phonological sub-component.

It should be noted that the processing architecture described above is far from a complete characterization of the processes involved in all spoken language tasks. One prominent omission is syntactic processes—that is, the processes and representations that manipulate and construct meaningful sequences of lexical items. Clearly such processes form part of our spoken production system. The tasks that I will discuss here involve the production of single words and nonwords, as well as the production of meaningless sequences of words and nonwords. I therefore assume that syntactic processes do not influence processing during these tasks.

Internal structure of the phonological sub-component

The focus of this investigation is the types of phonological regularities that are encoded by spoken production processes. I assume that phonological regularities are encoded by the output probability of structures within the phonological sub-component. Specifically, regular structures have higher output probability than irregular structures. Evidence for this assumption will be discussed in more detail in section 3. Based on this assumption, the structure of the phonological sub-component constrains the types of regularities that could be encoded. With respect to granularity, regularities can be encoded only if there is a corresponding level of representation within the phonological

sub-component. For example, if there is no sub-segmental level of representation, the phonological processing function cannot vary output probabilities as a function of sub-segmental regularities. The ability to encode contextual restrictions can also be influenced by the structure of the phonological sub-component. For example, if the processing function cannot make simultaneous reference to syllabic and sub-segmental representations, it cannot encode supra-segmental contextual restrictions on sub-segmental structures (e.g., a regularity banning voiced stops in coda position).

In the next few sub-sections, I argue that the phonological sub-component can encode the full range of regularities defined by the granularity feature. I will first argue that the phonological sub-component has three levels of representation supporting the encoding of regularities at different levels of granularity:

- Sub-segmental representations: The most basic level of phonological representation, expressing a limited number of distinctions related to articulatory dimensions.
- Segmental representations: A level of representation that organizes sub-segmental representations into groups.
- Supra-segmental representations: A level of representation that organizes segments into groups and expresses prosodic structure.

To argue for independent representations corresponding to these different levels, I present evidence suggesting that these representations can independently influence output probabilities. Subsequently, I present studies suggesting that the phonological sub-component can encode contextual restrictions. Specifically, I will show that output probabilities can be influenced by multiple representational levels simultaneously.

Evidence for phonological representations

I have characterized a cognitive process as a function that assigns probabilities to mappings from input representations to output representations. How can we infer the structure of representations within the phonological processing function?

One source of evidence will be speech production errors. Recall that if a cognitive process is working under less than ideal conditions, it can generate a variety of output patterns (not just the pattern with the highest output probability—the optimal output pattern). By examining the differences between the optimal and non-optimal outputs, we can learn about the structure of the function’s representations. For example, if a lexical semantics process generates lexical items as outputs (e.g., [BUS], [TRAIN], etc.), its non-optimal outputs will differ from the optimal output in terms of lexical identity. We might observe errors such as “bus” → “train.” By observing speech production errors where one whole word is replaced by another whole word, we can infer that there is some processing function in which distinct patterns correspond to distinct lexical items. Of

course, it is important to establish that some other representation could not give rise to the errors. For example, whole word errors that only differ in terms of a single segment (e.g., “cat” → “hat”) might arise in a segment-based representation purely by chance (this would predict that nonword errors like “cat” → “zat” would also be observed). Furthermore, for the purposes at hand, we must determine if the errors arise in the phonological sub-component. Spoken production errors can be produced by errors at many different levels of processing. We must be able to eliminate error sources in processes outside of the phonological sub-component.

The second type of data will be reaction times in priming paradigms. Recall that if a cognitive process has recently produced an output, the output probabilities associated with that output representation will increase (and its associated processing time will decrease). By examining what types of outputs can be used to prime other outputs, we can uncover the structure of representations. For example, if some representational level distinguishes outputs only in terms of their stress pattern, then we should be able to prime an output with an output that shares only its stress pattern (e.g., “motor” should be primed by words like “livid”). Of course, the same concerns apply as above; we must be certain that some other representational level cannot account for the priming effect, and that it arises within the phonological sub-component.

Evidence for independent sub-segmental representations

Evidence from speech errors supports the claim that sub-segmental representations are an independent level of representation (but see Shattuck-Hufnagel & Klatt, 1979, 1980). Guest (2001) had English-speaking participants quickly read aloud sequences of four consonant-vowel (CV) nonwords (e.g., “gee tay vu nai”) and recorded their speech errors involving consonants. To look for unambiguous sub-segmental errors, Guest used the following coding scheme. If participants replaced a consonant with another consonant in the target string, it was recorded as a segmental error. Sub-segmental errors fit the following two criteria: one, the participant used a single sub-segmental feature from one target consonant to replace a single sub-segmental feature on another consonant (e.g., for the example above, combining velar /g/ with unvoiced /t/ to create “kay” instead of “tay”); two, the resulting consonant was not in the target string (e.g., /k/ is not in “gee tay vu nai”). Using these criteria, Guest found that 33% of the errors produced by the participants were sub-segmental errors.

Do these errors arise within the phonological sub-component? Participants read the sequence once slowly to verify that the nonwords had been perceived correctly. The sequence was not visible during the fast repetitions, so there is a possibility that some errors arise as a result of memory failures (and thus may be due to factors outside the spoken form component). However, the sequence has a very low memory load, making

this explanation less than plausible. Within the spoken form component, the use of nonwords argues against a purely morphological basis for these speech errors.

Furthermore, errors in tongue twisters appear to arise prior to articulatory processing. To eliminate the role of the articulatory sub-component, Dell & Repka (1992) had participants silently produce tongue twisters and monitor their errors in inner speech. This experimental manipulation was most likely successful in eliminating the articulatory sub-component; Wheeldon & Levelt (1995) report that participants' monitoring of their own speech is sensitive to phonological, not articulatory, representations. Although fewer errors were found in the inner speech monitoring condition (as compared to overtly produced tongue twisters)²⁰, the distribution of errors was qualitatively similar (e.g., initial consonant errors occurred more frequently than medial or final errors). These results suggest that many errors in the tongue twister task arise prior to the articulatory sub-component). In light of these studies, Guest's results supports the existence of an independent sub-segmental representation specifically within the phonological sub-component.

Evidence for independent segmental representations

Evidence from speech errors supports the segment as an independent level of representation within the phonological sub-component. Many studies have observed a

²⁰ The reduced number of errors (relative to that observed in overt production) may be the result of the fallibility of the speech monitoring system (Postma & Noordanus, 1996).

“repeated phoneme” effect on speech errors. If a segment is repeated in a sequence, there is an increased likelihood of errors in the sequence. For example, a sequence like “time line,” where the vowel /aɪ/ is repeated, is more likely to have errors than a sequence like “heat pad”, where two different vowels (/i/ and /æ/) are used. This has been found both in analyses of spontaneous speech errors (Dutch: Nooteboom, 1969; English: Dell, 1984; MacKay, 1970; Shattuck-Hufnagel, 1979; Vousden, Brown, & Harley, 2000; German: MacKay, 1970; but see Motley, 1973, for a null result in English), as well as in experimentally-induced errors (SLIPs paradigm, English speakers: Dell, 1984, 1986; but see Camden, 1980, Appendix A; see section 3 for description). Stemberger (1990) attempted to assess if this effect was influenced by sub-segmental similarity. He examined the similarity of the two words involved in English spontaneous speech errors on word-initial singleton consonants. Replicating previous studies, he found that repeated segments increased error rates above chance levels. The novel result was that repetition of highly similar segments had no effect on error rates. That is, two words with minimally different segments (e.g., /t/ and /d/) did not show error rates above chance levels. In contrast, words with two identical segments showed error rates significantly greater than chance. This suggests that the repeated phoneme effect derives from repetition of segmental, not sub-segmental, units.

Independent segmental representations are also supported by priming results. Roelofs (1999) showed that segment overlap produced priming in Dutch speakers' picture naming, whereas no priming was observed with minimally different segments.

Can the repeated phoneme and segmental priming effects be localized to the phonological sub-component? With respect to the repeated phoneme effect, Stemberger (1990) examined spontaneous speech errors, which are notoriously difficult to attribute to a particular processing component. However, given that the repeated phoneme effect is found under experimental conditions that attempt to minimize the influence of other processing components (Dell, 1984), it is likely that Stemberger's results pertain to spoken form processing. The fact that the errors are insensitive to fine-grained aspects of structure argues against an articulatory locus for this effect (i.e., if the effect is insensitive to sub-segmental similarity, it is not influenced by articulatory similarity). Although there is no control for morphological effects, it is clear that the units involved in these speech errors are smaller than morphemes. These results appear to suggest an independent role for segmental representations within the phonological sub-component.

Evidence for independent supra-segmental representations: Syllabic structure

The strongest support for independent syllabic representations comes from Sevald, Dell, & Cole (1995). They used a priming paradigm to investigate the role of syllable structure. English-speaking participants were asked to repeat pairs of nonwords. One

member of the pair was a monosyllable (e.g., KILP) and the second a disyllable (e.g., KILPNER). Across conditions, the two nonwords shared segmental structure but differed in terms of syllabic structure. In one set of conditions, the two nonwords shared syllable structure (e.g., KILP, KILPNER share the first syllable KILP). In the second set, they differed (e.g., KILP, KILPLER, where the first syllables are KILP and KIL, respectively). Participants were shown a pair of nonwords for four seconds and then asked to repeat the sequence as many times as possible within another four second interval. Pairs that shared syllable structure were repeated more quickly and accurately than those that did not share syllable structure. Subsequent experiments showed that a comparable size effect was found when the two members of the pair just shared syllable structure and did not share segments. Similar priming results (in different paradigms) have also been claimed to support independent syllabic representations in a variety of languages (Dutch: Meijer, 1996; but see Schiller, 1998; Roelofs & Meyer, 1998; English: Ferrand, Segui, & Humphreys, 1997; but see Schiller, 2000; French: Ferrand & Segui, 1998, experiment 2; Ferrand, Segui, & Grainger, 1996; Spanish: Costa & Sebastian-Gallés, 1998).

The results reported by Sevald et al. (1995) are most probably due solely to spoken form processing. Perceptual processes are excluded because participants view the sequences for a considerable length of time before repeating them. Within spoken form

processing, these results most likely arise in the phonological sub-component. The use of nonwords minimizes the role of morphological processing; an articulatory locus is ruled out because the effect is insensitive to lower-level properties of the stimuli (e.g., if the effect is not sensitive to segmental identity, it is presumably not a function of articulatory similarity). In sum, this priming study provides strong support for an independent representation of syllabic structure.

Evidence for independent supra-segmental representations: Metrical structure

Speech errors provide evidence that metrical structure is independent of segmental and sub-segmental structure. In most situations when a stressed and unstressed vowel exchange, the lexical stress does not shift with the vowel. For example, Stemberger (1983) reports the error “people were” → “purple...”; here, unaccented schwa+r replaces /i/, but accent remains on the first syllable (in Stemberger’s corpus, 32/36 exchange errors involving vowels of different stress levels fit this pattern). Based on these results, I assume that metrical structure is represented independently from segmental and sub-segmental representations.

Interim summary: Independent levels of phonological representation

In the preceding sections, I have reviewed evidence supporting three different types of independent phonological representations.

- Sub-segmental representations: The most basic level of phonological representation, expressing a limited number of distinctions related to articulatory dimensions (Guest, 2001).
- Segmental representations: A level of representation that serves to organize sub-segmental representations into groups (Stemberger, 1990).
- Supra-segmental representations: A level of representation that organizes segments into groups and expresses prosodic structure (syllabic structure: Sevald et al., 1995; metrical structure: Stemberger, 1983).

Simultaneous reference to multiple phonological representations

The preceding sections motivate the presence of three distinct levels of phonological representation, supporting the ability to encode regularities at different granularities. The second question to address is whether the phonological sub-component can encode contextual restrictions. As noted in section 1, some regularities are specific to particular contexts. For example, the Lexical Distribution theory claims that phonological regularity depends on simultaneous reference to syllable structure, word position and metrical position. The next few sections demonstrate that such regularities could be encoded; output probabilities can simultaneously depend on multiple levels of representation.

Simultaneous reference to supra-segmental and segmental representations: Syllabic structure

Two studies show how shared syllabic structure can influence segmental errors in spontaneous speech. As discussed above, Stemberger (1990) analyzed the repeated phoneme effect in English spontaneous speech errors. Stemberger also found a repeated syllable structure effect on segmental errors; if two words shared syllabic structure, there was an increase in segmental error rates relative to chance. Additional evidence comes from Hartsuiker (2002). He found that the syllable structure produced by deletion errors in Dutch and Spanish was identical to that of many (50% or more) syllables in nearby contexts, suggesting that nearby syllable structure influences segmental errors.

In addition, several priming studies have suggested an interaction of segmental and supra-segmental representations. These studies suggest that priming effects are found only when both segmental and supra-segmental structure overlap (Dutch: Cholin, Schiller, & Levelt, 2002; Roelofs & Meyer, 1998; French: Ferrand & Segui, 1998, experiment 1).

Simultaneous reference to segmental and supra-segmental representations: Metrical structure

Studies in several different languages have noted that interacting segments in spontaneous speech errors tend to occur in syllables of the same stress (Dutch:

Nooteboom, 1969; English: Boomer & Laver, 1969; Garrett, 1975; MacKay, 1969; Shattuck-Hufnagel, 1983, 1987; German: MacKay, 1969). Similar effects have been reported in tongue twister experiments (English: Shattuck-Hufnagel, 1987, 1992; Frisch, 2000; Wilshire, 1999). However, stress and word position are often confounded. Shattuck-Hufnagel (1987, 1992) unconfounded these variables in a series of tongue twister experiments. She found that stress independently contributed to error rates²¹, supporting an effect of metrical structure on segmental errors.

In addition, some priming studies find effects only when primes share both segments and stress (Dutch speakers: Meijer, 1994; Roelofs & Meyer, 1998). These results appear to require that output probabilities be simultaneously sensitive to both segmental and metrical structure.

Simultaneous reference to segmental and sub-segmental representations

Sub-segmental representations encode the similarity of segments. This similarity appears to interfere with the planning and execution of speech; the more similar two

²¹ Shattuck-Hufnagel (1992) reports the effects of word position and stress are additive, further supporting their independence.

segments are, the more likely it is that they will interact in speech errors²². This result has been replicated by a number of studies of spontaneous speech errors in a number of different languages (Arabic: Abd-El-Jawad & Abu-Salim, 1987; Dutch: Nootboom, 1967; Van den Broecke & Goldstein, 1980; English: Boomer & Laver, 1968; Frisch, 1996, 1997; Fromkin, 1971; Garrett, 1975; Levitt & Healy, 1985; MacKay, 1970; Shattuck, 1975; Shattuck-Hufnagel & Klatt, 1979; Stemberger, 1991; Van den Broecke & Goldstein, 1980; Vousden et al., 2000; German: Berg, 1991a; MacKay, 1970; Van den Broecke & Goldstein, 1980; Spanish: García-Albea, del Viso, & Igoa, 1989; Swedish: Söderpalm, 1979) Such effects have also been observed in experimentally-induced²³ speech errors (English: Kupin, 1982; Frisch, 1996, 2000; Levitt & Healy, 1985; Stemberger, 1991; Wilshire, 1998, 1999). Finally, similar findings have been reported in studies of individuals with acquired language disorders in a variety of languages (see Blumstein, 1998, for a review). The findings from spontaneous speech errors and many

²² Not all dimensions of sub-segmental similarity are equal. This was noted in post-hoc analyses of spontaneous speech errors in various studies showing that some sub-segmental representations are more often involved in segmental exchanges than others (relative to some chance level: MacKay, 1970; Shattuck, 1975). Why some dimensions are more often involved than others is unclear. One particular hypothesis is that dimensions of similarity that establish a contrast between speech sounds in the language (e.g., for English, voicing in stops) appear to be more heavily weighted than dimensions that do not establish contrasts (e.g., for English, voicing in nasals). Some studies of spontaneous and experimentally elicited speech have supported this hypothesis (German, spontaneous errors: Berg, 1991a; English, spontaneous errors: Frisch, 1996, 1997; English, tongue twister paradigm: Frisch, 1996; English, SLIPs paradigm: Stemberger, 1991).

²³ Kupin (1982), Frisch (1996, 2000), Levitt & Healy (1985, Experiment 2), and Wilshire (1998, 1999) induced errors in tongue twisters. Levitt & Healy (1985, Experiment 1) and Stemberger (1991) induced errors using the SLIPs paradigm. Effects of sub-segmental similarity were also reported by Fowler (1987) using a paradigm developed by Carter & Bradshaw (1984); see footnote 19 for discussion of the uncertain locus of effects in this procedure.

studies of language disorders are a bit difficult to interpret, due to the uncertainty as to where in the processing system such effects arise. However, we can be reasonably confident that the experimentally-induced errors arise in spoken form processing. Sub-segmental similarity interference between two segments has also been seen in a response-priming paradigm (see below; Meyer & Gordon, 1985; Yaniv, Meyer, Gordon, Huff, & Sevald, 1990).

Simultaneous reference to all levels of structure

Results from Yaniv et al.'s (1990) study also suggest that effects can span all three levels of representation. Yaniv et al. asked English-speaking participants to prepare to produce a single syllable (e.g., “peak”, the ‘primary response’) as quickly as possible following a signal. They were also given a second syllable (e.g., “pick,” the ‘secondary response’), and occasionally asked to produce it instead of the primary response. Yaniv et al. compared response times for secondary responses that were sub-segmentally similar to the primary response (e.g., “peak” and “pick” have similar vowels /i/ and /ɪ/) to secondary responses that were dissimilar to the primary response (e.g., “peat” and “pat” have very different vowels). The first finding (referenced above) was that there were longer latencies and more errors when the secondary response was sub-segmentally similar to the primary response (as compared to cases where the secondary response was dissimilar). In addition, Yaniv et al. found effects of supra-segmental structure. The sub-

segmental similarity effect for the vowel was greatly reduced when the primary and secondary responses did not share the final consonant in the syllable (e.g., peat-pick showed much less interference than peak-pick). No such effect was found for differences in the first consonant in the syllable (e.g., peak-tick, with different initial consonants, showed just as much interference as teak-tick, with identical initial consonants). Thus, sub-segmental similarity effects on segmental structure are modulated by differences in supra-segmental structure (namely, the rime). This interference effect relies on references to all three levels of phonological representation.

Summary: Interaction between phonological representations

In both speech error and priming studies, we find that the phonological sub-component can be simultaneously sensitive to:

- Segmental and supra-segmental information (Syllable structure and segmental structure: Stemberger, 1990. Metrical structure and segmental structure: Shattuck-Hufnagel, 1987, 1992).
- Sub-segmental and segmental information (Supported most strongly by: Frisch, 1996, 2000; Kupin, 1982; Levitt & Healy, 1985; Meyer & Gordon, 1985; Stemberger, 1991; Yaniv et al., 1990).
- Sub-segmental, segmental, and supra-segmental information (Yaniv et al., 1990).

These results suggest that the phonological sub-component could encode contextual restrictions on regularities; it is capable of being sensitive to a conjunction of two or three of these levels of representation.

Although these studies suggest that the processing can be simultaneously sensitive to multiple levels of structure, the results reported in the previous section suggest that processing may also be insensitive to multiple levels of structure (as each level of representation is independent). These results are not inconsistent; the spoken production system is capable of exhibiting both types of effects. For example, the interaction of morphological and phonological sub-components was discussed above. Each sub-component can be independently disrupted, suggesting that the two sets of processes and representations are independent. We also find evidence that under certain situations, the two sub-components interact with one another. To account for both of these findings, we must assume that the interaction between these components is restricted; phonological retrieval processes exhibit interaction with morphological processes, while other phonological processes do not. A similar solution can be found for the phonological sub-component; by limiting interaction, the system can support both interactive and independent effects.

Summary: Spoken production processing

This section has examined the spoken production system at multiple levels of description. At the highest level, spoken production processing involves the interaction of perceptual, semantic and production components. Within the spoken production component, there is division into three sub-components (morphological, phonological, and articulatory) that exhibit varying degrees of interaction. Within the phonological sub-component, there are 3 types of representations (sub-segmental, segmental, and supra-segmental). Phonological processing can make reference to any single representational level, as well as to any conjunction of these levels. It is within this framework that we consider the encoding of phonological regularities.

SECTION 3:

PREVIOUS STUDIES OF THE ENCODING OF PHONOLOGICAL REGULARITIES

Does the spoken production system encode phonological regularities? The preponderance of evidence suggests that this is indeed the case; the output probabilities of phonological structures that are well-formed (hereafter, “regular” structures) are higher than those of those that are ill-formed (“irregular” structures). This is most certainly true for categorical within-language regularities. With respect to other regularities, the evidence is mixed. Some studies suggest that gradient within-language regularities and/or cross-linguistic regularities are encoded. Data supporting the encoding of these various types of regularities will be reviewed first. Subsequent sections review other studies that have failed to find an effect of either gradient within-language or cross-linguistic regularities, as well as studies that have found anti-regularity effects (where the output probabilities of irregular structures are higher than those of regular structures).

It should be noted that several types of (potentially relevant) studies are excluded from this review. First, I do not include studies of “wordlikeness” or “well-formedness” judgments. These involve tasks where participants are asked to judge the acceptability of a given form (either orthographic or phonological); e.g., “How much like an English word is zlorm?” or “How acceptable is zlorm as an English word?” There are several

concerns about whether these studies reveal information specifically about the phonological sub-component. First, these judgments appear to be influenced not only by phonological properties (e.g., phonological regularity) but also by lexical properties (e.g., neighborhood density) of the forms (Bailey & Hahn, 2001; Frisch & Zawaydeh, 2001). Thus, the judgments may reflect the influence of both the phonological and morphological sub-components. Furthermore, the extent to which the results reflect perceptual properties of the forms is unknown, making it unclear whether these tasks reveal properties of the spoken form component specifically. For similar reasons, I exclude studies that make use of other meta-linguistic tasks (e.g., blending two nonwords together: Treiman, Kessler, Knewasser, Tincoff, & Bowman, 2000; transforming one segment of a nonword into another: MacKay, 1978). It is unclear what processes are used to perform these tasks.

I also do not review studies involving the influence of phonological regularities on children's productions (e.g., Beckman & Edwards, 2000). Instead, I focus on the encoding of phonological regularities by the fully-developed, adult production system.

Evidence supporting encoding of regularities

Encoding of categorical within-language regularities

The first indication that phonological regularities are encoded in speech production processes comes from spontaneous speech errors. Phonotactically illegal structures are

those that violate the categorical phonological regularities of a particular language. For example, any word that starts with /ŋ/ is phonotactically illegal in English. Numerous studies in various languages have reported that spontaneous speech errors do not violate categorical phonological regularities; that is, speech errors rarely result in phonotactically illegal structures (Arabic: Abd-El-Jawad & Abu-Salim, 1987; English: Boomer & Laver, 1968; Fromkin, 1971; Garrett, 1975; MacKay, 1972; Motley, 1973; Stemberger, 1983; Vousden et al., 2000; Wells, 1951; German: MacKay, 1972; Mandarin: Wan & Jaeger, 1998). However, some researchers have questioned how infrequently phonotactically illegal errors occur (Hockett, 1967). In fact, some studies have found significant numbers of these errors in tongue twister experiments (Butterworth & Whittaker, 1980; Laver, 1980). Nevertheless, in these studies phonotactically illegal errors are vastly outnumbered by phonotactically legal ones (but see Mowrey & MacKay, 1990). This strong bias in speech errors suggests that regularities defining phonotactically legal structures are encoded by the spoken production system.

Of course, it is difficult to ascertain what processes are contributing to spontaneous speech errors. The phonotactic regularity effects may arise from the influence of processes other than the phonological sub-component. Studies of individuals with acquired language deficits eliminate some of these alternatives. For example, Hanlon & Edmondson (1996) review the case of JL. She had great difficulty with a variety of

comprehension as well as spoken production tasks following a stroke. Her speech consisted of “fluent strings of phonemes...with virtually no intelligible utterances (p. 201)” and containing “virtually no lexical items of English (p. 208).” Although perceptual, semantic and morphological processes appeared to be damaged, her speech was phonotactically legal, obeying the categorical within-language regularities of her dialect of English (Southern United States). These data suggest that the tendency for speech errors to be phonotactically legal sequences does not result from processes preceding the phonological sub-component. However, since JL’s articulatory sub-component appeared to be intact²⁴, the articulatory sub-component may contribute to this effect.

The encoding of regularities characterizing phonotactically legal sequences can occur even in adulthood. Dell, Reed, Adams, & Meyer (2000) asked adult participants to read aloud sequences of four CVC syllables. These syllables reflected not only the regularities of English (e.g., /ŋ/ occurred only in coda as in “meng”) but also regularities not found in English. For example, in one condition, the consonant /f/ occurred only in onset position—syllables like “fem” were presented, but others like “mef” were not. Speech errors were then induced by having the participants read the sequences quickly. Dell et al. found that the errors respected the regularities of English (e.g., no erroneously

²⁴ With respect to the articulatory sub-component, JL exhibited normal patterns of utterance-final vowel lengthening and frequency declination (i.e., the fundamental frequency of her speech declined across each utterance at a rate comparable that of normal speakers).

produced /ŋ/ segments occurred in onset) as well as the regularities specific to the experimental syllables. For example, when /f/ is restricted to onset, approximately 3% of the erroneously produced /f/ segments occurred in coda (in comparison, for segments that were not associated with a regularity, nearly 30% of the erroneously produced segments violated target syllable position)²⁵. This finding was not specific to /f/ in onset position. Across conditions, participants in Dell et al.'s experiments were able to encode regularities associating any one of four consonants /f,s,g,k/ to either syllable margin position (onset or coda). In a final experiment, Dell et al. showed that participants could also encode regularities involving specific consonant-vowel combinations (e.g., /f/ precedes /ɛ/ but not /ɪ/, whereas it can follow either)²⁶.

The effects seen in Dell et al.'s study appear to arise within the phonological sub-component. As with other tongue twister experiments, they appear to arise in the spoken form component (not in perceptual processes). Within this component, morphological effects do not appear to be driving the encoding of regularities. Most of the stimuli were nonwords, which lack morphological representations—at least prior to the experiment. Perhaps participants developed new morphological representations for the nonwords in

²⁵ Dell et al. (2000) do not report results specific to /f/ in onset; these figures are taken from collapsed results for /f/ and /s/.

²⁶ Similar results are reported by Onishi, Chambers, & Fisher (2002; see also Chambers, Onishi, & Fisher, 2002). They had participants listen to sequences respecting experiment specific regularities. After this exposure, participants initiated repetition of unstudied items respecting the regularities more quickly than repetition of unstudied items that did not respect regularities. As discussed below, it is unclear whether effects observed in repetition tasks (such as this one) derive from production (as opposed to perceptual) processes.

the stimulus set; a lexical bias for these new “words” could then account for the findings. To argue against this possibility, Dell et al. showed that errors were just as likely to result in syllables that did not appear in the stimulus set as syllables that did appear in the stimulus set (as long as the syllables respected the regularities of the stimulus set). Thus, those syllables that were “words” with respect to the stimulus set were no more likely to be produced than “nonwords,” arguing against a lexical bias account. With respect to the articulatory sub-component, note that participants are capable of encoding structure-sensitive regularities (e.g., encoding /f/ precedes /ε/ but not /ɪ/ requires sensitivity to supra-segmental structure). This suggests (but does not require) that the effects arise at a more abstract level of processing than the articulatory sub-component.

In sum, these studies suggest that the phonological sub-component encodes regularities that characterize phonotactically legal sequences. However, it is unclear at what level(s) of granularity these regularities are encoded. For example, in Dell et al.’s (2000) final experiment, participants were able to encode regularities that were sensitive to the identity of the following vowel. What regularities were encoded in the previous experiments? One possibility is that regularities about syllabic context (e.g., /f/ occurs in onset, not in coda) were encoded; alternatively, only consonant-vowel context were encoded (e.g., /f/ precedes /ε/, not /ɪ/). The final experiment shows that the second type

of regularity can be encoded, but this does not entail the first type of regularity cannot be encoded.

Encoding of gradient within-language regularities

Within the set of phonotactically legal structures of a language, there are also regularities. Some syllable structures are more frequent than others; some segments co-occur more frequently than others. Studies claiming to show that gradient within-language regularities are encoded are reviewed below, grouped by experimental methodology.

Repetition studies

Several studies examine the performance of adult participants who are asked to repeat spoken forms. In a series of studies, Vitevitch, Luce and colleagues (Vitevitch & Luce, 1998, 1999; Vitevitch, Luce, Charles-Luce & Kemmerer, 1997) examined the effect of a gradient within-language regularity (based on phonotactic probability) on latency in nonword repetition tasks in English-speaking adults. Vitevitch et al. classified syllables as having high phonotactic probability if they satisfied the following conditions: one, the syllable was made of segments with high positional token frequency; two, the syllable was made up of biphones with high positional token frequency. They found that repetition of nonwords with high probability syllables was initiated more quickly than repetition of nonwords with low probability syllables. Similar effects are reported by

Munson (2002) for biphone sequences in word internal (i.e., VCCV) contexts. Repetition of high frequency sequences was initiated more quickly, and was more accurate, than repetition of low frequency sequences.

The major concern about this task is that it conflates perception and production effects. For example, it could be that construction of a perceptual representation is slowed for low frequency structures, and that this effect increases the response latency. This interpretation is, in fact, favored by Vitevitch, Luce and colleagues. In support of this conclusion, they demonstrate that the same results occur in “pure” recognition tasks (e.g., same-different judgment). Since the production component is (presumably) eliminated in such tasks, any latency effects derive solely from the construction and evaluation of the perceptual representation. In sum, results from repetition suggests the existence of some form of sub-lexical regularity effects, but it is unclear as to whether this influence occurs in perception, production, or both.

Paired associates

Levelt & Wheeldon (1994) used a paired-associates paradigm to examine whether the spoken production system of Dutch speakers encodes a gradient within-language regularity based on syllable frequency. Participants were trained to associate symbols (e.g., //) with words from a list (e.g., “apple”). On test trials, the symbols were presented and the latency to initiate production of the associated word was measured. Lexical

frequency, number of phonemes and initial phonemes were controlled across high and low frequency syllables. Levelt & Wheeldon found that words composed of high frequency syllables have shorter latencies than words composed of low-frequency syllables. It should be noted that at least one study has failed to replicate these results (experiments by Levelt & Meyer, reported in Hendriks & McQueen, 1996).

Note that, in contrast to the repetition task, it is hard to interpret these regularity effects as deriving from perceptual processes (as the stimuli were not phonological). On the other hand, it is not entirely clear what perceptual and memorial mechanisms are used to encode each abstract symbol set and its association with the target. For example, lexical frequency effects were found in the experiment, suggesting a role for the morphological sub-component. Another concern is that the effect could be driven by articulatory properties of the high and low frequency syllables. With these caveats, we can tentatively localize these effects to the phonological sub-component.

Speech errors in SLIPs and tongue twister tasks

Levitt & Healy (1985) used two experimental speech error induction tasks to examine the effects of within-language regularities on spoken production. First, they made use of the SLIPs procedure with CV syllables (e.g., reading pairs like “ta si”). In the SLIPs paradigm (Spoonerisms of Laboratory Induced Predisposition: Baars, 1992; Baars & Motley, 1974; Motley, 1986; Motley & Baars, 1976), pairs of words or nonwords are

presented to participants, one pair at a time. On some trials, the participant is cued to produce the last pair of words as quickly as possible. Exchange errors are induced by priming subjects to reverse the order of the initial consonants of a critical pair of words. Before critical pairs, participants are shown several pairs of words that share the same initial consonant-vowel sequence²⁷ in the opposite order (e.g., for critical pair “tag sin,” prime pairs might be “sap tiff” and “sass tick”). After these primes are shown, participants see the critical pair and are cued to produce it as quickly as possible. In this situation, participants often produce exchange errors (e.g., “tag sin” → “sag tin”; but see Robins, 1980; Sinsabaugh & Fox, 1986). Levitt & Healy also made use of a tongue twister paradigm with a subset of the CV syllables used in the SLIPs experiment (e.g., reading four syllable sequences like “ta si sa ti”).

Levitt & Healy examined the errors for effects of segmental frequency. Both the SLIPs and tongue twister paradigms are intended to produce contextual errors based on other segments in the target sequence (e.g., “ta si” → “sa ti”). However, non-contextual errors can also occur in this task (e.g., “ta si” → “ta zi”). The first analysis examined both kinds of errors. Levitt & Healy selected pairs of target segments where one was high frequency (e.g., /t/) and the other low frequency (e.g., /tʃ/). They found that high frequency phonemes substituted for low frequency phonemes more often than the

²⁷ In some variations, the vowel is not shared (e.g., Camden, 1980), nor is the initial consonant of both words presented in each prime pair (Baars, 1992; Baars & Motley, 1974; Motley, 1986; Motley & Baars, 1976).

reverse. This suggests that errors are biased towards more frequent segments, supporting the encoding of gradient regularities based on within-language segment frequency.

In a second analysis, Levitt & Healy compared how context influenced the bias toward more frequent segments. Context appears to be more important for errors in which high frequency segments are replaced by low frequency segments (i.e., errors violating the segment frequency-based regularity). In the SLIPs task, 58% of the high frequency → low frequency errors occurred when the low frequency segment was present in the surrounding context, whereas only 42% of the reverse errors occurred when the high frequency segment was present in the immediate context. (The same is found for the tongue twister task: 50% of high→low errors are contextual vs. 37% for the reverse.) Thus, the presence of errors that do not reflect phonological regularities is highly dependent on context. This suggests that regularity-violating errors are motivated primarily by proximity of the low frequency segment. In contrast, errors that respect that regularity do not require a nearby source; these errors will be generated in all situations.

Do these effects arise within the phonological sub-component? First, there was no control for perceptual effects (although participants were pre-trained on pronunciation of these nonwords). Such a control would be desirable; with respect to the SLIPs paradigm, there is some question as to whether the effects are localizable to spoken form processing. Camden, Motley, & Baars (1982; see also Camden, 1980) present evidence

that performance in this task is at least partially a result of perceptual confusions (see also Carter & Bradshaw, 1984). Some studies have controlled for these perceptual confusions by asking participants to slowly repeat the correct target; trials on which participants cannot recall the target are excluded (Dell, 1986, 1990). Due to the absence of such a control from this study, perceptual errors could be contributing to the results. However, given the low memory demands of this study (e.g., the nonwords are CV syllables), it would appear likely that participants have correctly encoded the target sequences. With this caveat, we can assume that these effects derive from spoken form processing.

Within spoken form processing, there is some uncertainty regarding the source of the effects. Since Levitt & Healy used mixed word/nonword sequences (e.g., words like “see” spelled si, and nonwords like ta) the results may be influenced by morphological effects. Furthermore, there is no control for articulatory effects, although (as with other SLIPs and tongue twister experiments; Dell & Repka, 1992) many of the errors are likely to be generated with the phonological sub-component.

Speech errors in an aphasic group

Blumstein (1973) examined the conversational productions of a group of English-speaking aphasic individuals. She found that, as a group, they tended to more accurately produce frequent segments, supporting the encoding of gradient within-language regularities. However, interpretation of these results is rather difficult because it is

unclear how the errors are being generated. Presumably, a large number of cognitive processes could be engaged during the conversational task. Since we have no evidence as to which of these particular processes (or particular group of processes) is damaged in these individuals, the error patterns may arise from a source other than phonological processing. For example, suppose that one of the individuals in this study had damage to the morphological sub-component (which I assume is sensitive to lexical frequency). Assume that as a result, this individual randomly substitutes higher frequency words for lower frequency words. Differences in the phonological content of high and low frequency words (Landauer & Streeter, 1973) could lead to apparent regularity effects. For example, if the average frequency of segments within words is correlated to the frequency of the word, the aphasic individual would show higher error rates on low frequency segments. In this way, the mechanism generating errors could appear to be sensitive to segment frequency when in fact it is sensitive only to word frequency. Because the locus of damage is so uncertain, these data do not provide strong support for the encoding of gradient within-language regularities.

Interim summary: Evidence supporting the encoding of gradient within-language regularities

A variety of studies have examined the influence of within-language regularities on phonological processing. I have noted that many of these studies suffer from

methodological problems. With these caveats in mind, the studies appear to suggest that compared to irregular structure, phonological structure that conforms to gradient within-language regularities is:

- More accurately produced (Blumstein, 1973; Munson, 2001)
- Initiated more quickly (Levelt & Wheeldon, 1994; Vitevitch & Luce, 1998, 1999; Vitevitch, et al. 1997).
- More likely to be produced as an error outcome (Levitt & Healy, 1985).

Note that these studies failed to control for the effects of other regularities (e.g., cross-linguistic regularities). Thus, it is difficult to ascertain whether the effects observed here are due to encoding of the specific regularities that were proposed, or to other regularities.

Encoding of cross-linguistic regularities

Many aphasia researchers have examined the question of whether or not cross-linguistic regularities are encoded. This derives (at least partially) from the interest that the linguist Roman Jakobson (one of the fathers of modern-day phonological theory) took in the possible relationship between cross-linguistic regularities identified by linguistic theory and “disordered speech” (as well as language acquisition; Jakobson, 1941/1968). I focus on these studies here.

Aphasic speech errors: Group studies

Analysis of speech errors made by groups of aphasic individuals have found effects of cross-linguistic regularities at many levels of granularity. The studies report that errors occur more often on structures that are cross-linguistically marked (Blumstein, 1973; Carter, Gerken, & Holland, 1998; den Ouden, 2002) and that errors often result in structures that are cross-linguistically unmarked (Béland, 1990; Béland & Favreau, 1991; Béland, Paradis, & Bois, 1993; Blumstein, 1973; Christman, 1994; Code & Ball, 1994; Favreau, Nespoulos & Lecours, 1990; Kohn, Melvold, & Smith 1995; Nespoulos, Jeanette, Béland, Caplan, & Lecours, 1984; Nespoulos, Jeanette, Ska, Caplan, & Lecours, 1987; Nespoulos & Moreau, 1997, 1998). However, because the above studies report the results of group data, they do not provide strong evidence regarding the localization of these effects to the phonological sub-component. The problem is that the individuals within these groups could have had any number of deficits that resulted in spoken production errors (e.g., deficits to perceptual processes, morphological processes, articulatory processes; see the discussion of Blumstein, 1973, above). Two of these studies (Béland, 1990; Béland & Favreau, 1991) even collapsed data from both neurologically intact and impaired individuals. We cannot assume that within such a diverse range of individuals errors result from the same underlying deficit. This uncertainty about the locus of errors makes it difficult to draw inferences about the

phonological sub-component specifically. Thus, these results do not provide strong support for the encoding of cross-linguistic regularities.

Aphasic speech errors: Case studies

The single-case methodology provides a better opportunity to identify the locus of spoken production errors. Here, I review two such cases. B eland & Paradis (1997) describe the case of a French-speaking aphasic. Using assumptions derived from the Theory of Constraints and Repair Rules (Paradis, 1988), B eland & Paradis claim that just as certain phonological structures are cross-linguistically marked or unmarked, particular “repair strategies” are classified as cross-linguistically marked or unmarked. “Repair strategy” refers to the structural manipulation used to convert marked structures into unmarked structures. B eland & Paradis claim that, with respect to marked syllable structure, preservation of segmental material (e.g., through insertion of a vowel) is the unmarked repair strategy and deletion the marked strategy. They examined the performance of an aphasic individual, HC, to see if her errors were influenced by the markedness of phonological structures, as well as markedness of the repair strategy.

HC was a French-speaking primary progressive aphasic with probable dementia of the Alzheimer type. Initially, she showed mild atrophy of the left hemisphere, although this progressed in severity during the course of study. At the outset of the study, HC’s deficit in word production tasks appeared to be confined to the spoken form component.

She could discriminate auditory and orthographic stimuli, and her comprehension was intact. No report of her articulation is provided; it is thus possible that the results reported below do not reflect the effects of damage not to phonological processing, but rather a disruption to the articulatory sub-component.

HC was examined over a period of several years after these initial observations. During this time her condition deteriorated; her error rate in spoken output tasks increased markedly. Picture naming became a very difficult task, so data from the latter stage of her testing come from only repetition and reading. In addition, in these later stages of illness, her auditory discrimination was no longer perfect, and her auditory and written comprehension were severely impaired. The data from the latter period of the study should therefore be viewed with some caution.

Béland & Paradis examined HC's spoken word production across a variety of tasks (reading, repetition, naming, conversation). The data were divided into two phases (I and II), with the division roughly corresponding to the onset of her difficulties in auditory discrimination (although auditory and visual comprehension difficulties began in Phase I). HC's errors on different syllable structures were analyzed. During both phases, HC's errors tended to convert marked syllable structures into unmarked syllable structures. HC respected markedness of the repair strategy during Phase I: she preferred to insert (rather than delete) material. However during Phase II, she preferred to delete (rather than

insert) material to eliminate irregular syllable structures. As noted above, the results from Phase II may be due to HC's severe decline in cognitive processes not associated with the spoken form component.

Romani & Calabrese (1998) examined whether cross-linguistic regularities affected the errors of an Italian-speaking aphasic individual. DB suffered a stroke while scuba diving; this produced an extensive left fronto-parietal lesion. The study was conducted 6 years after the stroke, and his condition remained stable throughout testing. The data used in the study draw entirely on his repetition performance. DB's deficit in this task appeared to be confined to the spoken form-component, and was not due to damage to the articulatory sub-component. He exhibited no auditory discrimination deficits, suggesting that his perceptual representations are intact. Furthermore, he exhibited no difficulties in moving muscles in the face or mouth, suggesting that he did not have a articulatory deficit. With respect to the morphological sub-component, DB made relatively few morphological errors in production. He was equally accurate at producing words and nonwords. However, there is some suggestion of a lexical frequency effect (although the words on the list used to evaluate frequency effects are not controlled for phonological regularity). The pattern of results suggests that his primary deficit is due to the phonological sub-component. To guard against possible morphological effects, Romani & Calabrese removed all morphological errors prior to the analyses reported below.

Romani & Calabrese examine the influence of sonority regularities on DB's errors. Sonority is an abstract sub-segmental property roughly corresponding to the "resonance" or openness of the vocal tract (with more resonant segments being more sonorous than less resonant segments; e.g., nasal stops like /n/ are more sonorous than oral stops like /t/). Sonority is used by linguists to explain generalizations about the ordering of segments within a syllable (Clements, 1990). Cross-linguistically, syllables tend to have less sonorous segments in onset, and more sonorous segments in coda. Moreover, the most cross-linguistically common onset clusters have rising sonority profiles and the most common coda clusters have falling sonority profiles. Romani & Calabrese found that DB's errors tended to improve the sonority profile of syllable onsets. Furthermore, DB's errors tended to eliminate other marked aspects of syllable structure (e.g., onset clusters).

Other case studies (Hatfield & Walton, 1975; den Ouden, 2002) have also reported cross-linguistic regularity effects. However, like the group studies reported above, these studies do not establish the processing locus of the errors produced by the aphasic individuals; as such, they do not provide strong support for localizing the regularity effects specifically within the phonological sub-component.

Interim summary: Evidence supporting the encoding of cross-linguistic regularities

A variety of studies have examined the influence of cross-linguistic regularities on phonological processing. I have described the methodological problems associated with many of these studies. With these caveats in mind, the results suggest that compared to marked structure, cross-linguistically unmarked phonological structure is:

- More accurately produced (Blumstein, 1973; Carter et al., 1998; den Ouden, 2002).
- More likely to be produced as an error outcome (Béland, 1990; Béland & Favreau, 1991; Béland & Paradis, 1997; Béland et al., 1993; Blumstein, 1973; Christman, 1994; Code & Ball, 1994; Favreau et al., 1990; Hatfield & Walton, 1975; Kohn et al., 1995; Nespoulous et al., 1984, 1987; Nespoulous & Moreau, 1997; 1998; Romani & Calabrese, 1998).

Note that these studies failed to control for the effects of other regularities (e.g., gradient within-language regularities). It is possible that regularities defined in some other manner provide an account that is as good as, or even better than, the account afforded by cross-linguistic regularities.

Studies considering multiple types of regularities

Many of the studies reviewed above note that some other regularity may be correlated with the regularities examined in their study (e.g., Levitt & Healy, 1985, note the

correlation between within-language segment frequency and cross-linguistic markedness). However, none of these studies have attempted to analyze the differential effects of these related regularities. Here, I review two studies that have explicitly examined the encoding of different types of regularities.

Kupin (1982) examined the effect of within-language and cross-linguistic regularities on accuracy in tongue twisters. Kupin had participants repeat, as quickly as possible, nonwords made up of two CVC syllables. Each syllable began and ended with the same sound, but the two syllables in each target used different consonants (e.g., target sequences had the form “tatdad” but not “tattat”). Kupin examined correlations between various measures of phonological regularity and error rates on this task. Within-language well-formedness was measured by the product of the frequency of the two consonants in the target (e.g., for “tatdad”, the frequency of /t/ times the frequency of /d/); irregular target sequences had lower frequencies than regular sequences. Cross-linguistic well-formedness was indexed by the sum of the marked sub-segmental features that made up the two consonants; target sequences with more irregular features were less regular than those with fewer irregular features. Both of these factors were significantly correlated with error rate. These results appear to derive from the phonological sub-component. First, Kupin found no significant correlation with several articulatory variables (e.g., number and difficulty of changes of articulatory configurations; error rates of individuals

with articulatory disorders such as dysarthria). He did not control for misperceptions in this task; however, given how short the targets were, it is unclear whether participants would have difficulty correctly encoding the targets. Finally, the use of nonwords eliminates gross morphological effects (although phonological neighborhood effects may still influence the results). With these caveats in mind, these results suggest an influence of both within-language and cross-linguistic phonological regularities on processing within the phonological sub-component.

Motley & Baars (1975) used the SLIPs paradigm with CVC nonword primes and targets. They hypothesized that in this paradigm errors would be more likely to occur when the initial consonant of the second nonword is more regular than the first, relative to pairs in which the initial consonant of the second nonword is less regular. Their logic was that if errors are biased towards regular phonological structures, regular consonants will be more likely to “overpower” irregular consonants (and be produced first) than vice versa. To measure the number of cases in which “overpowering” occurred, they analyzed anticipation errors (e.g., sav tiz → tav tiz) as well as exchange errors (e.g., sav tiz → tav siz). They compared the rate of these errors on pairs where the irregular segment was first to the error rate on pairs where the regularity relationship was reversed. If their logic is correct, and errors are biased towards regular structure, there should be higher error rates on pairs where the irregular structure is first in the pair.

Motley & Baars tested the effect of two different types of regularity. Cross-linguistic well-formedness was indexed by the relative number of marked sub-segmental features; consonants with more irregular features were considered less regular than those with fewer irregular features. Gradient within-language well-formedness was indexed by relative frequency of the two segments. In each condition, word-initial transitional probability of the two phonemes was matched so that any differences would be attributable to the properties of the initial consonants. Motley & Baars found no markedness effect; there was no difference between error rates on pairs where more marked preceded less marked segments compared to pairs where the opposite markedness relation held. In contrast, there was a significant effect of frequency; more errors were found on pairs where the less frequent segment preceded the more frequent segment (as compared to pairs with the opposite frequency relationship). A subsequent experiment suggested effects of supra-segmental frequency. When the initial phoneme of the second nonword had a higher word-initial transitional probability with the vowel of the first word than that of the target, exchange errors were more likely.

It should be noted that, like many SLIPs studies, Motley & Baars did not control for perceptual errors during this task. A second concern is the role of similarity. As discussed in section 2, sub-segmental similarity influences the likelihood of segment interaction in speech errors. It is unclear whether Motley & Baars controlled for this

factor. This is especially important because the consonants pairs used in each condition do not appear to appear to be the same; thus, it is possible that differences in similarity contributed to the difference in results. Morphological effects were partially controlled for in that nonwords were used; however, the role of phonological neighborhoods of the nonwords cannot be discounted. Thus, although these results generally support the encoding of phonological regularities by the phonological sub-component, they do not constitute strong evidence against the encoding of cross-linguistic regularities.

Interim summary: Evidence supporting the encoding of phonological regularities

The evidence from speech errors, aphasia, and implicit learning studies all suggest that categorical within-language regularities are encoded. Errors are more likely to result in phonotactically legal structures than illegal structures. There is no clear evidence regarding the granularity of these regularities.

With respect to gradient within-language and/or cross-linguistic regularities, the results are more mixed. In some cases, it is unclear whether the results derive from the spoken form component at all, much less the phonological sub-component specifically. Nevertheless, several studies report effects that are likely to have arisen in the phonological sub-component (e.g., Béland & Paradis, 1997; Kupin, 1982; Levitt & Healy, 1985; Motley & Baars, 1975; Romani & Calabrese, 1998). These studies suggest that the output probabilities within the phonological sub-component are higher for

structures that are well-formed (with respect to graded within-language and/or cross-linguistic regularities) than for structures that are ill-formed. It is unclear which of these types of regularities best accounts for the data. I have found only one study (Motley & Baars, 1975) that attempts to contrast different definitions of regularity; it finds that gradient within-language regularities provide a better account of the data. Unfortunately, this study may not control for sub-segmental similarity effects.

Null results regarding encoding of phonological regularities

Some studies of spontaneous speech errors have found no evidence (at the segmental level) for an effect of either gradient within-language or cross-linguistic regularities. If the production system encoded a regularity associated with /k/ and /g/ (e.g., /k/ is more frequent or less marked than /g/), the output probability of /g/ would be lower than that of /k/. This asymmetry in output probabilities should produce asymmetries in errors; /g/ → /k/ errors should be more likely than /k/ → /g/. Several studies have reported no such asymmetry in their corpus of errors, reporting equal numbers of both types of errors (Arabic: Abd-El-Jawad & Abu-Salim, 1987; English: Frisch, 1996, 1997; Shattuck-Hufnagel & Klatt, 1979, 1980; Swedish: Söderpalm, 1979). On the basis of these data, researchers have concluded that gradient within-language and/or cross-linguistic regularities are not encoded. This conclusion is not necessarily correct. Because error corpora do not provide information about correct responses, they cannot provide

information about the rate or probability of errors. This makes interpretation of the results difficult; just because similar numbers of each type of error are observed does not entail that the probability of each type of error is the same (see Levitt & Healy, 1985, for a similar argument). If the frequency of segments in the language is different, an equivalent number of errors is meaningless; given potential frequency differences, the probabilities of making each error can be very different²⁸. For example, suppose the error rate on /z/ is 50% and the error rate on /s/ 5%. If /s/ is 10 times more frequent than /z/ in the language, we will not observe any difference in the number of /s/ and /z/ errors (e.g., 100 /s/ segments will yield 5 errors, and 10 /z/ segments will also yield 5 errors). Such a scenario is not unlikely; frequency is inversely correlated with markedness (see section 1). In all likelihood, these data in fact reflect higher error rates on irregular segments.

This criticism is somewhat blunted by null results that have been reported in experimental paradigms (namely, the tongue twister task, with English speakers: Frisch, 1996²⁹; Kupin, 1982³⁰; Wilshire, 1999) where the opportunities for errors are known.

²⁸ Some studies (e.g., Frish, 1996, 1997) do not compare raw numbers of errors, but compare probabilities—using frequency in the corpus of errors to estimate frequency in the language. As argued by Levitt & Healy (1985), this is a suspect assumption. If irregular segments are more likely to be targeted for errors, irregular error targets will have much higher relative frequency in the error corpus than in the language as a whole. Similarly, if irregular segments are substituted, even at low rates, for regular segments, then the irregular segment will have a higher than normal relative frequency as an error outcome in the corpus. The high frequency of regular segments will counteract the low rate of regular→irregular errors.

²⁹ Frisch (1996) found null results for nonword tongue twisters, but anti-regularity effects for words (see below).

³⁰ Although error outcomes were not influenced by regularity, Kupin (1982) did find that error rates were influenced (see above).

However, tongue twisters (and spontaneous speech errors) involve many contextual errors. Levitt & Healy (1985; see above) reported that errors that turn regular structure into irregular structure are more likely to occur when the irregular structure is present. Thus, the magnitude of the regularity effect may be reduced when contextual errors form most or all of the data set. Observation of significant effects may require examining a larger number of tongue twister productions than examined in these studies.

Null results have also been reported in aphasic data. However, it is unclear what the locus of damage is in these individuals. Wilshire & Nespoulos (1997) report no effect of syllable frequency on the productions of an aphasic French speaker. However, his performance is affected by lexical frequency, suggesting the involvement of the morphological sub-component. Similarly the null results of Niemi & Koivuselkä-Sallinen's (1985) study of three Finnish-speaking aphasics may result from impairments to speech comprehension processes.

Reasoning from null results is notoriously difficult. Studies reporting the absence of phonological regularity effects suffer from a variety of flaws that could have eliminated the effects. Spontaneous speech error studies fail to analyze differences in error probabilities; tongue twister studies rely too heavily on contextual errors; and the two studies of aphasic speakers fail to localize impairment to the phonological sub-

component. Thus, none of these studies offer strong evidence against the encoding of phonological regularities.

Anti-regularity effects

I have assumed that phonological regularities are encoded in output probabilities of the phonological sub-component in the following way: structures that are more regular have higher output probabilities than those that are less regular. However, some studies have indicated the opposite pattern holds. The most extensive study is that of Stemberger (1991a). Stemberger examined the performance of English speakers in the SLIPs paradigm with word targets. Target words and exchange error outcomes were matched for lexical frequency as well as length in letters and segments. To look for effects of regularity, he contrasted relative contextual error rates on each initial consonant in a SLIPs pair (e.g. /s/ → /ʃ/ vs. /ʃ/ → /s/ for “sift shuck” or “shift suck”).

Stemberger examined consonant pairs where one member was regular with respect to both gradient within-language and cross-linguistic regularities, while the other was irregular. He found that, along a dimension of contrast, the most regular consonant tended to be replaced by less regular consonants. For example, along the dimension of manner, English contrasts nasal (/m/, /n/) fricative (/f/, /s/) and stop (/p/, /t/). Stemberger reports that stop is the most regular manner (i.e., least marked and most frequent), followed by nasal and then fricative. He found that nasal or fricative → stop substitutions

were less likely than stop→nasal or fricative. The most regular manner (stop) was more likely to be eliminated by less regular manners (nasal, fricative) than vice versa—an anti-regularity effect. Similar results were found for other contrasts (place of articulation; palatalization; voicing). Similar results are also reported by Stemberger & Treiman (1986) and Stemberger (1991a) for consonant clusters (errors tend to produce marked/infrequent clusters rather than singleton consonants).

Two different hypotheses have been put forward to explain the anti-regularity effect in these studies. First, there are potential transcriber biases. Frisch & Wright (2002) analyzed English speakers' /s/-/z/ errors in a tongue twister task (using both words and nonwords). Broad phonological transcription of these errors replicated Stemberger's results. However, transcribers are biased, tending to interpret gradient error outcomes as /z/. This transcription bias towards the irregular outcome counteracts the bias towards the regular outcome in categorical speech errors³¹. These studies suggest that Stemberger's (1991a) results may be due, in part, to transcriber biases. It is unclear whether similar transcriber biases can explain the results for consonant clusters reported by Stemberger & Treiman (1986) and Stemberger (1991a).

³¹ Pouplier & Goldstein (2002) also find a similar transcription bias for English speakers' gradient speech errors on some consonant pairs (/t/-/k/). However, they fail to find a similar bias on other pairs (/s/-/ʃ/). See also Chang, Plauché, & Ohala (2001) and Goldstein (1980) for findings of a regularity bias when identifying consonants in noise.

A second explanation involves potential morphological effects. Frisch (1996) attempted to replicate Stemberger's results using a tongue twister paradigm. Frisch embedded some of Stemberger's consonant pairs in nonwords as well as words. He found no anti-regularity effect with the nonword stimuli, and a strong anti-regularity effect with the word stimuli (see also Wilshire, 1998, as well as Frisch & Wright, 2002, for effects of lexicality on tongue twister results). It is important to note that anti-regularity effects are not always found when words are used as stimuli (e.g., Béland & Paradis, 1997; Levelt & Wheeldon, 1994; Romani & Calabrese, 1998; see above). However, the sensitivity of biases on these particular consonant pairs to the morphological status of the stimuli raises further doubts about whether these findings are due to properties of the phonological sub-component specifically.

Anti-regularity effects have also been reported in some studies of language impairment³². Béland & Favreau (1991), in a study of normal and aphasic French speech errors, found that coronals (the most frequent, least marked segments) were the most likely to be deleted in errors—an anti-regularity effect. However, as noted above, it is unclear whether these results bear on the phonological sub-component. The errors observed in this group could have been generated by many different cognitive processes.

³² Anti-regularity effects have been reported in other cases that appear to involve disruption of the articulatory sub-component (e.g., Dogil & Mayer, 1998). These are not reviewed here.

Summary: Previous research on the encoding of phonological regularities

The phonological sub-component appears to encode phonological regularities by assigning higher output probability to phonologically regular structures. There is clearly a tendency for speech errors to result in phonotactically legal structures, suggesting that categorical, within-language regularities are encoded. Some studies suggest that, in addition, gradient, within-language regularities are encoded; others suggest cross-linguistic regularities are encoded. However, there is no conclusive evidence that favors the encoding of one these types of regularity, and not the other.

It should be noted that many studies have reported “tendencies” to follow regularities. The output probability of regular structures is greater than that of irregular structures, but not absolutely so. Irregular structures are not impossible, just less likely. This suggests that gradient regularities are encoded; well-formedness is defined on a continuous, rather than categorical, scale.

SECTION 4: EXPERIMENTAL INVESTIGATIONS

Section 1 introduced three proposals regarding the types of regularities that are encoded by the phonological sub-component. These are summarized in Table 1 below.

Table 1. Theoretical proposals regarding what types of regularities are encoded.

Theory	Claims about regularities encoded by phonological sub-component	
	Feature	Types of regularities encoded
Instance-Based	<u>Scope</u>	Within-language, <u>token frequency</u> .
Theory	<u>Scale</u>	Categorical and <u>gradient</u> .
	<u>Granularity</u>	Segmental and supra-segmental.
Lexical	<u>Scope</u>	Within-language, <u>type frequency</u> .
Distribution	<u>Scale</u>	Categorical and <u>gradient</u> .
	<u>Granularity</u>	Segmental and supra-segmental.
Markedness	<u>Scope</u>	<u>Cross-linguistic</u> and within-language markedness.
Theory	<u>Scale</u>	Categorical.
	<u>Granularity</u>	<u>Sub-segmental</u> , segmental and supra-segmental.

The studies reviewed in section 3 provide some constraints on theories of regularities. Specifically, these studies suggest that gradient regularities are encoded. Ill-formed structures are not always assigned output probabilities of zero; in some cases, they are merely assigned lower output probabilities than well-formed structures. This suggests that gradient regularities can be encoded, supporting the Instance-Based and Lexical Distribution theories. However, with respect to scope and granularity, the results reviewed in Section 3 do not distinguish between the claims of these three theories. First, with respect to scope, these studies most clearly show that categorical within-language regularities are encoded; this is predicted by all three of these theories. Other studies suggest that graded within-language and/or cross-linguistic regularities are encoded, but fail to distinguish between the two. Second, with respect to granularity, the studies in section 3 reported effects at all levels of granularity; however, they failed to control for the effects of other levels of granularity (which may make similar well-formedness distinctions).

This section presents results from two experimental studies that attempt to distinguish between these theories with respect to scope and granularity. The first study examines the contrasting predictions of these theories in an implicit learning paradigm. The second examines biases in speech errors in a tongue twister task.

Experiment 1: Implicit learning of phonological regularities

As described in section 3, Dell et al. (2000) introduced an experimental paradigm in which adult participants implicitly learn new phonological regularities. In Dell et al.'s experiments, participants read aloud nonword sequences that respected phonological regularities not found in the participants' native language. For example, in one condition with English-speaking participants, Dell et al. restricted /f/ to onset position in the experimental materials. Participants were asked to read aloud the sequences quickly, which induced speech errors. Dell et al. found that the participants' speech errors reflected the regularities of the experimental stimuli (e.g., very few /f/ errors occurred in coda position).

Experiment 1 used the Dell et al. paradigm to examine the granularity of regularities encoded by the phonological sub-component. As in Dell et al., participants were exposed to 2 segmental regularities with a supra-segmental contextual restriction (e.g., /f/ is restricted to onset³³, /s/ is restricted to coda). I refer to the consonants associated with these regularities as restricted consonants. Unlike Dell et al.'s experiments, one of the restricted consonants (the restricted test consonant) shared sub-segmental structure with an unrestricted consonant (the unrestricted related consonant). The other restricted consonant (the restricted control consonant) was not highly similar to any of the

³³ Since all nonwords in the experiment are monosyllables, the context restriction is ambiguous between syllable and word (e.g., all onsets are also word initial consonants).

unrestricted consonants. For example, in one condition /f/ was the restricted test consonant; it was restricted to onset while /v/ (the unrestricted related consonant), sharing all of /f/'s sub-segmental features save voicing, occurred both in onset and in coda. In the same condition, /s/ was the restricted control consonant; it was restricted to coda while its related consonant /z/ was absent from the stimulus set.

The theories proposed above make different predictions regarding performance on the restricted test and control consonants. If sub-segmental regularities are not encoded (as claimed by the Instance-Based and Lexical Distribution theories), then participants should learn the regularities associated with the restricted test and control consonants with comparable ease. However, if sub-segmental regularities are encoded (as claimed by the Markedness theory), then there is conflicting information regarding the restricted test consonant. At the segmental level, there is a categorical regularity (e.g., /f/ is always in onset, never in coda), but at the sub-segmental level, there is no categorical regularity associated with this restricted test consonant (e.g., labiodental fricatives are found in both onset and coda due to the presence of unrestricted /v/). Because of the conflict between the segmental and sub-segmental levels, participants should have more difficulty encoding the regularity associated with the restricted test consonant than the regularity associated with the restricted control. Experiment 1 examines whether theories that claim

sub-segmental regularities are encoded better predict learning for a number of consonant pairs that differ in voicing.

Method

Participants

Forty undergraduate and graduate students from the Johns Hopkins University community participated in the experiment (16 males, 15 right-handed; 24 females, 19 right-handed). They were compensated with \$7 or received extra-credit in introductory courses for their participation. All participants reported that they were native speakers³⁴ of English (learned English prior to the age of five). All participants reported no history of speech/language impairment.

Materials

The restricted test and unrestricted related consonants were drawn from four pairs of voiced-voiceless consonants: /v/-/f/, /z/-/s/, /d/-/t/, and /g/-/k/. Each pair was used in two experimental conditions. In one condition, the voiced consonant was the restricted test, while the voiceless was the unrestricted related consonant; in the second, the situation was reversed. The use of four pairs created a total of 8 conditions; 5 participants were randomly assigned to each condition.

³⁴ It should be noted that several of the participants were native bilinguals and/or had extensive training in foreign languages. Languages spoken by the bilinguals in this group included: Estonian; Japanese; Korean (2 participants); Mandarin Chinese; Spanish ;Vietnamese.

In addition to the restricted test and unrestricted related consonants, there were three other kinds of consonants in each stimulus list. First, each list contained /h/ and /ŋ/, restricted to onset and coda respectively (following the constraints of English). Second, each list contained a restricted control consonant that was not similar to any unrestricted segments. Finally, each list contained three unrestricted consonants (including one voiced-voiceless pair) that were not similar to either of the restricted consonants. The particular unrestricted and restricted control consonants used for each pair of consonants is shown in table 2 below.

Table 2. Restricted control and unrestricted consonants, Experiment 1.

Consonant Pair	Restricted Control	Unrestricted
/v/-/f/	/s/	/g,k,m/
/z/-/s/	/f/	/g,k,m/
/d/-/t/	/f/	/g,k,m/
/g/-/k/	/s/	/f,v,m/

For the consonant pair /z/-/s/, the restricted test segment was in coda and the restricted control was in onset. In all other conditions the restricted test segment occurred in onset and the restricted control was in coda.

These consonants were presented to participants in sequences of four CVC syllables (the vowel was always /ε/). Each consonant appeared once per sequence, with equal frequency in each syllable (i.e., equally often in the first, second, third, and fourth syllables). For each condition, a total of 3456 sequences satisfied these constraints. For each participant, a list of 192 sequences was drawn from this set in a random order.

Each sequence of four nonwords was spelled (for visual presentation to participants) in the following way. The vowel /ε/ was always spelled “e”. /g/ was spelled “gh” in initial position (e.g., “ghem”); all other initial consonants were spelled with single letters. Consonants were doubled in final position except when this would violate English orthographic constraints (specifically, /d, v, k, m/ were spelled with single consonants).

Procedure

The experimental session took place in a sound attenuated chamber. Stimuli were presented on a laptop computer placed on a small table in front of the participant. Each participant’s spoken productions were recorded onto audio tape by a head-mounted microphone. Recordings were used to should improve error detection by transcribers;

several studies (Ferber, 1991, 1995; Tent & Clark, 1980) have noted the difficulty of on-line detection of errors without the assistance of recordings³⁵.

Each trial proceeded as follows:

1. Participants were shown a single sequence of four syllables (e.g., “heng fek meg ness”) centered on a computer monitor in black 18 point Charcoal type (white background). They were instructed to read the sequence aloud in time to metronome-like clicks from the computer.
2. After the participant pressed a key, a set of four clicks was played at a rate of 1/second. The participant read aloud the sequence in time to these slow-playing clicks. This was done to ensure that the participants correctly encoded the target sequence before repeating it quickly.
3. After the participant pressed a key again, a set of twelve clicks was played at a rate of 2.5/second. This allowed for three fast repetitions of the sequence. These repetitions were intended to elicit speech errors.

³⁵ Although recording eliminates some transcriber errors, it does not necessarily allow all errors in the task to be correctly categorized (Buckingham & Yule, 1987; Cutler, 1982). One concrete concern is that production errors can result in articulatory/acoustic configurations that are not found in normal (non-errorful) speech (Boucher, 1994; Frisch & Wright, 2002; Laver, 1980; Mowrey & MacKay, 1990; Pouplier, Chen, Goldstein, & Byrd, 2000). Since transcription analysis records responses using normal speech category labels (i.e., IPA characters), such errors may be incorrectly identified by transcribers as errors within a normal speech category. However, the available evidence suggest that these transcription errors show an anti-regularity effect; that is, they are biased towards less regular structures (Frisch & Wright, 2002; Pouplier & Goldstein, 2002; see section 3 for discussion). Therefore, such transcription errors should weaken the regularity effects reported here.

The sequence remained visible through the entire trial, minimizing the memory demands of the task.

A set of three practice trials preceded the experimental trials. Practice trials were identical to experimental trials except that the consonants within these sequences were not used in the experimental trials. Following the practice trials, participants were pre-trained on the pronunciation of syllables that occurred in experimental sequences. Each syllable was presented individually (centered on the screen). The participant was then asked to read the syllable aloud, and corrected if their pronunciation did not match the desired one. At no time were participants instructed regarding the distribution of segments within the experiment, nor the similarity between different segments in the experiment. The experimental session itself consisted of four blocks of 48 sequences apiece. The entire procedure took approximately 45 minutes to complete.

Results and Discussion

Recordings were examined, and consonantal substitution errors³⁶ were transcribed using broad IPA transcription (vowel errors were rare, and not recorded). Each error is referred to by the segment that was inserted. For example, suppose “heng fek meg nes” was produced as “heng sek meg nes” (the error segment is underlined). This is referred to as an /s/ error; /s/ replaces the segment /f/, and the error is produced in onset.

³⁶ Cutoff errors, where both the error segment and the target were produced (e.g., /kɛɪ/ → /m...kɛɪ/) were included.

Analyses contrast the rate at which inserted segments violate target syllable position. For example, in the target sequence above, /s/ occurs in the coda of the fourth syllable. In the error “sek,” the /s/ error violates /s/’s syllable position in the target sequence (the error is produced in onset, not coda). In contrast, if the error was “fes,” the /s/ error would not violate target syllable position.

Note that if the inserted segment is a restricted segment, violating target syllable position is equivalent to violating a regularity of the stimulus set. For example, if /s/ is confined to coda in all target sequences, the error “sek” would violate the regularity, while “fes” would respect the regularity.

Voicing errors were excluded from the analysis. In a voicing error, the target consonant differs from the inserted consonant only with respect to voicing (e.g., /k/→/g/, /g/→/k/). These were excluded because sub-segmental similarity affects segmental speech errors (i.e., similar segments are more likely to be involved in speech errors; see section 2). Given the composition of the stimulus list, the restricted test segment has an opportunity to interact with a highly related segment (the unrestricted related segment), whereas the restricted control does not have the opportunity to interact with an equivalent segment. To equate the two restricted consonants in terms of error opportunities, errors in which the restricted test segment was replaced by the unrestricted related consonant (and vice versa) were excluded. Furthermore, since two unrestricted consonants also

differ only in voicing, all voicing errors were excluded to make the different segment groups as similar as possible. (Including these errors only strengthens the results reported below³⁷.)

Each participant was randomly assigned to one of two transcribers. Inter-transcriber reliability was good. Both transcribers analyzed two 10 minute recordings (randomly selected from two participants). Out of 2880 consonants, both transcribers agreed on 2878 consonants (99.9% agreement). For errors alone, the agreement rate is 77.8% (out of 9 errors identified by one or both transcribers, they agreed on 7).

Results: Collapsed across all conditions

Overall consonantal error rate across conditions was 4.8% (8912/184320). After voicing errors were excluded, a total of 6762 errors were analyzed. The first result is that no errors violate the regularities of English. All 867 /h/ errors occurred in onset (e.g., /fɛŋ/ → /hɛŋ/: target /f/ in onset is replaced by an /h/); all 771 /ŋ/ errors occurred in coda.

Second, replicating Dell et al. (2000), restricted control segment errors almost always respect the regularity. Collapsing across conditions, only 2.6% (18/691) of the errors violate the regularity associated with the restricted control segment. For example, when /s/ is confined to coda position (in the absence of /z/), errors like “fem ves → sem ves” are

³⁷ Inter-transcriber agreement on voicing errors was poor. When these are included, the overall agreement rate drops to 99.2%, while agreement rate on errors drops to 38.7%.

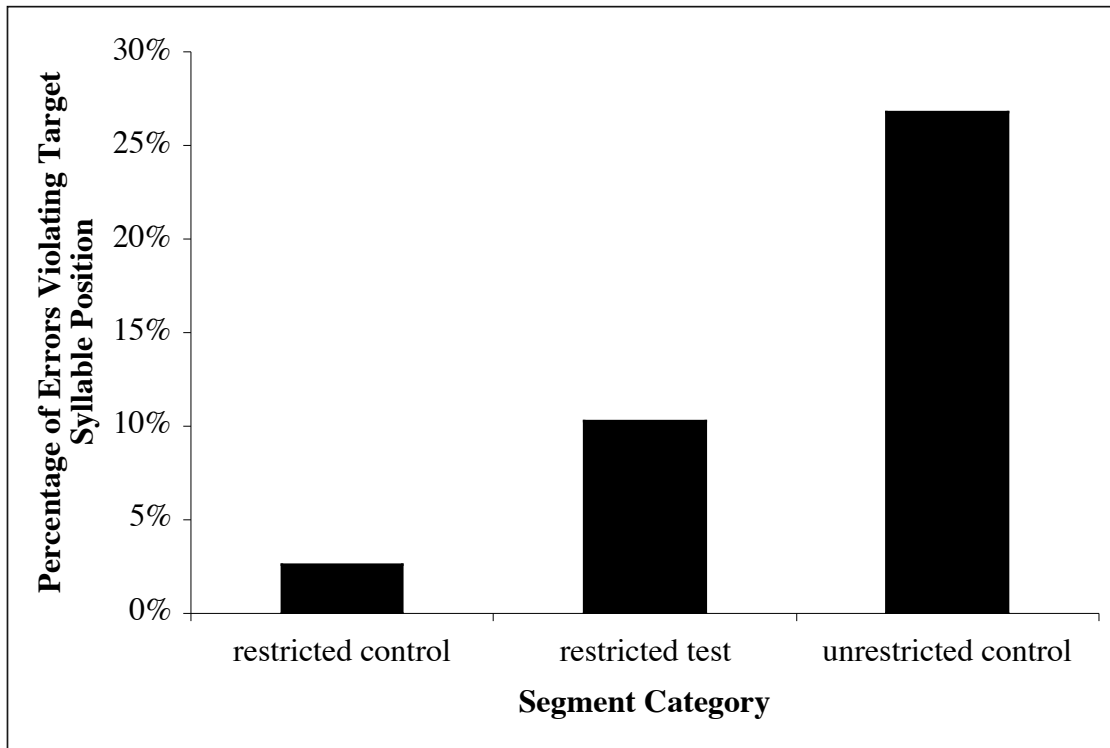
extremely unlikely. Participants appear to have extracted the regularities associated with the restricted control segments.

Third, also replicating Dell et al. (2000), unrestricted consonant errors violate target syllable position at rates significantly less than chance levels, suggesting that consonants tend to preserve their syllable position in errors. 26.8% (775/2896) of unrestricted consonant errors violate target syllable position. This is significantly less than the percentage predicted by chance (56.1%; see Appendix A for a discussion of chance levels; binomial test, $Z = -31.84$ (continuity corrected), two-tailed $p < .0001$). However, the rate at which unrestricted consonants violate target syllable position is significantly greater than that associated with the restricted control segment ($\chi^2(1, N = 3587) = 187.6$ (continuity corrected), one-tailed $p < .0001$; by participants, paired t-test comparing proportion of illegal errors across restricted control and unrelated segments; $t(39) = -8.4$, two-tailed $p < .0001$)³⁸. This shows that the tendency to respect the restricted control segment regularity does not reduce to a tendency to respect target syllable position.

The fourth, novel, result (shown in Figure 2) is that restricted test segment errors are distributed differently from both the restricted control and unrestricted segment errors.

³⁸ If anything, the by participants comparison is conservative. As discussed in Appendix A, restricted control segment errors are slightly more likely not to share target syllable position (57.1%) than unrestricted errors (56.1%).

Figure 2. Performance on all segment pairs, Experiment 1.



Note: All pairwise differences are significant.

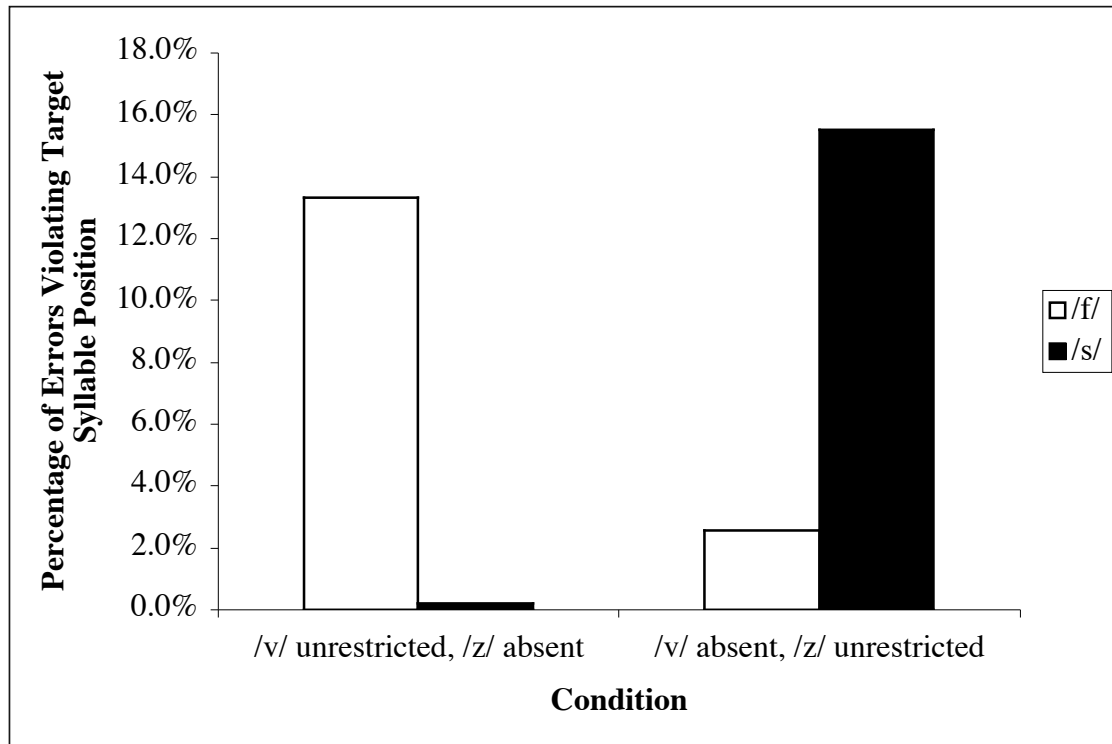
10.3% (66/642) of the restricted test segment errors violate the corresponding regularity. This is significantly greater than the percentage of comparable errors on the restricted control ($\chi^2(1, N = 1333) = 31.9, p < .0001$; by participants, $t(39) = -2.6, p < .02$). Although the percentage of errors violating the restricted test segment regularity is higher than that associated with the restricted control, participants have extracted the regularity. The percentage of restricted test segment errors that violate target syllable position is significantly less than that associated with unrestricted errors ($\chi^2(1, N = 3538)$

= 77.8, $p < .0001$; by participants, $t(39) = -4.37$, $p < .0001$)³⁹. This shows that although the tendency to respect the restricted test segment regularity is weaker than the tendency to respect the restricted control regularity, it is still stronger than the general tendency to respect target syllable position.

Is the difference between the restricted control and restricted test segment errors due to the presence of the unrestricted related consonant, or to some other factor? One concern is that for any given condition, different segments are used as restricted test and restricted control segments (e.g., in one condition, the restricted control is /s/ in coda position, while the restricted test is /f/ in onset). Perhaps intrinsic differences between these segment groups produce the results found here. We can test this possibility by comparing the same consonant, restricted to the same syllable position, when it is a restricted control versus a restricted test segment. This can be done for two consonants: /s/ and /f/, shown in figure 3 below.

³⁹ This comparison is conservative. As discussed in Appendix A, restricted test segment errors are slightly more likely not to share target syllable position by chance (58.3%) than unrestricted segment errors (56.1%).

Figure 3. Comparison of /f/ and /s/ across conditions, Experiment 1.

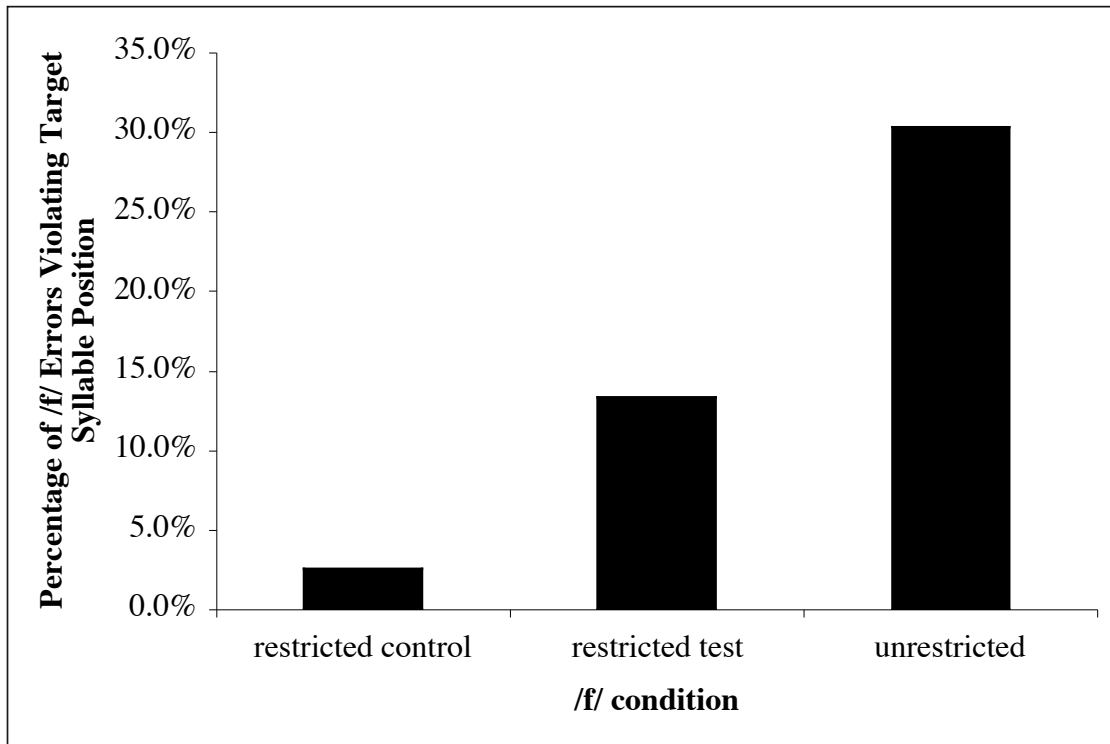


Note: All pairwise differences are significant.

Both consonants show significant differences across conditions (/f/ restricted test vs. /f/ restricted control: $\chi^2(1, N = 230) = 8.4, p < .004$; /s/ restricted test vs. /s/ restricted control: $\chi^2(1, N = 131) = 8.2, p < .005^{40}$). A similar comparison is shown in Figure 4 for /f/ as a restricted control, restricted test, and an unrestricted consonant.

⁴⁰ In both of these comparisons, expected cell counts in the χ^2 test are less than 5; at this point, the χ^2 approximation to the goodness-of-fit test becomes poor. Results were therefore re-calculated using Fisher's exact test; for both comparisons, $p < .0001$.

Figure 4. Comparison of /f/ across conditions, Experiment 1.



Note: All pairwise differences are significant.

The significant difference between /f/ restricted control and restricted test was reported above. /f/ restricted control errors violate target syllable position significantly less frequently than unrestricted /f/ (2.6% (4/155) vs. 30.3% (47/155); χ^2 (1, N = 310) = 41.4, $p < .0001$). In addition, a significant difference is found between /f/ restricted test and unrestricted /f/ errors (13.3% (10/75) vs. 30.3%; χ^2 (1, N = 230) = 6.9, $p < .009$).

The results shown in Figures 3 and 4, comparing the same segment across different conditions, show that the differences between the restricted control, restricted test, and

unrestricted segments are not due to differences in the particular segments making up each category.

Another concern is that differences between segment types is due to the influence of the morphological sub-component. As discussed in section 2, the interaction of the morphological and phonological sub-components creates a lexical bias effect.

Differences in the distribution of lexical items among consonant categories could influence the results. To control for this possibility, a new analysis was carried out in which all word responses were excluded. The results were unchanged⁴¹.

Another potential issue is that within each trial, participants repeated each sequence of four syllables three times. Dependencies among these repetitions may introduce artifacts into the analyses reported here (e.g., if an error is made on repetition one, the error may become more likely on subsequent repetitions)⁴². To control for such effect, a new analysis was performed, excluding the second and third repetitions of each sequence. The results were unchanged⁴³.

⁴¹ All overall differences remained significant (restricted control 3.3% (18/548) vs. unrestricted 27.9% (666/2390): $\chi^2(1, N = 2938) = 149.4, p < .0001$; restricted control 3.3% vs. restricted test 9.6% (49/510): $\chi^2(1, N = 1058) = 16.8, p < .0001$; restricted test 9.6% vs. unrestricted 27.9%: $\chi^2(1, N = 2900) = 74.4, p < .0001$) and were significant or marginal across participants (restricted control vs. unrestricted: $t(39) = -8.8, p < .0001$; restricted control vs. restricted test: $t(39) = -1.9, p < .07$; restricted test vs. unrestricted: $t(39) = -5.2, p < .0001$).

⁴² Thanks to Sanjeev Khudanpur for pointing out this potential confound.

⁴³ All overall differences remained significant (restricted control 2.4% (3/127) vs. unrestricted 31.2% (151/484): $\chi^2(1, N = 611) = 42.9, p < .001$; restricted control 2.4% vs. restricted test 10.4% (12/115): $\chi^2(1, N = 242) = 5.4, p < .03$; restricted test 10.4% vs. unrestricted 31.2%: $\chi^2(1, N = 599) = 19.2, p < .001$). Analysis by participants was not performed due to the small number of errors.

A final issue is that some participants in the study were bilinguals. Perhaps English-speaking monolinguals do not show the interference effect. To control for this possibility, a new analysis was performed in which data from bilingual participants were excluded; the results were unchanged⁴⁴.

Summary: Results collapsed across all conditions

Replicating Dell et al. (2000), participants are able to extract regularities present in the stimulus set. In addition, the results show that participants have difficulty extracting regularities in the presence of a highly similar unrestricted segment. This interference effect does not, however, eliminate learning; the restricted test segment preserved its target syllable position more frequently than the unrestricted segments. These results are significant across participants, and could not be attributed to differences between the particular segments in each group, nor to the influence of the morphological sub-component.

Results: Consonant pairs

Separate analyses were conducted for each consonant pair. These are summarized in Table 3 below. The results for each pair are similar to the collapsed results. For all pairs, restricted control and restricted test segment errors are less likely to appear in non-target

⁴⁴ All overall differences remained significant (restricted control 2.2% (7/315) vs. unrestricted 27.5% (383/1294): $\chi^2(1, N = 1709) = 91.6, p < .001$; restricted control 2.2% vs. restricted test 12.3% (36/292): $\chi^2(1, N = 607) = 22.0, p < .001$; restricted test 12.3% vs. unrestricted 27.5%: $\chi^2(1, N = 1686) = 28.9, p < .001$), as did analyses by participants (restricted control vs. unrestricted: $t(32) = -7.3, p < .0001$; restricted control vs. restricted test: $t(32) = -2.1, p < .05$; restricted test vs. unrestricted: $t(32) = -5.4, p < .0001$).

syllable positions than unrestricted consonants (i.e., within each consonant pair, participants could extract the regularity⁴⁵). For all pairs, restricted control segments are numerically less likely to violate their associated regularity than restricted test segments. This result is significant for all pairs⁴⁶ save /g/-/k/ ($\chi^2(1, N = 234) = .4, p > .50$). These null results are discussed in more detail below.

Table 3. Percentage of errors violating target syllable position for each consonant pair, Experiment 1.

Consonant Pair	Restricted Control	Restricted Test	Unrestricted
/v/-/f/	1.5% (2/132)	12.9% (18/140)	30.0% (202/678)
/z/-/s/	2.6% (6/235)	12.3% (16/130)	26.4% (233/882)
/d/-/t/	2.8% (7/247)	9.8% (21/215)	26.3% (240/913)
/g/-/k/	3.9% (3/77)	7.0% (11/157)	23.6% (100/423)

Note: Ratio of these errors to total errors shown in parentheses.

⁴⁵ /v/-/f/: restricted control $\chi^2(1, N = 810) = 45.4, p < .0001$; restricted test $\chi^2(1, N = 818) = 16.1, p < .0001$; /z/-/s/: restricted control $\chi^2(1, N = 1117) = 61.4, p < .0001$; restricted test $\chi^2(1, N = 1012) = 11.4, p < .0001$; /d/-/t/: restricted control $\chi^2(1, N = 1160) = 62.4, p < .0001$; restricted test $\chi^2(1, N = 1128) = 25.8, p < .0001$; /g/-/k/: restricted control $\chi^2(1, N = 500) = 14.3, p < .0002$; restricted test $\chi^2(1, N = 580) = 19.4, p < .0001$.

⁴⁶ /v/-/f/: $\chi^2(1, N = 272) = 11.2, p < .0001$; /z/-/s/: $\chi^2(1, N = 365) = 12.4, p < .0005$; /d/-/t/: $\chi^2(1, N = 462) = 8.5, p < .004$.

Results for individual consonants

Numerical differences between the restricted test and restricted control are also found in the analysis of each individual consonant except /k/. These are shown in Table 4 below.

Table 4. Percentage of errors violating target syllable position for each condition, Experiment 1.

Restricted Test Consonant	Restricted Control	Restricted Test
/v/	2.8% (2/72)	12.3% (8/65)
/f/	0% (0/60)	13.3% (10/75)
/z/	2.5% (2/78)	8.4% (5/59)
/s/	2.6% (4/155)	15.5% (11/71)
/d/	2.8% (5/179)	7.3% (11/151)
/t/	2.9% (2/68)	15.6% (10/64)
/g/	4.3% (1/23)	14.6% (7/48)
/k/	3.7% (2/54)	3.7% (4/109)

Note: Ratio of these errors to total errors shown in parentheses.

Experiment 1A: /g/-/k/ conditions, excluding /ŋ/

One possible reason for the null result on /g/-/k/ is that one of the restricted consonants (/ŋ/, restricted to coda position in English and in the experiment) is highly similar to both of these consonants. In fact, it only differs from /g/ with respect to nasality. Perhaps interactions with this consonant are eliminating the interference effect. To guard against this possibility, two new conditions using /g/-/k/ were designed, eliminating /ŋ/.

Participants

15 participants were drawn from the same pool as the conditions above, and compensated with \$7 or received extra-credit in introductory courses for their participation. 1 participant was excluded due to equipment failure. The remaining 14 participants consisted of 10 males (8 right handed) and 4 females (4 right handed). All participants reported that they were native speakers⁴⁷ of English and had no history of speech/language impairment. Each participant was randomly assigned to one of 2 conditions (yielding 7 participants per condition).

Materials

Materials were similar to the /g/-/k/ condition above. Instead of including /h/ and /ŋ/, restricted to onset and coda, /d/ and /n/ were added as unrestricted consonants. A total of

⁴⁷ It should be noted that several participants were native bilinguals and/or had extensive training in foreign languages. Languages spoken by the bilinguals in this group included: Farsi, Japanese, and Spanish.

11,520 sequences respected the constraints outlined in the text. For each participant, a list of 192 sequences was drawn from this set in a random order.

Procedure

Procedure was identical to the previous conditions.

Results and Discussion

Scoring and transcription methods are identical to the previous conditions. Restricted control segment errors violate the associated regularity at a rate of 1.5% (2/133); restricted test segment errors violate their regularity at a rate of 3.8% (9/234). This difference, although in the expected direction, is not significant ($\chi^2(1, N = 367) = .9, p > .30$). The presence of /ŋ/ in the stimulus set does not cause the null result observed in Experiment 1.

There are several other reasons that the null result could be found for this pair. First, /ŋ/ is restricted throughout English; this could influence the results. Second, note that in table 4 /g/ is numerically more likely to violate its associated regularity than the restricted control; no such difference is found for /k/. This was also found in Experiment 1A. /g/ violated its associated regularity at a rate of 4.3% (7/164), while the restricted control never violated its regularity (0/63). In contrast, /k/ violated its regularity at the same rate as the restricted control (/k/: 2.9%, 2/70; restricted control: 2.9%, 2/70). This suggests that the null result may be primarily due to the influence of /k/. Future experiments

should examine what aspects of English dorsal consonants (and /k/ in particular) contributes to the null effects observed here.

General Discussion: Experiment 1

Implications for granularity

The performance on the restricted test segment shows that the phonological sub-component encodes sub-segmental regularities. This favors the Markedness theory, and disfavors the Instance-Based and Lexical Distribution theories. Is there no way for these theories to account for these data? Note that the Instance-Based and Lexical Distribution theories are both sensitive to the context of segmental structure. For example, in the Instance-Based theory, the regularity of structures is partially determined by the transitional probability of segments. In the Lexical Distribution theory, the word position and prosodic environment of syllables affects regularity. One possibility is that this contextual information could be used to infer the similarity between segments that share sub-segmental structure⁴⁸. If sub-segmentally similar segments are distributed in a similar way (e.g., similar transitional probabilities; similar patterns of distribution in certain word positions/prosodic environments), it would be unnecessary to appeal to sub-segmental representations.

⁴⁸c.f. American structuralist notion of “pattern congruity” as a means for identifying similarity between phonemes (Anderson, 1985).

To test this possibility, the distribution of restricted test segments was compared to that of other segments within each condition. The distribution of a segment was indexed by its CELEX type frequency (Baayen et al., 1995) adjacent to stressed vowels⁴⁹. These frequency counts provide a measure of how often a segment co-occurs with different vowels. If similar segments tend to co-occur with the same vowels, it would be unnecessary to appeal to sub-segmental representations; segmental co-occurrence patterns would suffice. As shown in table 5, this does not appear to be the case.

⁴⁹ For /z/-/s/, frequency was calculated in word-final position following stressed vowels (these segments are restricted to coda). For all other segments, frequency was calculated in word-initial position preceding stressed vowels. This measure of distribution is intended to capture elements of both the Instance-Based and Lexical Distribution theories by combining position in the word (both theories), identity of the adjacent vowel (Instance-Based theory), and prosodic information (Lexical Distribution theory).

Table 5. Distributional similarity of consonants, Experiment 1.

Restricted Test Consonant	Segment in condition with most similar distribution
/v/	/m/ (.63)
/f/	/m/ (.46)
/z/	/m/ (.31)
/s/	/m/ (.15)
/d/	/m/ (.46)
/t/	/h/ (.72)
/g/	/k/ (.77)
/k/	/g/ (.77)

Note: Correlation (r^2) of frequency counts adjacent to stressed vowels shown in parentheses.

With the exception of /g/-/k/, restricted test consonants do not tend to co-occur with the same vowels as their sub-segmentally similar counterparts. Distributional similarity does not reveal sub-segmental similarity.

In one final attempt to salvage theories that exclude sub-segment representations, note that in several conditions above, the restricted test segment distributions are most similar to one of the unrestricted segments (/m/). Perhaps the difference between restricted test

and control segments is caused not by the unrestricted segment which is sub-segmentally similar, but the unrestricted segment which is distributionally similar (/m/). This account of the data fails as well. Consider the condition where /f/ is a restricted test segment versus when it is a restricted control segment. /m/ is an unrestricted consonant in both of these conditions. If it produced the interference effect, we would expect /f/ errors to show the same pattern in both conditions. Instead, /f/ errors vary depending on the presence or absence of the sub-segmentally similar segment /v/. Clearly, distributional similarity is insufficient; the phonological sub-component must be sensitive to sub-segmental regularities.

Implications for scale

Participants successfully encoded a categorical regularity associated with the restricted control segment. Errors that violated this regularity occurred at extremely low rates, suggesting that the encoded regularity makes a near-categorical distinction between regular and irregular structures. In contrast, for the restricted test segment, participants encoded a gradient regularity. Although the regularity was clearly encoded (i.e., restricted test segment errors violated target syllable position less than unrestricted segment errors), errors that violated the regularity occurred at significant levels (i.e., approximately 10%). This is consistent with the results reported in section 3. The phonological sub-component encodes gradient well-formedness; ill-formed structures can

have output probabilities greater than zero. This is consistent with the claims of the Instance-Based and Lexical Distribution theories, and inconsistent with the claims of the Markedness theory.

Is there no way for the Markedness theory to account for the results? A potential counter-argument is that the gradient performance results from the limitations of the mechanisms encoding regularities. Under this account, a categorical regularity is encoded, but it is implemented in an error-prone system; these errors prevent the system from absolutely obeying a regularity⁵⁰. The problem with this account is performance on the restricted control consonant. The spoken production system is clearly capable of obeying a regularity at near-absolute levels; only 3% of errors violate the regularity associated with the restricted control. The gradient performance on the restricted test consonant is therefore not a function of intrinsic error in the system encoding regularities; the system is encoding a gradient well-formedness distinction.

The restricted test segment regularity is encoded in a gradient fashion despite the fact that the segment is associated with a categorical regularity in the stimulus set (i.e., it always occurs in one syllable position). This is due to the presence of sub-segmental features associated with the restricted test consonant in both syllable positions. Since the

⁵⁰ This proposal is analogous to the Chomsky's "competence/performance" distinction. Chomsky & Halle (1968:3) define this contrast as follows: performance is "what the speaker-hearer actually does...based not only on his [competence], but on many other factors as well;" in contrast, competence is the knowledge that supports "the potential performance of an idealized speaker-hearer who is unaffected by grammatically irrelevant factors." Here, the "grammatically irrelevant factor" is computational accuracy; the system encoding regularities is unable to absolutely obey a regularity .

phonological sub-component encodes sub-segmental as well as segmental regularities, it does not precisely reflect the segmental statistics of the restricted test consonant. The combination of gradient sub-segmental and categorical segmental statistics yields the gradient performance on the restricted test consonant. There is some evidence that over the course of the experiment, the restricted test segment regularity is being encoded in a more categorical fashion. Errors from blocks 1 and 2 are more likely to violate the restricted test regularity than errors from block 3 and 4 (13.3% errors in the first two blocks versus 7.6% errors in the second two blocks; $\chi^2(1, N = 642) = 5.1, p < .03$). No comparable difference was found between restricted control and unrestricted segment errors⁵¹. This suggests that over time the phonological sub-component is more closely reflecting the segmental statistics, making the regularity associated with the restricted test segment more categorical.

Although the distribution of sub-segmental features influences performance on the restricted test consonant, it does not influence performance on the unrestricted related consonant. Note that the sub-segmental features associated with both the restricted test and unrestricted related segments occur in both onset and coda, biased towards the syllable position of the restricted test consonant. If this sub-segmental regularity directly influenced performance, the unrestricted related segment errors should be biased towards

⁵¹ Restricted control: 2.8% versus 2.4% errors violate the associated regularity; $\chi^2(1, N = 691) = 1.8, p > .8$; unrestricted: 26.5% vs. 27% errors do not share target syllable position; $\chi^2(1, N = 2896) = 0.1, p > .75$.

the syllable position of the restricted test segment (as the sub-segmental features shared by the two consonants are more likely to occur in that position). This is not the case. When the unrestricted similar segment is in the same target syllable position as the restricted test segment, 32% (164/513) of the errors violate target syllable position; when the unrestricted similar segment is the other syllable position, 33% (125/382) of the errors violate target syllable position ($\chi^2(1, N = 895) = 0.3, p > .8$). Thus, the sub-segmental regularity influences performance only in conjunction with a categorical segmental regularity (interfering with the encoding of the restricted test segment regularity). By itself, the sub-segmental regularity does not appear to influence error patterns. It is unclear why this is the case; I return to this question in the general discussion.

Implications for scope

This experiment examines the learning of new regularities; as such, it necessarily bears on the encoding of within-language regularities. All three theories assume that such regularities are encoded; the results therefore fail to distinguish these theories with respect to scope.

Locus of effects in this task

It is important to assess whether the effects observed in this experiment arise within the phonological sub-component. As with Dell et al.'s (2000) study, it is unlikely that effects arise outside of the spoken form component, or from the morphological sub-

component (see the discussion in section 3 above). The articulatory sub-component is also not likely to be the source of these effects. I assume that articulatory errors are most likely to occur between segments that are composed of very similar articulations. Many of these errors are excluded from the analysis, as voicing errors are not analyzed. Furthermore, the remaining interactions often involve very distinct consonants. For example, when /f/ is restricted to onset, the majority (40/75) of the errors that are analyzed involve interactions with the consonants /g,k,m,ŋ/—all of which are quite dissimilar from /f/, sharing at most voicing with the target (the remaining 35 errors involve interactions with /s/ or /h/, both of which share voicing and continuancy with /f/). It is difficult to imagine how such articulatorily dissimilar errors could be produced by the articulatory sub-component. However, even if the errors are generated by the phonological sub-component, the articulatory sub-component may produce the regularity biases (e.g., by filtering out irregular phonological errors). With this caveat in mind, the phonological sub-component is the most likely locus for the effects reported here.

Summary

The findings of this experiment suggest that sub-segmental regularities must be encoded. This is consistent with the predictions of the Markedness theory, and is inconsistent with the predictions of the Instance-Based and Lexical Distribution theories. The findings regarding scale support the encoding of categorical as well as gradient

regularities. With respect to the restricted control consonant regularity, structures were categorically ill-formed; the probability of irregular structures was nearly zero. In contrast, the restricted test segment regularity specified gradient well-formedness; ill-formed structures merely had a lower probability than well-formed structures. With respect to scope, the experiment showed that within-language regularities can be encoded; this does not distinguish between the three theories.

Experiment 2: Biases in speech errors

The previous experiment provides evidence that sub-segmental representations are used to encode regularities. However, a potential concern is that this result is specific to learning tasks. After learning has been completed, the phonological sub-component may no longer encode regularities at the sub-segmental level. Experiment 2 explores this possibility using speech errors in a tongue twister task. As discussed in section 3, there is a regularity bias in speech errors; regular structures are more likely to replace irregular structures than vice versa. This provides a tool for contrasting theories of regularities. This study assesses the three theories proposed above using pairs of consonants. For some of these pairs, the Markedness theory predicts that errors should be biased towards one consonant, while the Instance-Based and Lexical Distribution theories predicts a bias towards the other member of the pair. Thus, for each of these pairs, the two types of theories predict opposite biases in speech errors.

Note that unlike the studies reported in section 3, this design allows a direct contrast between the predictions of graded within-language and cross-linguistic regularities. Almost every study reviewed in section 3 examined the predictions of a single type of regularity, failing to control for the effects of other regularities. This study controls for these effects by contrasting the predictions of different theories of regularities. Furthermore, unlike Experiment 1, all critical consonants are distributed in the same fashion within the experimental materials. Since the statistics of the immediate environment do not distinguish the different consonants in each pair, any biases found in speech errors must already be present in the speech production system. This allows us to tap into the regularities encoded by the adult phonological sub-component, outside of a learning task.

Method

Participants

Fifty undergraduate and graduate students from the Johns Hopkins University community participated in the experiment and were compensated with \$7 or received extra-credit in introductory courses for their participation. 5 participants were excluded: 3 due to equipment failure, 1 due to failure to learn the pronunciation of nonwords used in the experiment, and 1 because her dialect of English failed to distinguish between two of the vowels used in the study (/æ/ and /ɑ/). The remaining 45 participants (12 males,

11 right-handed; 33 females, 32 right-handed) reported that they were native speakers⁵² of English and that they had no history of speech/language impairment.

Materials: Consonant pairs

Appendix B describes the method used to determine regularities for each of the three theories (i.e., references for cross-linguistic generalizations and frequency counts). For the Markedness theory, I examined cross-linguistic regularities at the sub-segmental level (these particular regularities have no contextual restrictions). Three sub-segmental contrasts were identified. For each contrast, one type of sub-segmental structure is marked (i.e., cross-linguistically irregular), while the other is unmarked (cross-linguistically regular). This is shown in Table 6.

⁵² It should be noted that several of the participants were native bilinguals and/or had extensive training in foreign languages. Languages spoken by the bilinguals in this group included: Cantonese Chinese, Farsi, French, and Mandarin Chinese. There was also one French-Spanish-English trilingual.

Table 6. Cross-linguistic regularities, Experiment 2.

<u>Sub-segmental markedness contrast</u>	<u>Associated segment pairs</u>
(unmarked > marked)	
Unvoiced stop > voiced stop	t-d; k-g
Stop > fricative	t-s; d-z
Labial stops > dorsal stops	p-k ⁵³ ; b-g

For one of the consonant pairs associated with each sub-segmental contrast, the Instance-Based and Lexical Distribution theories predict a different bias; the unmarked consonant is less frequent than the marked one. In other words, the consonant that is regular with respect to the Markedness theory is irregular with respect to the other theories. This is shown in Table 7. Since the Markedness and frequency-based theories associate opposite regularities with each consonant pair, they make contrasting predictions regarding error biases on these pairs.

⁵³ The status of the relative markedness of labial and dorsal stops is a matter of some debate. Gamkrelidze (1978) and Sherman (1975; cited by Ohala, 1983) present typological evidence that among unvoiced stops, labial is marked, and dorsal unmarked. Among voiced stops, the preference is reversed. This markedness proposal therefore differs from the predictions of the Markedness theory on the pair /p-/k/.

Table 7. Test pairs, Experiment 2.

Consonant pair with markedness contrast	Relative token frequency:		Relative type frequency:	
	Instance-Based theory		Lexical Distribution theory	
Unmarked > Marked	Unmarked	Marked	Unmarked	Marked
t > d	4.4%	5.7%	4.0%	5.2%
t > s	4.4%	11.2%	4.0%	6.9%
p > k	7.6%	8.9%	6.2%	6.9%

As noted in section 3, some studies have suggested an anti-regularity effect, where the preferred direction of errors is opposite the one reported by many other studies (i.e., irregular structure is more likely to replace regular structure than vice versa). To examine this possibility, 3 control pairs were selected. All three theories agree on the relative regularity of consonants in these pairs (statistics are shown in Table 8). Using these pairs, we can determine whether errors are biased towards regular consonants (a regularity effect) or towards irregular consonants (an anti-regularity effect). This will show whether bias in errors on the test pairs reveals the more regular or less regular consonant.

Table 8. Control pairs, Experiment 2.

Consonant pair with markedness contrast	Relative token frequency:		Relative type frequency:	
	Instance-Based theory		Lexical Distribution theory	
Unmarked > Marked	Unmarked	Marked	Unmarked	Marked
k > g	8.9%	2.6%	6.9%	2.1%
d > z	5.7%	0.1%	5.2%	0.2%
b > g	5.3%	2.6%	5.1%	2.1%

Materials: Stimulus list

For each consonant pair, two nonword bodies were selected, made up of a vowel and coda consonant. Test nonwords were formed by pairing each consonant with each nonword body (yielding four nonwords for each consonant pair). As described in Appendix C, nonword bodies were selected to control for morphological effects (e.g., neighborhood density), as well as contextual effects (e.g., transitional probability from initial consonant to vowel). Note that to control for segment frequency, two non-critical control pairs were added to the stimulus list (leading to a total of 8 consonant pairs; see Appendix C). Errors involving consonants within these pairs were not analyzed.

To generate errors on each pair, tongue twister sequences were created using the nonwords. Each sequence was composed of four nonwords generated from a single pair

of consonants and nonword bodies. Sequences were constructed to respect an alliterating pattern: the initial consonants followed one pattern, while the nonword bodies followed another. For example, one sequence was “gaysh bofe gofe baysh.” The initial consonants follow one pattern, ABAB (g-,b-,g-,b-), while the nonword bodies followed another, ABBA (-aysh,-ofe,-ofe,-aysh). This is similar to the pattern of many “natural” tongue twisters (Kupin, 1982), and is designed to induce speech errors. Eight different sequences were generated for each consonant pair. These 8 sequences were produced by crossing consonant order (e.g., g- first vs. b- first), alliteration pattern (ABAB/ABBA vs. ABBA/ABAB), and nonword body order (e.g., -aysh first vs. -ofe first). This yields a total of 64 sequences (8 sequences per consonant pair, with 8 consonant pairs total). Each participant received a randomized list of these 64 tongue twisters.

Each sequence was spelled out for visual presentation. A description of spelling conventions is provided in Appendix C.

Procedure

Stimulus presentation, trial procedure, and data recording were similar to that of experiment 1. There were two main differences. First, the experiment did not take place in a sound-attenuated chamber; participants were seated in a quiet room with the experimenter. Second, a more extensive practice session preceded the experiment trials (to ensure that participants learned the correct pronunciation for the nonword targets).

The practice session began with an extensive training session. The set of nonwords used in the experiment was broken up into 9 subsets, corresponding to the nine vowels used in the experiment. Each subset was introduced by pairing the vowel sound with its spelling (e.g., “In this set, ‘e’ is pronounced /ɛ/.”) Three examples of this vowel spelling in English were provided (e.g., “end, seven, bet”; the same examples were used for all participants). The participant was then shown each nonword in this subset (e.g., “kev, gev”). For each nonword, the participant saw its orthographic form and heard a recording of its pronunciation; s/he then repeated the nonword back to the experimenter. Feedback was provided if necessary. After all nine sets were presented, the pairing of vowel spelling and pronunciation was reviewed. A brief test was then administered. The entire set of nonwords was presented (orthographically) to the participant in two different random orders. For each nonword, the participant had to produce the correct phonological form; feedback was provided if necessary.

Following the training session, three practice trials were administered. These practice trials used sequences from the non-critical control pairs. The set of 64 experimental trials was then administered, broken up into 4 blocks. The experimenter remained with the participant throughout the experiment; if the participant forgot how to pronounce any of the nonwords during the slow repetition of the sequence, they were prompted by the experimenter. The entire experiment required approximately 25 minutes to complete.

Results and Discussion

Analysis was similar to that of experiment 1, except that voicing errors were included in the analysis (as some critical pairs contrasted in voicing) and vowel errors were transcribed. Each participant was randomly assigned to one of two transcribers. Inter-transcriber reliability was good. Both transcribers examined a section of a randomly selected recording. The overall agreement rate is 98.8% (261/264). For errors alone, the agreement rate is 70.0% (out of 10 errors identified by one or both transcribers, they agreed on 7).

Statistical analysis of bias in errors

Two statistical tests were conducted to test for a bias in errors. If there is no bias in the errors, then there should be equal numbers of regular→irregular and irregular→regular errors. The collapsed analysis was therefore a binomial test comparing the rate of regular→irregular errors to the chance rate of 50% of total errors. If the rate of regular→irregular errors is significantly less than 50%, errors are biased towards producing regular structures (i.e., regular→irregular errors are less likely than irregular→regular errors). If the rate of regular→irregular errors is significantly greater than 50%, errors are biased towards irregular structures (an anti-regularity effect). The second analysis (by participants) used a paired t-test to see if there was a significant

difference between the number of regular→irregular errors and the number of irregular→regular errors produced by each participant.

Results

The first analysis considers all errors involving consonants that occurred within any control or test pair. The results for the control pairs are shown in Table 9. The third column reports the results of the collapsed analysis, showing the percentage of total errors that produced irregular structure (i.e., regular→irregular errors).

Table 9. Performance on control pairs, Experiment 2.

Pair	Regular→Irregular errors	Irregular→Regular errors	Percent errors producing irregular
/k/-/g/	246 (5.5)	340 (7.6)	42.0%*
/d/-/z/	167 (3.7)	370 (8.2)	31.1%*
/b/-/g/	213 (4.7)	185 (4.1)	53.5%

Note: Mean number of errors per participant shown in parentheses.

*Significant difference ($p < .05$) from 50% of total errors.

Two of these pairs show a regularity effect; irregular segments are more likely to be replaced by regular segments than vice versa (/k/-/g/: collapsed, $Z = -3.8$ (continuity-

corrected), two-tailed $p < .0002$; by participants, $t(44) = -2.4$ (continuity-corrected), two-tailed $p < .02$; /d/-/z/: collapsed, $Z = -8.7$, $p < .0001$; by participants: $t(44) = -7.3$, $p < .0001$). The remaining pair, /b/-/g/, shows no significant effect (collapsed, $Z = 1.35$, $p > .17$; by participants $t(44) = .2$, $p > .80$).

The results for the test pairs are shown in Table 10. Here, the collapsed analysis (shown in the third column) examines the rate of errors that result in segments that are cross-linguistically marked but frequent within English (i.e., segments that are irregular with respect to the Markedness theory, and regular with respect to the Instance-Based and Lexical Distribution theories). If this rate is significantly less than 50%, errors are biased towards less marked structure; if the rate is greater than 50%, errors are biased towards more frequent structure.

Table 10. Performance on test pairs, Experiment 2.

Pair	Unmarked/infrequent	Marked/infrequent →	Percent errors producing marked/frequent
	→Marked/frequent	Unmarked/frequent	
/t/-/d/	248 (5.5)	234 (5.2)	51.5%
/t/-/s/	244 (5.4)	301 (6.7)	44.8%*
/p/-/k/	208 (4.6)	171 (3.8)	54.9%

Note: Mean number of errors per participant shown in parentheses.

* Significant difference ($p < .05$) from 50% of total errors.

Given the regularity effect shown in the control pairs, results for the test pair /t/-/s/ favor the Markedness theory. Marked/frequent /s/ is significantly more likely to be replaced by unmarked/infrequent /t/ than vice versa (collapsed, $Z = -2.4$, $p < .02$), but the difference is not significant by participants ($t(44) = -1.3$, $p > .20$). With respect to the by participants analysis, note that the effect size is small (difference in means = 1.3). The 0.5 continuity correction may be obscuring the small effect. Re-computation of the t-statistic without the continuity correction yields a significant difference (uncorrected $t(44) = -2.1$, $p < .05$). Results for the pair /p/-/k/ provide some support for the Instance-Based and Lexical Distribution theories: numerically, marked/frequent /k/ is less likely to be replaced by unmarked/infrequent /p/ than vice versa. This difference is marginally

significant overall (collapsed, $Z = 1.8$, $p < .07$), but the difference is not significant by participants ($t(44) = .7$, $p > .50$). As above, the t-test comparison was re-computed removing the correction for continuity. In this analysis, the effect was at best marginally significant (uncorrected $t(44) = 1.7$, $p < .10$). The other pair shows no significant differences (/t/-/d/: collapsed, $Z = .7$, $p > .45$; by participants: uncorrected $t^{54}(44) = 0.6$, $p > .50$).

The marginal effects associated with /p/-/k/ may result from perceptual errors made by transcribers. During the process of coding data, the transcribers noted the difficulty of detecting cutoff errors. Often, the initial consonant in these errors is not released (i.e., followed by aspiration or a reduced vowel), making it extremely difficult to detect. This raises the possibility that asymmetries in errors could result from perceptual errors by transcribers, not participant production errors. To control for this possibility, a further analysis was performed, excluding cut-off errors. When these errors were excluded, the marginally significant bias effect associated with /p/-/k/ disappeared ($p > .6$ for both collapsed and by participants analysis⁵⁵). In contrast to the elimination of the marginal /p/-/k/ bias, the three pairs that showed significant differences in the overall analysis (/d/-/z/, /k/-/g/, and /t/-/s/) showed significant or marginally significant differences in this

⁵⁴ Unless otherwise noted, t statistics are uncorrected because the mean difference in errors is less than the continuity correction of 0.5.

⁵⁵ 125 /p/→/k/ errors, 117 /k/→/p/ errors, $Z = .4$, $p > .60$; mean of 2.8 /p/→/k/ errors per participant, 2.6 /k/→/p/, uncorrected $t(44) = .5$, $p > .60$.

analysis⁵⁶. This suggests that, unlike the other three pairs, there is no significant bias associated with the pair /p/-/k/.

Is the bias on the three pairs due to one consonant being more regular, or to some other factor? In constructing the experimental materials, care was taken to match the different nonword targets for morphological and contextual factors. In many of the above errors, however, the nonword bodies were produced incorrectly, which may have led to differences between targets. To control for this possibility, an analysis was conducted excluding all errors where the body of the nonword was produced incorrectly (following the previous analysis, cut-off errors were excluded). The results were unchanged; the three pairs /d/-/z/, /k/-/g/, and /t/-/s/ all showed significant or marginally significant bias effects⁵⁷, while no significant bias was observed on the other pairs (collapsed analysis, Zs ranging from .1 to 1.3; by participants, uncorrected ts ranging from .05 to 1.3). This

⁵⁶ /d/-/z/: 150 /d/→/z/ errors, 289 /z/→/d/ errors, $Z = -6.6$, $p < .0001$; mean of 3.3 /d/→/z/ errors per participant, 6.4 /z/→/d/, $t(44) = -5.0$, $p < .0001$; /k/-/g/: 236 /k/→/g/ errors, 327 /g/→/k/ errors, $Z = -3.8$, $p < .0002$; mean of 5.2 /k/→/g/ errors per participant, 7.3 /g/→/k/, $t(44) = -2.4$, $p < .03$; /t/-/s/: 214 /t/→/s/ errors, 254 /s/→/t/ errors, $Z = -1.8$, $p < .08$; mean of 4.8 /t/→/s/ errors per participant, 5.6 /s/→/t/, uncorrected $t(44) = -1.6$, $p < .12$.

⁵⁷ All pairs showed significant or marginally significant overall differences in this analysis (/d/-/z/: 96 /d/→/z/ errors, 171 /z/→/d/ errors, $Z = -4.5$, $p < .0001$; ; /k/-/g/: 180 /k/→/g/ errors, 271 /g/→/k/ errors, $Z = -4.1$, $p < .0001$; /t/-/s/: 80 /t/→/s/ errors, 104 /s/→/t/ errors, $Z = -1.7$, $p < .09$). By participants, /d/-/z/ and /k/-/g/ showed significant differences (/d/-/z/: mean of 2.1 /d/→/z/ errors per participant, 3.8 /z/→/d/, $t(44) = -3.1$, $p < .004$; /k/-/g/: mean of 4.0 /k/→/g/ errors per participant, 6.0 /g/→/k/, $t(44) = -2.8$, $p < .008$). By participants, the asymmetry for /t/-/s/ failed to reach significance with the continuity correction (mean of 1.8 /t/→/s/ errors per participant, 2.3 /s/→/t/, $t(44) = -0.1$, $p > .9$), but was marginally significant without the correction (uncorrected $t(44) = -1.7$, $p < .09$).

suggests that the biases observed in the overall results do not result from morphological effects.

Another concern is the inclusion of non-contextual errors—errors involving segments that are in different sequences (e.g., /d/→/t/ errors in a sequence with /d/-/z/ targets). It is possible that these errors are not comparable; these differences may introduce bias effects. To control for this possibility, an analysis was performed excluding these errors (cut-off errors were also excluded⁵⁸). Results were unchanged; the three pairs /d/-/z/, /k/-/g/, and /t/-/s/ showed significant bias effects⁵⁹, while no significant bias was observed on the other pairs (collapsed analysis, Zs ranging from $-.05$ to 1.4 ; by participants, uncorrected ts ranging from $-.1$ to 1.3). This suggests that the biases found in the overall error analysis do not result from the influence of non-contextual errors.

Finally, this experiment shares two potential confounds with Experiment 1: the use of a paradigm involving repetition of target sequences; and the inclusion of bilingual participants. These do not appear to influence the results. Analysis of the first repetition of each target sequence do not yield significant differences (due to the low number of

⁵⁸ Due to insufficient numbers of errors, analysis eliminating all three potentially confounding factors (cut-off errors, non-contextual errors and nonword body errors) were not performed.

⁵⁹ All pairs showed significant overall differences in this analysis (/d/-/z/: 145 /d/→/z/ errors, 287 /z/→/d/ errors, $Z = -6.8$, $p < .0001$; /k/-/g/: 221 /k/→/g/ errors, 302 /g/→/k/ errors, $Z = -3.5$, $p < .0005$; /t/-/s/: 208 /t/→/s/ errors, 252 /s/→/t/ errors, $Z = -2.0$, $p < .05$). By participants, /d/-/z/ and /k/-/g/ showed significant differences (/d/-/z/: mean of 3.2 /d/→/z/ errors per participant, 6.4 /z/→/d/, $t(44) = -5.2$, $p < .0001$; /k/-/g/: mean of 4.9 /k/→/g/ errors per participant, 6.7 /g/→/k/, $t(44) = -2.2$, $p < .04$). By participants, the asymmetry for /t/-/s/ failed to reach significance with the continuity correction (mean of 4.6 /t/→/s/ errors per participant, 5.6 /s/→/t/, $t(44) = -0.9$, $p > .35$), but was marginally significant without the correction (uncorrected $t(44) = -1.8$, $p < .08$).

errors), but the direction of error biases was consistent with the results reported for /d-/z/, /k/-/g/, and /t/-/s/⁶⁰. Furthermore, exclusion of the bilingual participants did not alter results on these three pairs⁶¹.

General Discussion: Experiment 2

The results on two of the three control pairs establish a regularity effect; structure that is classified as irregular by all three theories is replaced by structure classified as regular more often than the reverse. In light of the regularity effect on the control pairs, the results for the test pairs are most consistent with the Markedness theory. The Instance-Based and Lexical Distribution theories predict the opposite bias on the test pair /t/-/s/. Since /t/ replaces /s/ more often than the reverse, we can infer that /t/ is in fact more regular than /s/—just as predicted by the Markedness theory, not by the other theories.

Can the Instance-Based and/or Lexical Distribution theories offer any account for the results on /t/-/s/? Note that the frequency counts used to determine the regularities of

⁶⁰ /d/-/z/: 25 /d/→/z/ errors, 39 /z/→/d/ errors. /k/-/g/: 53 /k/→/g/ errors, 60 /g/→/k/ errors. /t/-/s/: 43 /t/→/s/ errors, 51 /s/→/t/ errors

⁶¹ All pairs showed significant overall differences in this analysis (/d/-/z/: 147 /d/→/z/ errors, 332 /z/→/d/ errors, $Z = -8.4$, $p < .0001$; /k/-/g/: 233 /k/→/g/ errors, 310 /g/→/k/ errors, $Z = -3.3$, $p < .001$; /t/-/s/: 223 /t/→/s/ errors, 278 /s/→/t/ errors, $Z = -2.4$, $p < .02$). By participants, /d/-/z/ and /k/-/g/ showed significant differences (/d/-/z/: mean of 2.7 /d/→/z/ errors per participant, 6.0 /z/→/d/, $t(39) = -5.8$, $p < .0001$; /k/-/g/: mean of 3.4 /k/→/g/ errors per participant, 5.3 /g/→/k/, $t(39) = -2.8$, $p < .009$). By participants, the asymmetry for /t/-/s/ failed to reach significance with the continuity correction (mean of 3.7 /t/→/s/ errors per participant, 5.1 /s/→/t/, $t(39) = -1.7$, $p > .09$), but was significant without the correction (uncorrected $t(39) = -2.6$, $p < .02$). A marginally significant bias was also found for in the overall analysis for /b/-/g/ (193 /b/→/g/ errors, 155 /g/→/b/ errors, $Z = 2.1$, $p < .04$). This bias failed to reach significance in the by participants analysis (mean of 4.8 /b/→/g/ errors per participant, 3.9 /g/→/b/ errors, $t(39) = .9$, $p > .35$; uncorrected $t(39) = 1.9$, $p > .05$). Given that this difference is not significant by participants, it is unclear whether it is due to differences between monolinguals and bilinguals (as opposed to variation within the participants). Further experiments should examine this issue in more detail.

Lexical Distribution theory are based on Frisch et al. (2000), who only distinguished initial, medial, and final syllables. In their formulation of the Lexical Distribution theory, Coleman & Pierrehumbert (1997) drew finer-grained positional distinctions, distinguishing syllables that were initial and not final (e.g., /kæt/ in “Kathmandu”) from syllables that were both initial and final (e.g., /kæt/ in “cat”). Frequency statistics were therefore re-computed for the 6 consonant pairs used in the study considering only the word-initial onsets of monosyllabic words. This version of the Lexical Distribution theory agrees with the predictions of the Markedness theory on all pairs except /t/-/s/ (relative frequency of /t/ in monosyllables: 3.8%; /s/: 4.3%). This is, of course, the only test pair which showed a significant difference—in the direction predicted by the Markedness theory. The Instance-Based and Lexical Distribution theories therefore appear to be unable to account for performance on this pair. This is not say that no within-language theory could account for the data. The data from this task do not indicate whether the within-language theories fail because of their claims concerning scope (as opposed to their commitments along other features of regularities such as granularity). I return to this point in the general discussion.

Implications for scale

As with Experiment 1 and the studies reported in section 3, speech errors show a tendency to respect regularities. This suggests that irregular structures are not assigned

zero output probability by the phonological sub-component, but merely have a lower probability than regular structures. Well-formedness is defined using a gradient, rather than a categorical scale. This supports the predictions of the Instance-Based and Lexical Distribution theories, and is inconsistent with the predictions of the Markedness theory.

Accounting for null effects

How can the Markedness theory account for the null results observed for the one control pair and two test pairs? With respect to the labial-dorsal contrast, one possibility is that the null results are due to perceptual errors on the part of the transcribers. As shown above, the results on the test pair /p/-/k/ were significantly altered by the elimination of cutoff errors. The null results on the labial-dorsal pairs may therefore result from contamination by perceptual errors, not due to the properties of the phonological sub-component.

Another possibility⁶² is that we have incorrectly characterized the predictions of the Markedness theory. It may be that labial consonants are no more regular than dorsal consonants. There is some evidence to support this; as discussed in footnote 53, the typological relationship between dorsal and labial stops is unclear. If there is no difference in the regularity of labials as compared to dorsals, the Markedness theory would predict no bias in errors involving these two consonants. Note that for the pairs

⁶² This is based on a suggestion made by Colin Wilson.

that do show regularity effects, there are clear reasons to suppose the presence of a regularity relationship. With respect to /k/-/g/, there are many languages that replace /g/ with /k/ in certain positions (e.g., German devoicing; see section 1). With respect to /d/-/z/ and /t/-/s/, there are many languages that allow stops but not fricatives in certain contexts (e.g., in Spanish, stops but not fricatives are allowed following nasals, e.g., hom/b/re, *hom/β/re; Kenstowicz, 1993). Thus, for the pairs that show significant effects, there are clear cross-linguistic generalizations; for /p/-/k/, /b/-/g/, there is no clear cross-linguistic generalization. However, this account has one significant problem—the null result for /t/-/d/. /d/ is clearly less regular than /t/ (as shown by processes such as German devoicing). In sum, although the significant differences found in the experiment are consistent with a Markedness theory, it is unclear how such a theory could explain the full extent of null results.

Locus of effects in this task

Finally, it is important to assess whether the effects observed in this experiment arise within the phonological sub-component. As with the previous experiment, the use of the slow repetition condition ensures that errors are arising within the spoken form component. The materials controlled for morphological effects, eliminating that sub-component as a locus of the effects. However, since the pairs contrasted here (outside of the labial-dorsal pairs) are highly similar articulatorily, it is possible that errors can arise

within the articulatory sub-component. Two points would argue against an articulatory basis for these effects. First, contrasting results on /t/-/d/ and /k/-/g/ would not be expected; comparable movements are required for both voiceless-voiced pairs. However, since none of the proposed phonological sub-component theories can accommodate this results, this objection is a bit weak. Second, as noted in section 2, many errors in the tongue twister task appear to be generated in the phonological sub-component, supporting a phonological locus for the asymmetries observed in this task. However, the articulatory sub-component may contribute to the asymmetries by filtering out irregular phonological errors. With these caveats in mind, I will assume these results reflect the properties of the phonological sub-component.

Although the production errors may be arising in the phonological sub-component, we must still be concerned about the influence of perceptual errors of transcribers. Although some were eliminated by excluding cut-off errors (see above for analysis), it is quite likely that non-categorical speech errors were still included within the results. If perception of such errors is biased, it may contribute to the observed asymmetries. I do not believe this to be a concern. As discussed in section 3, two studies (Frisch & Wright, 2002; Pouplier & Goldstein, 2002) report that these transcriber errors exhibit an anti-regularity effect. Therefore, the presence of a regularity effect in the control pairs suggests that non-categorical errors are not causing the error biases observed here.

Summary

Some of the results of Experiment 2 favor the Markedness theory. Control pairs show a regularity effect in this task, and results on one of the test pairs fit the predictions of the Markedness theory, contradicting the predictions of the Instance-Based and Lexical Distribution theories. However, it is unclear whether the Markedness theory can account for the absence of error biases found on one control pair and two test pairs. With respect to scale, the results favor the Instance-Based and Lexical Distribution theories. Errors tend to follow regularities, but not absolutely; this suggests that gradient regularities are encoded by the phonological sub-component.

General Discussion: Experimental investigations

What types of regularities are encoded?

With respect to granularity, the Markedness theory best characterizes the types of regularities encoded in the two experiments. In experiment 1, we found that regularities are encoded over sub-segmental representations. This level of granularity is omitted from both the Instance-Based and Lexical Distribution theories. Furthermore, experiment 2 found a significant error bias that reflected the regularity associated with the Markedness theory, not the regularity associated with the other two theories.

Clearly, the Instance-Based and Lexical Distribution theories make an incorrect assumption regarding granularity. To accommodate the results of experiment 1, we must

modify these theories to include sub-segmental regularities. Once this modification has been made, the within-language theories predict that errors should be biased in the same direction as predicted by the Markedness theory. Table 11 illustrates the predictions of these modified theories for Experiment 2. Predictions were derived by contrasting the sum frequency statistics for all segments sharing the sub-segmental structure of consonants in each experimental pair.

Table 11. Sum frequency of segment classes, control and test pairs.

Pairs	Sub-segmental contrast	Sum frequency of	Sum frequency of segment
		segment class: Instance-Based theory	class: Lexical Distribution theory ⁶³
/k/-/g/;	Unvoiced stop	unvoiced stop: 20.1%	unvoiced stop: 17.1%
/t/-/d/	/k, t, p/ > Voiced stop /g, d, b/	voiced stop: 12.4%	voiced stop: 12.4%
/d/-/z/;	Stop	stop: 29.2%	stop: 29.5%
/t/-/s/	/k, g, p, b, t, d/ > Fricative /ð, θ, z, s, f, v, ʃ, ʒ/	fricative: 25.7%	fricative: 15.1%
/b/-/g/;	Labial stop	labial stop: 12.9%	labial stop: 11.3%
/p/-/k/	/p, b/ > Dorsal stop /k, g/	dorsal stop: 11.5%	dorsal stop: 9.0%

Note: ‘>’ denotes ‘less marked than.’

⁶³ The same qualitative predictions are made when only monosyllabic words are considered (following Coleman & Pierrehumbert, 1997; see above). Here, the differences are: unvoiced stops, 12.8%; voiceless stops; 11.1%; stops, 23.9%; fricatives, 12.1%; labial stops, 12.8%; dorsal stops, 11.1%.

As can be seen in the table, when within-language regularities are stated at the sub-segmental level, these theories make the same predictions as the Markedness theory; the segment class of the unmarked segment is more frequent than that of the marked segment. Furthermore, if we assume that strength of the regularity increases with the difference in frequency, we find that that weakest regularity is the labial-dorsal regularity. This is shown in Table 12. If strong asymmetries in errors are dependent on strong asymmetries in frequencies, these theories would predict the smallest asymmetries on the labial-dorsal pairs—just as observed in the data. However, it should be noted that, just as with the Markedness theory, these reformulated theories cannot account for the null effect on /t/-/d/.

Table 12. Difference in sum frequency of segment classes for control and test pairs.

Pairs	Sub-segmental contrast	Difference in sum frequency of classes: Instance-Based theory	Difference in sum frequency of classes: Lexical Distribution theory ⁶⁴
/k/-/g/;	Unvoiced stop >	7.7%	4.7%
/t/-/d/	Voiced stop		
/d/-/z/;	Stop >	3.5%	14.4%
/t/-/s/	Fricative		
/b/-/g/;	Labial stop >	1.4%	2.3%
/p/-/k/	Dorsal stop		

Note: '>' denotes 'less marked than.'

The results from the two experiments can therefore be accommodated equally well by within-language and cross-linguistic regularities, as long as the regularities are stated at the sub-segmental level of granularity.

⁶⁴ Different quantitative predictions are made when only monosyllabic words are considered (following Coleman & Pierrehumbert, 1997; see above). Here, the differences are: voiced vs. voiceless stops, 1.8%; stop vs. fricative, 11.8%; labial vs. dorsal stops, 3.0%. Thus, contrary to the data, the weakest effect is predicted for the voiced-voiceless contrast instead of the labial-dorsal contrast.

With respect to scale, the results clearly favor the Instance-Based and Lexical Distribution theories. Results from Experiment 1 show that gradient as well as categorical regularities can be encoded; for the restricted test segment, errors tend to respect the associated regularity, but do not do so absolutely. Similar results are found in Experiment 2, suggesting that gradient well-formedness distinctions are encoded by the phonological sub-component.

To account for the results, theories of regularities must incorporate elements of all three theories. The Instance-Based and Lexical Distribution theories are correct in assuming that gradient regularities are encoded; the Markedness theory is correct in assuming that sub-segmental regularities are encoded. The results regarding scope are unclear; conclusions regarding this feature of regularities must rely on data from future experiments.

Limitations of this study

As noted above, the experimental studies reported here are not without their limitations. First, the results fail to constrain theories with respect to scope of regularities. Second, it is unclear why there is no significant bias associated with three of the consonant pairs in experiment 2. Third, the use of transcription-based data collection raises the possibility that some of the results are due to biases in transcribers. Future

studies should examine these issues to provide a more complete picture of processing within the phonological sub-component.

Extensions to this study

Some concrete steps can be taken to extend this study and overcome its limitations. First, recall that Experiment 1 had a gradient sub-segmental regularity associated with the restricted test and unrestricted similar consonant. The sub-segmental features associated with these consonants were biased towards one syllable position (that of the restricted test segment), but occurred in the other syllable position 25% of the time. This gradient regularity had no effect on the unrestricted related segment (i.e., it behaved no differently from other unrestricted segments). One possible reason for this null result is that the regularity is too weak to exert an independent effect on errors. To test this possibility, we could increase the strength of the regularity by halving the frequency of the unrestricted related consonant. This would decrease the frequency of its sub-segmental features by half in both positions, cutting the number of exceptions to the gradient regularity in half. For example, in the condition where /s/ is restricted to coda and /z/ unrestricted, the frequency of /z/ could be halved. This would mean that the associated sub-segmental features [continuant] and [coronal] would be found in onset position only 12.5% of the time, reducing the number of exceptions to the sub-segmental regularity by half.

Once all theories incorporate the assumption that sub-segmental regularities are encoded, the data from Experiment 2 cannot distinguish their predictions. This may be possible in subsequent experiments; specifically, some consonant pairs can distinguish the Instance-Based theory from both the Markedness theory and Lexical Distribution theory. Cross-linguistically, dorsal stops are marked relative to coronal stops, so the Markedness theory predicts an asymmetry favoring coronal stops. The Lexical Distribution theory makes the same predictions: dorsal stops are less frequent than coronal stops (sum relative frequency of coronal stops: 9.2%; dorsals: 9.0%⁶⁵). In contrast, for the Instance-Based theory, dorsal stops are more frequent than coronal stops (sum relative frequency of coronal stops: 10.1%; dorsals: 11.5%). Thus, these two groups of theories make different predictions for the consonant pairs /t-/k/ and /d-/g/. Error biases on these pairs would provide a further contrast between the three theories.

Summary

The experiments reported here provide support for a specific level of granularity—the sub-segmental level. Studies reviewed in section 3 had suggested that regularities at this level of granularity were encoded, but failed to examine whether correlated regularities at other levels of granularity could account for the data. By directly contrasting the predictions of segmental and sub-segmental regularities, the experiments have shown that

⁶⁵ The same prediction is made when only monosyllabic words are considered (following Coleman & Pierrehumbert, 1997; sum frequency coronal stops: 7.1%; dorsals: 6.9%).

sub-segmental regularities must be encoded. Future work, examining other contrasts between different theories of regularities, should provide further evidence regarding the types of regularities encoded by the spoken production system.

SECTION 5: GENERAL DISCUSSION

The symmetries that enchant us may be no more than good tools—
compact ways for brains to store information.

(Johnson, 1996: 324)

In this section, I discuss the implications of the results discussed in sections 2, 3, and 4 for theories of phonological processing mechanisms—that is, the processes that implement that phonological sub-component. After I discuss these implications, I briefly discuss some possibilities for extending this line of research.

Mechanisms of phonological processing

In section 2, I presented a framework for spoken production processing. In this framework, phonological regularities are encoded through variation in output probabilities of the phonological sub-component. This framework was specified at a functional level; although I assumed that some set of mechanical processes would implement the phonological sub-component, I did not specify the nature of these processes. What mechanisms are needed to process phonological representations? How is variation in output probabilities implemented computationally?

To examine these questions, I'll first discuss a set of 3 constraints on the mechanisms of phonological processing. In light of these constraints, I'll review how current

processing theories fail to satisfy them; I'll then discuss how one of these theories could be extended to incorporate all three constraints.

Constraints on phonological processing mechanisms

The results from studies in the literature (sections 2 and 3) as well as the experiments reported here (section 4) suggest 3 basic constraints on theories of phonological processing.

1. Multiple levels of phonological structure must be represented. Section 2 reviewed evidence supporting sub-segmental, segmental, and supra-segmental representations within the phonological sub-component. Results from speech errors, acquired speech impairments, and reaction time studies suggest that these levels of structure play an important role in phonological processing.
2. Regularities over these multiple levels of phonological structure must be encoded. Knowledge of phonological regularities clearly influences spoken production processing. Section 3 reviewed evidence suggesting regularity effects at multiple levels of structure. Section 4 provided evidence specifically supporting the encoding of sub-segmental regularities.

A third constraint has not been discussed previously, but is nonetheless crucial to any theory of phonological processing.

3. Phonological processing must encode serial order. A central problem in understanding behavior is explaining how behaviors are produced in sequence (Lashley, 1951). This can be seen at the most basic level of speech production, which involves the production of sequences of gestures. At higher levels, there is evidence that phonological representations themselves are retrieved in sequence. For example, Wheeldon & Levelt (1995) present evidence from a self-monitoring task suggesting that participants gain access to phonological strings on a syllable-by-syllable basis. These data highlight the importance of serial order for theories of phonological processing mechanisms.

Previous accounts of phonological processing mechanisms

There are a number of proposals for mechanisms that implement the phonological sub-component. Unfortunately, none of these satisfies all three constraints; current theories manage, at best, to satisfy two out of the three. Here, I present a brief overview of these theories, concentrating on a representative example of each class.

Regularity-based theories

The core assumption made by these theories is that regularity effects arise because phonological processing must “elaborate” or “specify” the abstract phonological material it receives as input (Béland, Caplan & Nespoulous, 1990; Butterworth, 1992; Caplan, 1987; Garrett, 1982, 1984; Kohn & Smith, 1994; Stemberger, 1985a, b; Wheeler &

Touretzky, 1997)⁶⁶. I focus on Wheeler & Touretzky's proposal, as it provides the most detailed account of the mechanisms that encode phonological regularities.

Wheeler & Touretzky assume that the phonological sub-component receives as input a specification of segments and their linear order (e.g., for “after,” /æ f t ə/; the segmental representations are assumed to contain their sub-segmental structure). The phonological sub-component elaborates this representation by specifying the supra-segmental constituents appropriate to this string (e.g., two syllables, one with /æ/ as the peak and /f/ as the coda; the other syllable with /ə/ as the peak and /t/ as the onset). This specification procedure occurs in two steps. First, for each vowel, a syllable unit is generated, and the vowel is assumed to be its peak; all consonants are then associated to the onset and coda positions of every syllable. This results in a complex representation (e.g., /f/ and /t/ would both be associated to the first and second syllables as both onset and coda). To reduce the complexity of the representation, the initial generation of syllable structure is followed by parallel application of a set of “licensing constraints.” These constraints specify that phonological structures must respect (for example): linear order (e.g., /f/ should precede /t/); binding relationships (e.g., /f/ should not be an onset and coda of the same syllable); as well as phonological regularities (e.g., /ft/ should not

⁶⁶ These theories are the most straight-forward “processing” interpretation of generative grammars, which define the phonological component of the grammar as a function mapping abstract underlying representations to fully-specified surface representations (Chomsky & Halle, 1968, et seq.).

be the onset of syllable). Structures that violate these constraints are eliminated, resulting in a representation with appropriate syllabic and segmental structure.

Errors arise in this system due to disruption of the application of licensing constraints. For example, failure to correctly apply the linear order constraint can result in segments being reversed in the output sequence. However, due to their parallel application, other licensing constraints may correctly apply and produce a well formed representation. This allows for regularity effects to arise; constraints specifying phonological regularities will eliminate those structures that violate regularities.

Although this system encodes phonological regularities (constraint 2), and it represents many levels of phonological structure (constraint 1), it fails to provide an account of serial order. The entire phonological representation is processed in parallel by the licensing constraints; this system does not produce phonological structures in a sequential fashion. Therefore, as currently implemented, these theories fail to satisfy constraint 3 (serial order).

Emergentist theories

Another class of theories focuses on mechanisms that encode regularities and serial order, claiming that phonological structure will emerge from the interaction of these mechanisms (Dell, Juliano, & Govindjee, 1993; Gupta & Dell, 1999; Joanisse, 2000;

Plaut & Kello, 1999). I focus on the proposal of Dell et al. (1993), as it has been the focus of extensive analysis (see below).

Dell et al. propose a system that is implemented in a neural network; in such a network, representations are instantiated as patterns of activity over simple processing units (see the discussion of spreading activation networks in section 2, above). The input representation is either a pattern chosen randomly for each word (i.e., a random set of activation values), or a pattern that is correlated with the desired output pattern. From this representation, a sequence of sub-segmental representations corresponding to each segment of the target word is generated.

A recurrent neural network is used to generate the sequence (for a detailed description, see Dell et al., 1993, as well as Elman, 1990; Jordan, 1986). Like other neural networks, processing in a recurrent neural network involves passing activation values along connections between different sets of simple processing units. In Dell et al.'s network, activation values pass from the input representation to a set of intermediate ("hidden") units, and then on to the output units. To produce sequences, the recurrent network makes use of context units containing the activation of the hidden and output units during the previous time step of processing. Activation flows from these context units to the hidden and output units, making the network sensitive to previous network states. Sensitivity to these previous states allows the network to produce sequences.

Constraint 3 is clearly satisfied by this system, as phonological structures are produced in a serial order. Constraint 2 is also satisfied to a certain degree. Part of how a recurrent network solves the problem of generating sequences is by encoding the sequential dependencies found in items in its training set. Phonological regularities that can be described as sequential dependencies are therefore easily learned by this network. For example, a system trained on English words will encode that vowels are unlikely to be followed by /h/ (as /h/ would never follow vowels in the training set).

There are two problems with such an approach. First, it is not clear whether recurrent neural networks are powerful enough to encode the full range of phonological regularities (i.e., can all regularities be reduced to sequential dependencies?). A second, more concrete, concern is that these systems do not appear to adequately satisfy constraint 1. The only structures explicitly represented in Dell et al.'s theory are sub-segmental (the units of the output representation) and segmental (in that the members of the output sequence correspond to segments). What of supra-segmental representations?

Dell et al. claim that such supra-segmental representations need not be explicitly represented; they will arise as the network learns to generate sequences. To support this claim, they show that many errors made by the network appear to involve the supra-segmental constituent rime (vowel and coda of a syllable) whereas very few involve adjacent segments that are not part of the same constituent (e.g., onset and vowel of a

syllable). They claim that this effect arises as the network acquires sequential dependencies; the dependencies between vowels and codas are much stronger than those between onsets and vowels. Thus, a supra-segmental effect emerges even though no supra-segmental structure is explicitly encoded.

However, analyses by Anderson, Milostan, & Cottrell (1998) argue against this interpretation. They generated outputs by adding Gaussian noise onto each member of the output sequence for a word. This method reveals error patterns that result purely from chance distortions of the output; no phonological regularities or phonological structure are encoded by the error-generating process. Anderson et al. found that errors generated by the noise method exhibited the same “rime effect” as the errors generated by the recurrent network. This suggests that the rime effect reported by Dell et al. is not a consequence of the organization of the network. Thus, it is unclear whether the emergentist approach can successfully satisfy constraint 1. At least for this specific instantiation of the emergentist position, supra-segmental phonological structure does not appear to emerge spontaneously from mechanisms encoding serial order and phonological regularities.

Encoding theories

The final set of theories focuses on how phonological structure at multiple levels is encoded for production. In doing so, they address the serial order question, making use

of well-articulated phonological representations. The question of phonological regularities, however, is not addressed in any general fashion⁶⁷. Encoding theories have offered two general solutions to the problem of serial order. The first solution is to associate sequences of phonological structures with a “frame” or structural schema. The frame encodes serial order, retrieving the appropriate phonological structures in sequence (Dell, 1986; Dell, Burger, & Svec, 1997; Levelt, Roelofs, & Meyer, 1999; Meijer, 1994; Roelofs, 1997; Shattuck-Hufnagel, 1987, 1992). Here, I focus on an alternative solution (Harris, 2002; Hartley & Houghton, 1996; Vousden et al., 2000). In particular, I focus on the proposal of Harris (2002, building on Vousden et al., 2000).

The theory of Harris is similar to that of Dell et al. (1993), in that phonological processing takes as input a memory representation (discussed below) and yields a sequence of sub-segmental representations (referred to here as the output sequence). Like Dell et al., each member of this output sequence corresponds to a segment in the target word. However, this theory uses quite different mechanisms to produce this sequence, and a very different input representation; I sketch these below.

⁶⁷ Some phonological regularities are encoded by these theories, but none offer a general mechanism to implement regularities. Several theories encode the irregularity of certain sequences by omitting units for such sequences. For example, Dell (1986) indexes segments by syllable position, so /ŋ/ errors cannot occur in onset (they are simply impossible). This is not a general solution to the problem of regularity, as many regularities are encoded in a gradient fashion (as shown in sections 3 and 4). Levelt et al. (1999) propose that syllable frequency is encoded; however, it is unclear how other regularities would be implemented. Similarly, Hartley & Houghton (1996) encode sonority sequencing constraints in their system, but omit other types of regularities.

First, the to-be-produced sequence is stored in a memory representation; this representation associates each member of the output sequence with an index⁶⁸. These indices specify the order of items within the sequence. To illustrate, imagine the sequence is stored using a stopwatch. Each member of the sequence could be associated with a number on the minute hand (e.g., for “dog”, /d/⁶⁹ is associated with 1 minute, /a/ with 2 minutes, and /g/ with 3 minutes). To retrieve a member of the output sequence, all one requires is the correct index (e.g., to find /a/, one must know that it is associated with 2 on the minute hand). To generate sequences in order, the memory representation is combined with an automatic procedure for generating these indices. For example, to retrieve “dog”, the automatic clockwork mechanisms of the stopwatch could be used. Once the stopwatch has been started at 0, it will automatically advance; the appropriate index for each member of the sequence will be generated as it counts off each minute. In order to produce a sequence in a given order, then, the phonological sub-component takes as input the memory representation, automatically generates each index, and outputs the correct member of the sequence.

To encode complex phonological representations, Harris adds structure to the indices used in the memory representation. Members of the output sequence that are structurally

⁶⁸ The indices are generated by combining the output of a set of repeating and non-repeating oscillators (see Vousden et al., 2000, for a detailed characterization).

⁶⁹ The same vector representing /d/ is associated with all words that contain /d/; the only difference between words with the same segments is the particular order they use the segments.

similar have similar indices. For example, suppose that we use the combination of the minute and second hand as an index. Let each minute represent a different syllable; within each syllable, onset is associated with 15 seconds, the vowel with 30 seconds, and the coda with 45 seconds. These indices encode structural similarity with the position of the second hand. Although each segment has a unique index (i.e., a unique combination of the minute and second hands), segments that are in similar syllabic positions will have similar indices (i.e., they will share the second hand position).

Serial order errors are caused by noise in the retrieval system. When this noise is introduced, the index cannot correctly access the memory representation; instead of retrieving the appropriate member of the output sequence, an item associated with a similar index is retrieved. Since similar indices are assigned to structurally similar segments, serial order errors will tend to respect target syllable position (as real speech errors do).

Although Harris' proposal tackles the serial order problem as well as the encoding of phonological structure at multiple levels, it does not contain a mechanism for encoding regularities. As with other encoding theories, phonological structure serves only as an index for the serial ordering mechanism. For example, Harris' system stores only the similarity between segments in the same syllable position; no other information about

phonological structure is stored. Thus, these theories fail to meet the second constraint on theories of phonological processing.

A proposal to satisfy all three constraints

Previous approaches have addressed two out of the three constraints; no single proposal has incorporated all three. Here, I outline an extension to Harris' (2002) proposal that may be capable of satisfying all of the constraints.

Recall that Harris' proposal builds on that of Vousden et al. (2000). Harris solves a problem associated with Vousden et al.'s proposal; namely, that their system requires a prodigious amount of memory to store sequences. Encoding each member of the sequence requires storing two items (the output sequence member—the sub-segmental representation to be output—as well as its index/position in the sequence); furthermore, the representation must store that these two items are associated. For sequences of any reasonable length, this can require storing a large amount of information. To confront this problem, Harris uses a memory representation that compresses this association—a holographic reduced representation (HRR; Plate, 1995). The HRR compresses the representation of the two items and their association so that it requires the same amount of memory as storing either item alone. For example, suppose the index and output sequence member were encoded as 3 numbers (e.g., index: 0 1 0; output sequence member: 0 1 1). The HRR can encode these two items and their association as a single

set of 3 numbers (e.g., 1 0 1; see Plate, 1995, for details). The use of HRRs addresses the memory problems of Vousden et al.'s proposal.

However, the memory-saving properties of HRRs come with a price. Due to the compression of information, the retrieval process will tend to be noisy. When an index is used to retrieve an output sequence member from an HRR, it will often retrieve a distorted version of the associated output sequence member. For example, if the output sequence member is (0 1 1), the retrieval process might yield (.2 .8 .7). This distorted representation must therefore be "cleaned up" using a memory system storing the possible output sequence members. By "possible" sequence, I refer to the set of representations that the system would store in the HRR (for example, if it was storing spellings, the set of representations could be the letters of English). In the simple example of strings of three numbers, if the HRR was used to store either (0 1 1) or (1 0 1), a simple memory system storing these items could simply pick the item most similar to the distorted representation; (.2 .8 .7) is clearly more similar to the correct output (0 1 1) than to the alternative (1 0 1). The memory system used by Harris' follows this "nearest-neighbor" principle.

An alternative to this simple clean-up system comes from the connectionist literature. There exists a family of network architectures that use optimization to solve the problem of generating a stored representation on the basis of incomplete and/or noisy input (e.g.,

attractor networks: Hopfield, 1982, 1984; stochastic Boltzmann machines: Ackley, Hinton, & Sejnowski, 1985; harmony networks: Smolensky, 1986). Here, I focus on the harmony network architecture of Smolensky (1986). Harmony networks store representations through a system of soft constraints that prefer network states corresponding to stored network states. Computation in a harmony network involves optimizing over these constraints; in other words, the network will generate the output state that best satisfies the constraints (given the input). This can be used to solve the clean-up problem above. Since the constraints of the harmony network encode the properties of stored states, optimization will yield the stored state that is most consistent with a given noisy input.

We can re-conceptualize the specific clean-up problem posed by the HRRs of Harris' system within this optimization framework. This architecture generates sequences of phonological representations; for English speakers, these sequences will be made up of possible words of English. This set includes words like /dag/ as well as nonwords like /dav/, but excludes nonwords like /ɲag/. The constraints that define this set are the phonological regularities of English (e.g., /ŋ/ can only occur at the end of words). We could therefore build a harmony network whose constraints were the phonological regularities of English; distorted phonological representations could be cleaned up using these regularities. In terms of memory requirements, this would improve over Harris'

system; rather than store the enormous set of representations corresponding to possible words of English, the harmony network would store the regularities that define this set. In addition, it would provide a source of regularity effects in phonological processing. For example, if any errors occur in the retrieval process, the harmony network will eliminate representations that violate phonological regularities.

This augmentation of Harris' proposal potentially satisfies all three constraints placed on theories of phonological processing. Harris' basic proposal satisfies the first and third constraints, tackling the problems of encoding phonological representations and serial order. The second constraint is met by the addition of a harmony network, which cleans up the retrieved memory representations using phonological regularities. An attractive feature of this proposal is that it provides a functional motivation for the encoding of phonological regularities; phonological regularities must be encoded to compensate for the compression of phonological representations within memory.

In future work, I plan to specify this proposal in greater detail, both functionally and computationally. Until the proposal is specified in this way, many questions remain. Is similarity between indices a sufficiently powerful mechanism for encoding supra-segmental representations? Within the context of cleaning up distorted memory representations, is a harmony network capable of encoding the full range of phonological

regularities⁷⁰? What mechanisms will allow the harmony network to adapt to new regularities in the environment? With these caveats in minds, this is the first proposal that may be capable of satisfying the three constraints on theories of phonological processing mechanisms.

Future directions

The preceding discussion has suggested many avenues for extensions of this work. One result from Experiment 1 was that a weak gradient sub-segmental regularity did not exert a significant influence on error patterns; testing with more robust gradient regularities should allow us to determine if any gradient regularities can be encoded. With respect to Experiment 2, examination of error biases in additional consonant pairs may distinguish the Instance-Based theory from the Lexical Distribution and Markedness theories. Finally, within this section, I have sketched a theory of mechanisms implementing the phonological sub-component; in future work, I plan to specify the functional and computational details of this proposal.

Finally, it is important to acknowledge that this work has touched on just a tiny fraction of the enormous space of regularities that are found in language. Having explored sub-segmental regularities, I plan to turn to regularities at higher levels of

⁷⁰ For example, one problem is that most clean-up memories used with HRRs (e.g., those used by Plate, 1995) are sensitive solely to properties of the output sequence member. This is clearly insufficient for encoding regularities; if the harmony network is only sensitive to the output sequence member, it cannot be sensitive to regularities that involve syllable structure (as these are encoded in the indices). However, there does not appear to be any principled reason (other than efficiency) to prevent the harmony network from being sensitive to properties of the index as well.

linguistic structure (e.g., supra-segmental). As with the sub-segmental level, different linguistic theories disagree as to the nature of supra-segmental regularities, providing a rich set of hypotheses for future experiments. For example, the Instance-Based theory postulates that transitional probabilities are encoded. In contrast, the Lexical Distribution theory does not encode transitional probabilities, but does encode the frequency of onset and rime syllable constituents. Testing these contrasting positions will further constrain theories of what regularities are encoded.

Conclusion: Patterns of sound, patterns in mind

Regularities in phonological structure are reflected by the language production system. First, we have seen that regularity is actively maximized by the phonological sub-component. In Experiments 1 and 2, we found that errors are biased by regularity; regular structure is more likely to be produced than irregular structure. Furthermore, the language production system actively encodes new regularities; in Experiment 1, we found that regularities within the experiment were encoded by the language production system. These results are consistent with many other studies (reviewed in section 3). The novel result from the experiments reported here is the clear evidence that sub-segmental representations are used by the spoken production system to encode regularities. Although previous studies (reviewed in section 2) have shown that sub-segmental representations are part of the phonological sub-component, no studies have shown that

such representations were necessary to encode regularities. The data reported in section 4 cannot be explained by theories that do not encode regularities at the sub-segmental level. By comparing and contrasting specific alternative theories, we have begun to map the portion of phonological regularity space that is encoded by the spoken production system.

APPENDIX A:

CHANCE RATES OF VIOLATING TARGET SYLLABLE POSITION,

EXPERIMENT 1

Restricted Control Consonants

In each sequence, there are seven possible positions that a consonant can appear in as an error (four syllables times two positions yields 8 positions total; one is subtracted because it is the target position). An error can occur in 4 out of these 7 positions (57.1% of positions) and not share syllable position with the target. This is the probability that the restricted control consonant will violate target syllable position by chance.

Unrestricted Consonants

Two of the unrestricted consonants are members of a voiced-voiceless pair (e.g., /k/-/g/). Since voicing errors are excluded from the analysis, the chance level for unrestricted consonants must reflect this exclusion. This was done by subtracting 1 error position that violates target syllable position from 2/3 of the trials. For the remaining 1/3 of the trials, 1 error position that respects target syllable position is subtracted. The proportions of subtraction for each type of position result from the design of the set of consonants. In each set, there are 4 consonants restricted to onset or coda. The remaining 4 unrestricted consonants are therefore distributed among 2 onset and 2 coda positions. If one unrestricted consonant appears in syllable position X, the other

unrestricted consonants occur equally frequently in the remaining 3 syllable positions.

Thus, any pair of unrestricted consonants (such as the voiced-voiceless pair) share syllable positions on $1/3$ of the trials.

Applying the corrections, each member of the voiced-voiceless pair can appear in 6 possible positions in a sequence. On average, $3\frac{1}{3}$ of these will violate target syllable position and $2\frac{2}{3}$ will respect it. Summing over both members of the voiced-voiceless pair and the other unrestricted consonant yields an average of 19 possible positions (6 for each member of the voiced-voiceless pair plus 7 for the other unrestricted consonant); $10\frac{2}{3}$ ($3\frac{1}{3}$ for each member of the voiced-voiceless pair + 4 for the other consonant) of these will violate target syllable position. Thus, by chance, 56.1% of the errors on unrestricted consonants will violate target syllable position.

Restricted Test Consonants

To correct for the presence of the unrestricted similar segment (which shares voicing with the restricted test consonant), 1 error position which violates target syllable position must be subtracted from $1/2$ of the trials; for the other $1/2$, 1 error position respecting target syllable position is subtracted. The proportions of subtraction for each position result from the design of the set of consonants (the unrestricted similar segment occurs equally frequently in both onset and coda). Applying the correction, the restricted segment can appear in 6 possible positions in a sequence. On average, $3\frac{1}{2}$ of these will

violate target syllable position and 2 1/2 will respect it. The chance of an error violating target syllable position is therefore 3.5/6 or 58.3%.

APPENDIX B:

DETERMINATION OF REGULARITIES FOR EACH THEORY,

EXPERIMENT 2

For the Markedness theory, I collected a set of cross-linguistic sub-segmental generalizations from the markedness literature (specifically, from Greenberg, 1966; Kenstowicz, 1994; Lombardi, 1999; Maddieson, 1984; Paradis & Prunet, 1991b).

The CELEX lexical database (Baayen et al., 1995) was used to calculate the predictions of the two frequency-based theories. Frequency was based on the frequency of “words” in CELEX, where “word” was defined as CELEX entries that did not contain all capital letters (or all capital letters and “s”), spaces, or punctuation marks (this was done to exclude acronyms and phrases). Frequency counts for CELEX entries are based on the COBUILD/Birmingham corpus, which contains both written and spoken texts of British and American English from a variety of sources. All words were counted in type frequency counts, and all words were assumed to have a log token frequency of at least 1. Frequency counts used here summed across all entries that were spelled and pronounced in the same way. Pronunciations used were the primary pronunciations listed in CELEX, corrected for the omission of “ghost-r” segments in British English (such ghost segments were replaced with /r/).

The Instance-Based theory required calculation of the frequency of segments in word-initial position. Following Luce et al. (2000), the log (base 10) weighted frequency of each word in the database that began with a given segment was summed. This provided the token frequency of the initial segment. From these raw frequency measures, the relative frequencies of word initial segments were calculated. The frequencies of all the word-initial phonemes were summed; the frequency of each individual phoneme was then divided by this number to provide relative frequency. Second, the Lexical Distribution theory required calculation of the frequencies of different onsets of stressed initial syllables. Following Frisch et al. (2000), this was calculated by counting the number of words with a stressed initial syllable with a particular onset. “Null” was not counted as an onset; that is, onsetless words were excluded from the count. This provided the type frequencies of onsets of stressed initial syllables. The frequencies of all of these onsets were then summed, and the frequency of each individual onset divided by this number to yield the relative frequency of each onset.

APPENDIX C: PREPARATION OF MATERIALS, EXPERIMENT 2

In selecting the nonword bodies, attempts were made to control for potential confounding effects. Two controls were introduced for effects of the morphological sub-component. First, the nonword bodies had to form nonwords when paired with each consonant in a pair. Second, within each pair, the two nonwords resulting from pairing each consonant with a nonword body were matched for neighborhood characteristics. Specifically, the resulting nonwords were matched in terms of the number of monosyllabic CVC words in CELEX (Baayen et al., 1995) sharing: one, the initial consonant-vowel sequence; and two, the initial and final consonants. (Note that since the two nonwords share bodies, they are equated in terms of properties of the vowel and final consonant.)

A second control attempted to eliminate contextual effects. This control made use of the Instance-Based theory's claim that regularity is based partly on transitional probabilities. Vowels in the nonword bodies were selected to control forward and backward transitional probabilities across the two consonants in each pair (calculated using token frequency). To control transitional probabilities within the stimulus set, the same consonant was never paired with the same vowel across pairs (e.g., if a nonword body used with the /t/-/d/ pair starts with /a/, no nonword body used with the /t/-/s/ pair starts with /a/).

Mean statistics for the nonwords used in each control and test pairs are shown in Tables A1 and A2. It was not possible to perfectly match nonwords on all variables (differences in variables were examined statistically using paired t-tests). For the control pairs, the nonwords formed by pairing regular consonants with the nonword bodies had a significantly higher mean backward transitional probability ($t(5) = 3.3$, two-tailed $p < .02$), as well as a significantly higher number of words sharing the initial consonant and vowel ($t(5) = 5.3$, $p < .004$). For the test pairs, the nonwords formed by pairing the unmarked/infrequent consonants with the nonword bodies had a significantly higher forward transitional probability ($t(5)=2.8$, $p < .04$). All other differences were not significant ($t(5)s < 1$).

Table A1. Mean statistics for control pairs, Experiment 2.

Consonant type	Forward transitional probability	Backward transitional probability*	Number of words sharing both initial consonant and vowel*	Number of words sharing both initial and final consonants
Regular	2.9%	3.8%	8.5	2.7
Irregular	2.9%	1.6%	3.8	1.8

* $p < .05$

Table A2. Mean statistics for test pairs, Experiment 2.

Consonant type	Forward transitional probability*	Backward transitional probability	Number of words sharing both initial consonant and vowel	Number of words sharing both initial and final consonants
Unmarked/Infrequent	3.0%	6.1%	7.3	2.3
Marked/frequent	2.2%	5.6%	6	2.8

*p < .05

Note that these statistics are not matched across pairs; therefore, all statistical comparisons of error rates are done within pairs.

A further control equated frequency of segments within the experiment, so as not to introduce any regularity differences within the stimulus set. This was an issue because several consonants occurred in two pairs, while others (/p, b, s, z/) occur in only one test or control pair. To equate all consonants for absolute frequency in word initial position, a set of non-critical control pairs (/p/-/b/ and /s/-/z/) was introduced. Two nonword bodies were created for each of these pairs, although no attempt was made to control transitional

probabilities or neighborhood density of these items (as such, errors involving consonants within each pair were not analyzed; e.g., /p/→/b/ errors were not analyzed).

Consonants in test and control pairs were spelled using single consonants. Final consonants were spelled using single consonants, except for /ʃ/, spelled “sh”, /θ/ spelled “th” and /tʃ/ spelled “tch.” Vowels were spelled as shown in Table A3.

Table A3. Spelling of vowels, Experiment 2 stimuli.

Vowel	Spelling	Example Nonword Pronunciation	Example Nonword Spelling
/aɪ/	ai, ai-e	/paɪʃ/	paish
		/saɪv/	saive
/ɛ/	e	/kɛv/	kev
/eɪ/	ay	/beɪʃ/	baysh
/aʊ/	ow	/taʊʃ/	towsh
/ɑ/	o	/dɑtʃ/	dotch
/oʊ/	o-e	/koʊf/	kofe
/æ/	a	/dæθ/	dath
/i/	ee	/dɪb/	deeb
/ɔɪ/	oy	/tɔɪtʃ/	toytch

REFERENCES

- Abd-El-Jawad, H. & Abu-Salim, I. (1987). Slips of the tongue in Arabic and their theoretical implications. Language Sciences, *9*, 145-171.
- Ackley, D. H., Hinton, G. E., & Sejnowski, T. J. (1985). A learning algorithm for Boltzmann machines. Cognitive Science, *9*, 147-169.
- Allen, M., & Badecker, W. (2001). Morphology: The internal structure of words. In B. Rapp (Ed.) The handbook of cognitive neuropsychology: What deficits reveal about the human mind (pp. 233-262). Philadelphia: Psychology Press.
- Anderson, K., Milostan, J., & Cottrell, G. W. (1998). Assessing the contribution of representation to results. In M. A. Gernsbacher & S. J. Derry (Eds.) Proceedings of the 20th annual conference of the Cognitive Science Society (pp. 48-53). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Anderson, S. R. (1981). Why phonology isn't "natural." Linguistic Inquiry, *12*, 493-539.
- Anderson, S. R. (1985). Phonology in the twentieth century. Chicago: University of Chicago Press.
- Archangeli, D., & Pulleyblank, D. (1994). Grounded phonology. Cambridge, MA: MIT Press.

- Baars, B. J. (1992). A dozen competing-plans techniques for inducing predictable slips in speech and action. In B. J. Baars (Ed). Experimental slips and human error: Exploring the architecture of volition (pp. 129-150). New York: Plenum.
- Baars, B. J., & Motley, M. T. (1974). Spoonerisms: Experimental elicitation of human speech errors. Abstract in Catalog of Selected Documents in Psychology, 4, 118.
- Baars, B. J., Motley, J. T., & MacKay, D. (1975). Output editing for lexical status from artificially elicited slips of the tongue. Journal of Verbal Learning and Verbal Behavior, 14, 382-391.
- Baayen, R.H., Piepenbrock, R., & Gulikers, L. (1995). The CELEX Lexical Database (Release 2) [CD-ROM]. Philadelphia: Linguistics Data Consortium.
- Bailey, T. M., & Hahn, U. (2001). Determinants of wordlikeness: Phonotactics or lexical neighborhoods? Journal of Memory and Language, 44, 568-591.
- Battistella, E. L. (1996). The logic of markedness. Oxford: Oxford University Press.
- Beckman, M. E., & Edwards, J. (2000). Lexical frequency effects on young children's imitative productions. In M. B. Broe & J. B. Pierrehumbert (Eds.) Papers in laboratory phonology V: Acquisition and the lexicon (pp. 208-218). Cambridge: Cambridge University Press.

- Béland, R. (1990). Vowel epenthesis in aphasia. In J.-L. Nespoulous & P. Villiard (Eds.), Morphology, phonology, and aphasia (pp. 235-252). New York: Springer-Verlag.
- Béland, R., Caplan, D., & Nespoulous, J.-L. (1990). The role of abstract phonological representations in word production: Evidence from phonemic paraphasias. Journal of Neurolinguistics, *5*, 125-164.
- Béland, R., & Favreau, Y. (1991). On the special status of coronals in aphasia. In C. Paradis & J.-F. Prunet (Eds.), Phonetics and phonology, vol. 2: The special status of coronals (pp. 201-221). San Diego: Academic Press.
- Béland, R., & Paradis, C. (1997). Principled syllabic dissolution in a primary progressive aphasia case. Aphasiology, *11*, 1171-1196.
- Béland, R., Paradis, C., & Bois, M. (1993). Constraints and repairs in aphasic speech: A group study. Canadian Journal of Linguistics, *38*, 279-302.
- Bell, A. (1971). Some patterns of occurrence and formation of syllable structures. In Working papers on language universal, *6*, pp. 23-137. Palo Alto: Language Universals Project, Stanford University.
- Berg, T. (1991a). Redundant-feature coding in the mental lexicon. Linguistics, *29*, 903-925.

- Berg, T. (1991b). Phonological processing in a syllable-timed language with pre-final stress: Evidence from Spanish speech error data. Language and Cognitive Processes, 6, 265-301.
- Berg, T. (1998). Linguistic structure and change: An explanation from language processing. Oxford: Clarendon Press.
- Best, W. (1996). When racquets are baskets but baskets are biscuits, where do the words come from? A single case study of formal paraphasic errors in aphasia. Cognitive Neuropsychology, 13, 443-480
- Blumstein, S. (1973). A phonological investigation of aphasic speech. The Hague: Mouton.
- Blumstein, S. (1998). Phonological aspects of aphasia. In M. T. Sarno (Ed.) Acquired aphasia (3rd edition). San Diego: Academic Press.
- Boersma, P. (1998). Functional phonology. The Hague: Holland Academic Graphics.
- Boomer, D. S., & Laver, J. D. M. (1968). Slips of the tongue. British Journal of Disorders of Communication, 3, 2-12.
- Boucher, V. (1994). Alphabet-related biases in psycholinguistic inquiries: Considerations for direct theories of speech production and perception. Journal of Phonetics, 22, 1-18.

Breen, G., & Pensalfini, R. (1999). Arrernte: A language with no syllable onsets.

Linguistic Inquiry, 30, 1-26.

Browman, C. P., & Goldstein, L. (1989). Articulatory gestures as phonological units.

Phonology, 6, 201-251.

Browman, C. P., & Goldstein, L. (1992). Articulatory phonology: An overview.

Phonetica, 49, 155-180.

Buckingham, H. W., & Yule, G. (1987). Phonemic false evaluation: Theoretical and

clinical aspects. Clinical Linguistics & Phonetics, 1, 113-125.

Burton-Roberts, N., Carr, P., & Docherty, G. (Eds.). (2000). Phonological knowledge:

Conceptual and empirical issues. Oxford: Oxford University Press.

Butterworth, B. (1992). Disorders of phonological encoding. Cognition, 42, 261-286.

Butterworth, B. & Whittaker, S. (1980). Peggy Babcock's relatives. In G. E. Stelmach

& J. Requin (Eds) Tutorials in motor behavior (pp. 647-656). Amsterdam: North

Holland.

Bybee, J. (2001). Phonology and language use. Cambridge: Cambridge University

Press.

Byrd, D. (1996). A phase window framework for articulatory timing. Phonology, 13,

139-169.

- Camden, C. T. (1980). Word recognition processes: Evidence from laboratory-induced slips of the tongue. Doctoral dissertation, Ohio State University.
- Camden, C. T., Motley, M. T., & Baars, B. J. (1982). Cognitive encoding processes: Evidence for a graphemically-based short-term memory. Human Communication Research, 8, 327-337.
- Caplan, D. (1987). Phonological representations in word production. In E. Keller & M. Gopnik (Eds.), Motor and sensory processes of language (pp. 111-124). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Carter, A., Gerken, L., & Holland, A. (1998, November). Markedness and syllable structure in aphasic word production. Paper presented at the Academy of Aphasia annual meeting, Santa Fe, NM. Abstract in Brain and Language, 65, 202-203.
- Carter, R. E., & Bradshaw, J. L. (1984). Producing “spoonerisms” on demand: Lexical, phonological and orthographic factors in a new experimental paradigm. Speech Communication, 3, 347-360.
- Chambers, K. E., Onishi, K. H., Fisher, C. (2002, June). Generalizing phonotactic regularities from brief auditory experience. Poster presented at the eighth Conference on Laboratory Phonology, New Haven, CT.

- Chang, S. S., Plauché, M. C., & Ohala, J. J. (2001). Markedness and consonant confusion asymmetries. In E. Hume & K. Johnson (Eds.) The role of speech perception in phonology (pp. 79-121). San Diego: Academic Press.
- Cholin, J., Schiller, N. O., & Levelt, W. J. M. (2002, June). The preparation of syllables in speech production. Poster presented at the eighth Conference on Laboratory Phonology, New Haven, CT.
- Chomsky, N., & Halle, M. (1968). The sound pattern of English. New York: Harper & Row.
- Clements, G. N. (1990). The role of the sonority cycle in core syllabification. In M. Beckman & J. Kingston (Eds.), Papers in laboratory phonology I: Between the grammar and physics of speech (pp. 283-333). Cambridge: Cambridge University Press.
- Code, C., & Ball, M. J. (1994). Syllabification in aphasic recurring utterances: Contributions of sonority theory. Journal of Neurolinguistics, 8, 257-265.
- Coleman, J. S. (1998). Cognitive reality and the phonological lexicon: A review. Journal of Neurolinguistics, 11, 295-320.
- Coleman, J. S., & Pierrehumbert, J. (1997). Stochastic phonological grammars and acceptability. In Computational phonology: Third meeting of the ACL special

- interest group in computational phonology (pp. 49-56). Somerset, NJ: Association for Computational Linguistics.
- Coltheart, M. (1999). Modularity and cognition. Trends in Cognitive Sciences, 3, 115-120.
- Costa, A., & Sebastian-Gallés, N. (1998). Abstract phonological structure in language production: Evidence from Spanish. Journal of Experimental Psychology: Learning, Memory, and Cognition, 24, 886-903.
- Cummins, R. (1986). Inexplicit information. In M. Brand & R. M. Harnish (Eds.), The representation of knowledge and belief (pp. 116-126). Tuscon, AZ: University of Arizona Press.
- Cummins, R. (1991). Meaning and mental representation. Cambridge, MA: MIT Press.
- Cutler, A. (1982). The reliability of speech error data. Linguistics, 19, 561-582.
- Dell, G. S. (1986). A spreading activation theory of retrieval in sentence production. Psychological Review, 93, 283-321.
- Dell, G. S. (1988). The retrieval of phonological forms in production: Tests of predictions from a connectionist model. Journal of Memory and Language, 27, 124-142.
- Dell, G. S. (1990). Effects of frequency and vocabulary type on phonological speech errors. Language and Cognitive Processes, 5, 313-349.

- Dell, G. S., Burger, L. K., & Svec, W. R. (1997). Language production and serial order: A functional analysis and a model. Psychological Review, *104*, 123-147.
- Dell, G. S., Juliano, C., & Govindjee, A. (1993). Structure and content in language production: A theory of frame constraints in phonological speech errors. Cognitive Science, *17*, 149-195.
- Dell, G. S., Reed, K. D., Adams, D. R., & Meyer, A. S. (2000). Speech errors, phonotactic constraints, and implicit learning: A study of the role of experience in language production. Journal of Experimental Psychology: Learning, Memory, and Cognition, *26*, 1355-1367.
- Dell, G. S., & Reich, P. A. (1981). Stages in sentence production: An analysis of speech error data. Journal of Verbal Learning and Verbal Behavior, *20*, 611-629.
- Dell, G. S., & Repka, R. J. (1992). Errors in inner speech. In B. J. Baars (Ed). Experimental slips and human error: Exploring the architecture of volition (pp. 237-262). New York: Plenum.
- del Viso, S., Igoa, J. M., & García-Albea, J. E. (1991). On the autonomy of phonological encoding: Evidence from slips of the tongue in Spanish. Journal of Psycholinguistic Research, *20*, 161-185.
- den Ouden, D.-B. (2002). Phonology in aphasia: Syllables and segments in level-specific deficits. Doctoral dissertation, Rijksuniversiteit Groningen, NL.

- Dogil, G., & Mayer, J. (1998). Selective phonological impairment: A case of apraxia of speech. Phonology, *15*, 143-188.
- Elman, J. L. (1990). Finding structure in time. Cognitive Science, *14*, 213-252.
- Favreau, Y., Nespoulous, J.-L., & Lecours, A. R. (1990). Syllable structure and lexical frequency effects in the phonemic errors of four aphasics. Journal of Neurolinguistics, *5*, 165-187.
- Ferber, R. (1991). Slip of the tongue or slip of the ear? On the perception and transcription of naturalistic slips of the tongue. Journal of Psycholinguistic Research, *20*, 105-122.
- Ferber, R. (1995). On the reliability and validity of slip-of-the-tongue corpora: A methodological note. Linguistics, *33*, 1169-1190
- Ferrand, L., & Segui, J. (1998). The syllable's role in speech production: Are syllables chunks, schemas, or both? Psychonomic Bulletin and Review, *5*, 253-298.
- Ferrand, L., Segui, J., & Grainger, J. (1996). Masked priming of word and picture naming: The role of syllabic units. Journal of Memory and Language, *35*, 708-723.
- Ferrand, L., Segui, J., & Humphreys, G. W. (1997). The syllable's role in word naming. Memory and Cognition, *25*, 458-470.
- Fodor, J. A. (1983). The modularity of mind. Cambridge, MA: MIT Press.

- Fowler, C. A. (1987). Consonant-vowel cohesiveness in speech production as revealed by initial and final consonant exchanges. Speech Communication, 6, 231-244.
- Frisch, S. (1996). Similarity and frequency in phonology. Doctoral dissertation, Northwestern University. ROA-198-0597, Rutgers Optimality Archive, <http://roa.rutgers.edu/>.
- Frisch, S. (1997). Against underspecification in speech errors. Studies in the Linguistic Sciences, 27, 79-97.
- Frisch, S. (2000). Temporally organized lexical representations as phonological units. In M. B. Broe & J. B. Pierrehumbert (Eds.) Papers in laboratory phonology V: Acquisition and the lexicon (pp. 283-298). Cambridge: Cambridge University Press.
- Frisch, S., Large, N. R., & Pisoni, D. B. (2000). Perception of wordlikeness: Effects of segment probability and length on the processing of nonwords. Journal of Memory and Language, 42, 481-496.
- Frisch, S. A., & Wright, R. (2002). The phonetics of phonological speech errors: An acoustic analysis of slips of the tongue. Journal of Phonetics, 30, 139-162.
- Frisch, S. A., & Zawaydeh, B. A. (2001). The psychological reality of OCP-Place in Arabic. Language, 77, 91-106.
- Fromkin, V. A. (1971). The non-anomalous nature of anomalous utterances. Language, 47, 27-52.

- Gagnon, D. A., Schwartz, M. F., Martin, N., Dell, G. S., & Saffran, E. M. (1997). The origins of formal paraphasias in aphasics' picture naming. Brain and Language, 59, 450-472.
- García-Albea, J. E., del Viso, S., & Igoa, J. M. (1989). Movement errors and levels of processing in sentence production. Journal of Psycholinguistic Research, 18, 145-161.
- Garrett, M. F. (1975). The analysis of sentence production. In G. H. Bower (Ed.), The psychology of learning and motivation: Advances in research and theory (vol. 9, pp. 133-177). New York: Academic Press.
- Garrett, M. F. (1976). Syntactic processes in sentence production. In R. J. Wales & E. Walker (Eds.), New approaches to language mechanisms (pp. 231-255). Amsterdam: North-Holland.
- Garrett, M. F. (1982). Production of speech: Observations from normal and pathological language use. In A. Ellis (Ed.), Normality and pathology in cognitive functions (pp. 19-76). New York: Academic Press.
- Garrett, M. F. (1984). The organization of processing structure for language production: Applications to aphasic speech. In D. Caplan, A. R. Lecours & A. Smith (Eds.) Biological perspectives on language (pp.172-193). Cambridge, MA: MIT Press.

- Gamkrelidze, T. V. (1978). On the correlation of stops and fricatives in a phonological system. In J. H. Greenberg (Ed.) Universals of human language (Vol. II, pp. 9-46).
- Goldrick, M., Rapp, B., & Smolensky, P. (1999, October). Lexical and postlexical processes in spoken word production. Poster presented at the Academy of Aphasia annual meeting, Venice, Italy. Abstract in Brain and Language, *69*, 367-370.
- Goldstein, L. (1980). Bias and asymmetry in speech perception. In V. A. Fromkin (Ed.) Errors in linguistic performance: Slips of the tongue, ear, pen, and hand (pp. 241-261). New York: Academic Press.
- Goldsmith, J. A. (1990). Autosegmental and metrical phonology. Oxford: Basil Blackwell.
- Greenberg, J. H. (1966). Language universals. The Hague: Mouton.
- Greenberg, J. H. (1978a). Typology and cross-linguistic generalizations. In J. H. Greenberg (Ed.) Universals of human language (Vol. I, pp. 33-59).
- Gupta, P., & Dell, G. S. (1999). The emergence of language from serial order and procedural memory. In B. MacWhinney (Ed.) Emergentist approaches to languages (pp. 447-481). Mahwah, NJ: Erlbaum.
- Hanley, J. R., Kay, J., & Edwards, M. (2002). Imageability effects, phonological errors, and the relationship between auditory repetition and picture naming: Implications for models of auditory repetition. Cognitive Neuropsychology, *19*, 193-206.

- Hanlon, R. E., & Edmondson, J. A. (1996). Disconnected phonology: A linguistic analysis of phonemic jargon aphasia. Brain and Language, 55, 199-212.
- Hartsuiker, R. J. (2002). The addition bias in Dutch and Spanish phonological speech errors: The role of structural context. Language and Cognitive Processes, 17, 61-96.
- Harley, T. A. (1984). A critique of top-down independent levels models of speech production: Evidence from non-plan-internal speech errors. Cognitive Science, 8, 191-219.
- Harris, H. D. (2002). Holographic reduced representations for oscillator recall: A model of phonological production. In W.D. Gray & C. D. Schunn (Eds.), Proceedings of the 24th annual meeting of the Cognitive Science Society (pp. 423-428). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Harris, J., & Lindsey, G. (2000). Vowel patterns in mind and sound. In N. Burton-Roberts, P. Carr, & G. Docherty (Eds.). (2000). Phonological knowledge: Conceptual and empirical issues (pp. 185-205). Oxford: Oxford University Press.
- Hartley, T., & Houghton, G. (1996). A linguistically constrained model of short-term memory for nonwords. Journal of Memory and Language, 35, 1-31.
- Hatfield, F. M., & Walton, K. (1975). Phonological patterns in a case of aphasia. Language and Speech, 18, 341-357.

- Hay, J., Pierrehumbert, J., Beckman, M., & West, J. F. (1999, January). Lexical frequency effects in speech errors. Paper presented at the Linguistic Society of America Annual Meeting, Los Angeles.
- Hayes, B. (1999). Phonetically-driven phonology: The role of Optimality Theory and inductive grounding. In M. Darnell, E. Moravcsik, M. Noonan, F. Newmeyer & K. Wheatly (Eds.), Functionalism and formalism in linguistics, volume I: General papers (pp. 243-285). Amsterdam: John Benjamins.
- Hendriks, H., & McQueen, J., eds. (1996). Annual report 1995. Nijmegen: Max Planck Institut für Psycholinguistik.
- Hillis, A. E. (2001). The organization of the lexical system. In B. Rapp (Ed.) The handbook of cognitive neuropsychology: What deficits reveal about the human mind (pp. 185-210). Philadelphia: Psychology Press.
- Hockett, C. F. (1967). Where the tongue slips, there slip I. In To honor Roman Jakobson: Essays on the occasion of his seventieth birthday (vol. II, pp. 910-936). Paris: Mouton.
- Hopfield, J. J. (1982). Neural networks and physical systems with emergent collective computational abilities. Proceedings of the National Academy of Sciences, 79, 2554-2558.

- Hopfield, J. J. (1984). Neurons with graded response have collective computational properties like those of two-state neurons. Proceedings of the National Academy of Sciences, 81, 3088-3092.
- Humphreys, K. R. (2002). Lexical bias in speech errors. Doctoral dissertation, University of Illinois at Urbana-Champaign.
- Jakobson, R. (1939/1984). Zero sign. (L. R. Waugh., Trans.) in L. R. Waugh & M. Halle (Eds.) Russian and Slavic grammar (pp. 151-159). Berlin: Mouton.
- Jakobson, R. (1941/1968). Child language aphasia and phonological universals. (A. R. Keiler, Trans.) The Hague: Mouton.
- Joanisse, M. F. (2000). Connectionist phonology. Doctoral dissertation, University of Southern California, Los Angeles, CA.
- Johnson, G. (1996). Fire in the mind: Science, faith, and the search for order. New York: Vintage.
- Jordan, M. I. (1986). Serial order: A parallel distributed processing approach. Institute for Cognitive Science Technical Report 8604. La Jolla, CA: University of California at San Diego.
- Jurafsky, D., Bell, A., & Girand, C. (in press). The role of the lemma in form variation. In C. Gussenhoven & N. Warner (Eds.) Papers in Laboratory Phonology VII. Berlin: Mouton.

- Kean, M.-L. (1975/1980). The theory of markedness in generative grammar. Doctoral dissertation, Massachusetts Institute of Technology. Bloomington, IN: Indiana University Linguistics Club.
- Kenstowicz, M. (1994). Phonology in generative grammar. Cambridge: Blackwell.
- Kessler, B., & Treiman, R. (1997). Syllable structure and the distribution of phonemes in English syllables. Journal of Memory and Language, *37*, 295-311.
- Kiparsky, P. (1985). Some consequences of Lexical Phonology. Phonology Yearbook, *2*, 83-138.
- Kohn, S. E., Melvold, J., & Smith, K. L. (1995). Consonant harmony as a compensatory mechanism in fluent aphasic speech. Cortex, *31*, 747-756.
- Kohn, S. E., & Smith, K. L. (1994). Distinctions between two phonological output deficits. Applied Psycholinguistics, *15*, 75-95.
- Kucera, H., & Francis, W. (1967). Computational analysis of present-day American English. Providence, RI: Brown University Press.
- Kupin, J. J. (1982). Tongue twisters as a source of information about speech production. Bloomington: Indiana University Linguistics Club.
- Landauer, T. K., & Streeter, L. A. (1973). Structural differences between common and rare words: Failure of equivalence assumptions for theories of word recognition. Journal of Verbal Learning and Verbal Behavior, *12*, 119-131.

- Lashley, K. S. (1951). The problem of serial order in behavior. In L. A. Jeffress (Ed.) Cerebral mechanisms in behavior: The Hixon symposium (pp. 112-136). New York: John Wiley & Sons.
- Laver, J. (1980). Slips of the tongue as neuromuscular evidence for a model of speech production. In H. W. Dechert & M. Raupach (Eds.) Temporal variables in speech: Studies in honour of Frieda Goldman-Eisler (pp.21-26). The Hauge: Mouton.
- Levelt, W. J. M. (1989). Speaking: From intention to articulation. Cambridge, MA: MIT Press.
- Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. Behavioral and Brain Sciences, *22*, 1-75.
- Levelt, W. J. M., & Wheeldon, L. (1994). Do speakers have access to a mental syllabary? Cognition, *50*, 239-269.
- Levitt, A. G., & Healy, A. F. (1985). The roles of phoneme frequency, similarity, and availability in the experimental elicitation of speech errors. Journal of Memory and Language, *24*, 717-733.
- Lindblom, B. (2000). Developmental origins of adult phonology: The interplay between phonetic emergents and the evolutionary adaptations of sound patterns. Phonetica, *57*, 297-314.

- Lombardi, L. (1995). Laryngeal features and privativity. The Linguistic Review, 12, 35-59.
- Lombardi, L. (1999). Positional faithfulness and voicing assimilation in optimality theory. Natural Language and Linguistic Theory, 17, 267-302.
- Luce, P. A., Goldinger, S. R., Auer Jr., E. T., & Vitevitch, M. S. (2000). Phonetic priming, neighborhood activation, and PARSYN. Perception & Psychophysics, 62, 615-625.
- MacKay, D. G. (1969). Forward and backward masking in motor systems. Kybernetik, 6, 57-64.
- MacKay, D. G. (1970). Spoonerisms: The structure of errors in the serial order of speech. Neuropsychologia, 8, 323-350.
- MacKay, D. G. (1972). The structure of words and syllables: Evidence from errors in speech. Cognitive Psychology, 3, 210-227.
- MacKay, D. G. (1973). Complexity in output systems: Evidence from behavioral hybrids. American Journal of Psychology, 86, 785-806.
- MacKay, D. G. (1978). Speech errors inside the syllable. In A. Bell & J. B. Hooper (Eds.) Syllables and segments (pp. 201-212). Amsterdam: North-Holland Publishing.
- MacKay, D. G. (1987). The organization of perception and action: A theory for language and other cognitive skills. New York: Springer-Verlag.

- MacKay, D. G., & MacDonald, M. C. (1984). Stuttering as a sequencing and timing disorder. In R. F. Curlee, & W. H. Perkins (Eds.) Nature and treatment of stuttering: New directions (pp. 261-282). Boston: Allyn & Bacon.
- MacNeilage, P. F. (1981). Feedback in speech production: An ecological perspective. In T. Myers, J. Laver, & J. Anderson (Eds.) The cognitive representation of speech (pp. 39-44). Amsterdam: North-Holland.
- Maddieson, I. (1984). Patterns of sounds. Cambridge: Cambridge University Press.
- Martin, N., & Saffran, E. M. (2002). The relationship of input and output phonological processing: An evaluation of models and evidence to support them. Aphasiology, *16*, 107-150.
- Meijer, P. J. A. (1994). Towards a new model of phonological encoding. In A. Ram & K. Eiselt (Eds.) Proceedings of the 16th annual conference of the Cognitive Science Society (pp.619-623). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Meijer, P. J. A. (1996). Suprasegmental structures in phonological encoding: The CV structure. Journal of Memory and Language, *35*, 840-853.
- Meyer, D. E., & Gordon, P. C. (1985). Speech production: Motor programming of phonetic features. Journal of Memory and Language, *24*, 3-26.
- Motley, M. T. (1973). An analysis of spoonerisms as psycholinguistic phenomena. Speech Monographs, *40*, 66-71.

- Motley, M. T. (1986). On replicating the SLIP technique: A reply to Sinsabaugh and Fox. Communication Monographs, 53, 342-351.
- Motley, M. T., & Baars, B. J. (1975). Encoding sensitivities to phonological markedness and transitional probabilities: Evidence from spoonerisms. Human Communication Research, 1, 353-361.
- Motley, M. T., & Baars, B. J. (1976). Laboratory induction of verbal slips: A new method for psycholinguistic research. Communication Quarterly, 24, 28-34.
- Mowrey, R. A., & MacKay, I. R. A. (1990). Phonological primitives: Electromyographic speech error evidence. Journal of the Acoustical Society of America, 88, 1299-1312.
- Munson, B. (2001). Phonological pattern frequency and speech production in adults and children. Journal of Speech, Language, and Hearing Research, 44, 778-792.
- Nespoulous, J.-L., Joanette, Y., Béland, R., Caplan, D., & Lecours, A. R. (1984). Phonological disturbances in aphasia: Is there a “markedness effect” in aphasic phonetic errors? in F. C. Rose (Ed.), Advances in neurology, vol. 42: progress in aphasiology (pp. 203-214). New York: Raven Press.
- Nespoulous, J.-L., Joanette, Y., Ska, B., Caplan, D., & Lecours, A. R. (1987). Production deficits in Broca’s and conduction aphasia: Repetition versus reading. In E. Keller & M. Gopnik (Eds.) Motor and sensory processes of language (pp. 53-81). Hillsdale, NJ: Lawrence Erlbaum Associates.

- Nespoulous, J.-L., & Moreau, N. (1997). Repair strategies and consonantal cluster production in Broca's aphasia. In Y. Lebrun (Ed.) From the brain to the mouth: Acquired dysarthria and dysfluency in adults (pp.71- 80.). Dordrecht: Kluwer.
- Nespoulous, J.-L., & Moreau, N. (1998). "Repair strategies" and the production of segmental errors in aphasia: Epenthesis vs. syncope in consonantal clusters. In E. G. Visch-Brink & R. Bastiaanse (Eds.) Linguistic levels in aphasiology (pp. 133-145). San Diego: Singular.
- Nickels, L., & Howard, D. (1995). Phonological errors in aphasic naming: Comprehension, monitoring, and lexicality. Cortex, 31, 209-237.
- Niemi, J., & Koivuselkä-Sallinen, P. (1985). Phoneme errors in Broca's aphasia: Three Finnish cases. Brain and Language, 26, 28-48.
- Nooteboom, S. G. (1969). The tongue slips into patterns. In A. G. Sciarone, A. J. van Essen, & A. A. van Raad (Eds.) Leyden studies in linguistics and phonetics (pp. 114-132). The Hague: Mouton. [Reprinted in V. A. Fromkin (Ed.) (1973) Speech errors as linguistic evidence (pp. 144-156). The Hague: Mouton.]
- Ohala, J. J. (1983). The origin of sound patterns in vocal tract constraints. In P. MacNeilage (Ed.) The production of speech (pp. 189-216). New York: Springer-Verlag.

- Onishi, K. H., Chambers, K. E., & Fisher, C. (2002). Learning phonotactic constraints from brief auditory experience. Cognition, 83, B13-B23.
- Palmer, S. E. (1978). Fundamental aspects of cognitive representation. In E. Rosch, & B. Lloyd (Eds.) Cognition and categorization (pp. 259-303). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Palmer, S. E., & Kimchi, R. (1986). The information processing approach to cognition. In T. J. Knapp & L. C. Robertson (Eds.), Approaches to cognition: Contrasts and controversies (pp. 37-77). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Paradis, C. (1988). On constraints and repair strategies. The Linguistic Review, 6, 71-97.
- Paradis, C., & Prunet, J.-F. (1991a). Introduction: Asymmetry and visibility in consonant articulations. In C. Paradis & J.-F. Prunet (Eds.) Phonetics and phonology, vol. 2: The special status of coronals (pp. 1-28). San Diego: Academic Press.
- Paradis, C., & Prunet, J.-F., eds. (1991b). Phonetics and phonology, vol. 2: The special status of coronals. San Diego: Academic Press.
- Plate, T. (1995). Holographic reduced representations. IEEE Transactions on Neural Networks, 6, 623-641.
- Plaut, D. C., & Kello, C. T. (1999). The emergence of phonology from the interplay of speech perception and comprehension: A distributed connectionist approach. In B.

- MacWhinney (Ed.) The emergence of language (pp. 381-415). Mahweh, NJ: Erlbaum.
- Postma, A., & Noordanus, C. (1996). Production and detection of speech errors in silent, mouthed, noise-masked, and normal auditory feedback speech. Language and Speech, 39, 375-392.
- Pouplier, M., Chen, L., Goldstein, L., & Byrd D. (2000, June). Articulatory evidence of gradient speech errors. Poster presented at the Seventh Laboratory Phonology meeting.
- Pouplier, M., & Goldstein, L. (2002). Asymmetries in speech errors: Production, perception and the question of underspecification. Manuscript submitted for publication.
- Prince, A., & Smolensky, P. (1993). Optimality theory: Constraint interaction in generative grammar. Technical report TR-2, Rutgers Center for Cognitive Science, Rutgers University, New Brunswick, NJ. Technical report CU-CS-696-93, Department of Computer Science, University of Colorado, Boulder. Revised version, 2002: ROA-537-0802, Rutgers Optimality Archive, <http://roa.rutgers.edu/>.
- Rapp, B., Folk, J., & Tainturier, M.-J. (2001). Word reading. In B. Rapp (Ed.) The handbook of cognitive neuropsychology: What deficits reveal about the human mind (pp. 233-262). Philadelphia: Psychology Press.

- Robins, F. L. (1980). A study of elicited speech errors in normal speakers. M.A. thesis, University of Minnesota.
- Roelofs, A. (1997). The WEAVER model of word-form encoding in speech production. Cognition, 64, 249-284.
- Roelofs, A. (1999). Phonological segments and features as planning units in speech production. Language and Cognitive Processes, 14, 173-200.
- Roelofs, A., & Meyer, A. S. (1998). Metrical structure in planning the production of spoken words. Journal of Experimental Psychology: Learning, Memory, and Cognition, 24, 922-939.
- Romani, C., & Calabrese, A. (1998). Syllabic constraints in the phonological errors of an aphasic patient. Brain and Language, 64, 83-121.
- Rumelhart, D. E., Hinton, G. E., & McClelland, J. L. (1986). A general framework for parallel distributed processing. In D. E. Rumelhart, J. L. McClelland & the PDP Research Group, Parallel distributed processing: Explorations in the microstructure of cognition (Vol. I, pp. 45-76). Cambridge, MA: MIT Press.
- Sagey, E. (1986). The representation of features and relations in nonlinear phonology. Doctoral dissertation, Massachusetts Institute of Technology.
- Saltzman, E. L., & Munhall, K. G. (1989). A dynamical approach to gestural patterning in speech production. Ecological Psychology, 1, 333-382.

- Schiller, N. O. (1998). The effect of visually masked syllable primes on the naming latencies of words and pictures. Journal of Memory and Language, *39*, 484-507.
- Schiller, N. O. (2000). Single word production in English: The role of subsyllabic units during phonological encoding. Journal of Experimental Psychology: Learning, Memory, and Cognition, *26*, 512-528.
- Sevald, C. A., Dell, G. S., & Cole, J. S. (1995). Syllable structure in speech production: Are syllables chunks or schemas? Journal of Memory and Language, *34*, 807-820.
- Shattuck, S. (1975). Speech errors and sentence production. Doctoral dissertation, Massachusetts Institute of Technology.
- Shattuck-Hufnagel, S. (1979). Speech errors as evidence for a serial-ordering mechanism in sentence production. In W. E. Cooper & E. C. T. Walker (Eds.) Sentence processing: Psycholinguistic studies presented to Merrill Garrett (pp. 295-341). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Shattuck-Hufnagel, S. (1983). Sublexical units and suprasegmental structure in speech production planning. In P. F. MacNeilage (Ed.) The production of speech (pp. 109-136). New York: Springer Verlag
- Shattuck-Hufnagel, S. (1987). The role of word-onset consonants in speech production planning: New evidence from speech error patterns. In E. Keller & M. Gopnik (Eds.)

- Motor and sensory processes of language (pp.17-51). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Shattuck-Hufnagel, S. (1992). The role of word structure in segmental serial ordering. Cognition, 42, 213-259.
- Shattuck-Hufnagel, S., & Klatt, D. H. (1979). The limited use of distinctive features and markedness in speech production: Evidence from speech error data. Journal of Verbal Learning and Verbal Behavior, 18, 41-55.
- Shattuck-Hufnagel, S., & Klatt, D. H. (1980). How single phoneme error data rule out two models of error generation. In V. A. Fromkin (Ed.) Errors in linguistic performance: Slips of the tongue, ear, pen, and hand (pp.35-46). New York: Academic Press.
- Sherman, D. (1975). Stop and fricative systems: A discussion of paradigmatic gaps and the question of language sampling. Stanford Working Papers in Language Universals, 17, 1-31.
- Sinsabaugh, B. A., & Fox, R. A. (1986). Reevaluating the SLIP paradigm: A research note. Communication Monographs, 53, 335-341.
- Smolensky, P. (1986). Information processing in dynamical systems: Foundations of harmony theory. In D. E. Rumelhart, J. L. McClelland & the PDP Research Group,

- Parallel distributed processing: Explorations in the microstructure of cognition (Vol. I, pp. 194-281). Cambridge, MA: MIT Press.
- Smolensky, P. (1990). Tensor product variable binding and the representation of symbolic structures in connectionist networks. Artificial Intelligence, 46, 159-216.
- Smolensky, P. (1996). On the comprehension/production dilemma in child language. Linguistic Inquiry, 27, 720-731.
- Söderpalm, E. (1979). Speech errors in normal and pathological speech. Travaux de l'Institute de Linguistique de Lund, 14. Lund: CWK Gleerup.
- Stemberger, J. P. (1982). The nature of segments in the lexicon: Evidence from speech errors. Lingua, 56, 235-259.
- Stemberger, J. P. (1983). Speech errors and theoretical phonology: A review. Unpublished manuscript, Carnegie-Mellon University. Distributed by the Indiana University Linguistics Club, Bloomington, IN.
- Stemberger, J. P. (1985a). An interactive activation model of language production. In A. W. Ellis (Ed.) Progress in the psychology of language (Vol.1, pp. 143-186). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Stemberger, J. P. (1985b). Phonological rule ordering in a model of language production. Unpublished manuscript, Indiana University. Distributed by the Indiana University Linguistics Club, Bloomington, IN.

- Stemberger, J. P. (1990). Wordshape errors in language production. Cognition, 35, 123-157.
- Stemberger, J. P. (1991a). Apparent anti-frequency effects in language production: The addition bias and phonological underspecification. Journal of Memory and Language, 30, 161-185.
- Stemberger, J. P. (1991b). Radical underspecification in language production. Phonology, 8, 73-112.
- Stemberger, J. P., & MacWhinney, B. (1988). Are inflected forms stored in the lexicon? In M. Hammond & M. Noonan (Eds.), Theoretical morphology: Approaches in modern linguistics (pp.101-116). San Diego, CA: Academic Press.
- Stemberger, J. P., & Treiman, R. (1986). The internal structure of word-initial consonant clusters. Journal of Memory and Language, 25, 163-180.
- Steriade, D. (1997). Phonetics in phonology: The case of laryngeal neutralization. Unpublished manuscript, University of California, Los Angeles.
- Tent, J. & Clark, J. E. (1980). An experimental investigation into the perception of slips of the tongue. Journal of Phonetics, 8, 317-325.
- Treiman, R., Kessler, B., Knewasser, S., Tincoff, R., & Bowman, M. (2000). English speakers' sensitivity to phonotactic patterns. In M. B. Broe & J. B. Pierrehumbert

- (Eds.) Papers in laboratory phonology V: Acquisition and the lexicon (pp. 269-282).
Cambridge: Cambridge University Press.
- Trubetzkoy, N.S. (1939/1969). Principles of phonology. (C. A. M. Baltaxe, Trans.)
Berkeley, CA: University of California Press.
- University of California, Los Angeles Phonetics Laboratory (2000). UCLA Phonetic
Segment Inventory Database (UPSID).
- Van den Broecke, M. P. R., & Goldstein, L. (1980). Consonant features in speech errors.
In V. A. Fromkin (Ed.) Errors in linguistic performance: Slips of the tongue, ear, pen,
and hand (pp.47-65). New York: Academic Press.
- Vitevitch, M. S. (2002). The influence of phonological similarity neighborhoods on
speech production. Journal of Experimental Psychology: Learning, Memory, and
Cognition, 28, 735-747.
- Vitevitch, M. S., & Luce, P. A. (1998). When words compete: Levels of processing in
spoken word recognition. Psychological Science, 9, 325-329.
- Vitevitch, M. S., & Luce, P. A. (1999). Probabilistic phonotactics and neighborhood
activation in spoken word recognition. Journal of Memory and Language, 40, 374-
408.

- Vitevitch, M. S., Luce, P. A., Charles-Luce, J., & Kemmerer, D. (1997). Phonotactics and syllable stress: Implications for the processing of spoken nonsense words. Language and Speech, *40*, 47-62.
- Vousden, J. I., Brown, G. D. A., & Harley, T. A. (2000). Serial control of phonology in speech production: A hierarchical model. Cognitive Psychology, *41*, 101-175.
- Wan, I-P., & Jaeger, J. (1998). Speech errors and the representation of tone in Mandarin Chinese. Phonology, *15*, 417-461.
- Wells, R. (1951). Predicting slips of the tongue. The Yale Scientific Magazine, *26*(3), 9-30.
- Wheeldon, L. R., & Levelt, W. J. M. (1995). Monitoring the time course of phonological encoding. Journal of Memory and Language, *34*, 311-334.
- Wheeler, D. W., & Touretzky, D. S. (1997). A parallel licensing model of normal slips and phonemic paraphasias. Brain and Language, *59*, 147-201.
- Wilshire, C. E. (1998). Serial order in phonological encoding: An exploration of the 'word onset effect' using laboratory-induced errors. Cognition, *68*, 143-166.
- Wilshire, C. E. (1999). The "tongue twister" paradigm as a technique for studying phonological encoding. Language and Speech, *42*, 57-82.
- Wilshire, C. E., & McCarthy, R. A. (1996). Experimental investigations of an impairment in phonological encoding. Cognitive Neuropsychology, *13*, 1059-1098.

- Wilshire, C. E., & Nespoulous, J.-L. (1997, October). The role of syllabic frequency in aphasic phonological errors. Poster presented at the Academy of Aphasia annual meeting, Philadelphia, PA. Abstract in Brain and Language, 60, 147-150.
- Yaniv, I., Meyer, D. E., Gordon, P. C., Huff, C. A., & Sevald, C. E. (1990). Vowel similarity, connectionist models, and syllable structure in motor programming of speech. Journal of Memory and Language, 29, 1-26.
- Yip, M. (1991). Coronals, consonant clusters, and the coda condition. In C. Paradis & J.-F. Prunet (Eds.) Phonetics and phonology, vol. 2: The special status of coronals (pp. 61-78). San Diego: Academic Press.
- Zipf, G. K. (1935). The psycho-biology of language: An introduction to dynamic philology. Boston: Houghton Mifflin.

CURRICULUM VITAE

Matthew Andrew Goldrick was born in East Lansing, Michigan, on March 3, 1975, and spent most of his formative years in Hockessin, Delaware. He received a Bachelor of Arts in Cognitive Science with honors from Johns Hopkins University in 1997 and a Master of Arts in Cognitive Science at the same institution in 1999. Following completion of his Ph.D., he will begin a post-doctoral appointment in the Department of Cognitive and Linguistic Sciences at Brown University.