NORTHWESTERN UNIVERSITY

# SIMILARITY AND FREQUENCY IN PHONOLOGY

A DISSERTATION

SUBMITTED TO THE GRADUATE SCHOOL

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

for the degree

DOCTOR OF PHILOSOPHY

Field of Linguistics

By

**STEFAN FRISCH**

EVANSTON, ILLINOIS

December 1996

## ABSTRACT

Similarity and Frequency in Phonology

Stefan Frisch

This thesis focuses upon parallels between phonology and phonological processing. I study phonological speech errors and a phonotactic dissimilarity constraint, demonstrating they have analogous similarity and frequency effects. In addition, I show that abstract phonological constraints are influenced by the phonological encoding of lexical items.

The results of this thesis are based on a metric of similarity computed using the representations of STRUCTURED SPECIFICATION (Broe 1993). This metric is quantitatively superior to traditional metrics of similarity which are based on feature counting. I also employ a probabilistic model of a gradient linguistic constraint which is based on categorical perception. In this model, the acceptability of a form is gradient, and acceptability is correlated with frequency. The most acceptable forms in a language are the most frequent ones. This constraint model provides a better fit to gradient phonotactic data than traditional categorical linguistic constraints. Together, the similarity metric and gradient constraint model demonstrate that statistical patterns in language can be relevant, principled, and formally modeled in linguistic theory.

Using the gradient constraint model, I show that similarity effects in phonotactics are stronger word initially than later in the word. A parallel pattern is experimentally demonstrated for speech errors. I claim that the effect for speech errors follows from the fact that production of segmental material in a lexical item is inherently temporal. I argue that segmental information in lexical representations is sequentially accessed even for abstract phonological purposes, like phonotactics. The effects of word position on similarity in both speech production and phonotactics are accounted for in a connectionist model of lexical access, which does not differentiate the storage of a representation from its use.

Structured specification is incompatible with UNDERSPECIFICATION (Kiparsky 1982, Archangeli 1984). In underspecification, features are left blank in a linguistic representation to capture redundancy relationships and phonological markedness. I demonstrate that models of similarity in phonotactics and speech errors which use underspecification do not model the data as well as the similarity metric based on structured specification.

ACKNOWLEDGMENTS

Like the subject about which it is written, this thesis is not a discrete entity, with a discrete author. To a greater or lesser degree, it consists of many papers I've read or seen presented, and is coauthored by the many people who have discussed the material with me. In particular, the participants at the 5th conference on Laboratory Phonology at Northwestern University provided inspiration and new ideas just as I was preparing to defend. This is my opportunity to thank you all for bringing me to where I am today.

I would like to give explicit thanks to Janet Pierrehumbert for being an excellent mentor while simultaneously treating me as an equal. My debt to her is tremendous, and can only be repaid to my future students by reminding them to "pick up your own dirty socks." Michael Broe has been a friend, a colleague, and an advisor. I would not have survived without our regular squash matches or his help in the final weeks in figuring out what I was "getting on about." Doug Medin introduced me to cognitive science, and served as guide through the literature on similarity and categorization. Gary Dell gave me a wealth of encouragement and some of the best discussions I have ever had. His input has radically influenced this thesis, and my own perspective on phonology.

Special thanks go to Stefanie Shattuck-Hufnagel for her encouragement and ideas while I was working out my own speech error experiments. She was also kind enough to share a portion of her natural speech error corpus with me, and is therefore doubly thanked.

I would also like to thank Kyle Johnson for awakening my interest in linguistics, and Sufian Husseini (in the Mathematics Department at the University of Wisconsin at Madison) for encouraging me to pursue it. I also owe very much to Beatrice Santorini, who got me started in quantitative modeling and supervised and encouraged my early professional development.

Thanks to Sarmad Hussein, Lewis Gebhardt, Bob Knippen, Lisa Lane, Shahrzad Mahootian, Rami Nair, Dan Silverman, Grace Song, Monica Wright, Richard Wright, and especially to Sue Triforo. Detailing their personal and professional contributions would be a thesis in itself. You all helped me to keep my sanity.

I owe a lot to my family, Nancy Frisch, Ronald Frisch, Eric Frisch, Cherie Goggin, and John Goggin, for supporting and encouraging my seemingly endless education.

Finally, I would like to thank Katherine for reminding me to "get going." Every day. Twice.

This work was supported by NSF Grant No. BNS 9022484 to Northwestern University and by a Northwestern University Dissertation Year Grant.

CONTENTS

# CHAPTER 1
# Linguistic and Cognitive Theory

Linguistics is the scientific study of language as a system. The phonology, or sound structure, of a language is one component of this system. Standard phonological theories treat linguistic knowledge as a symbolic grammar (Chomsky & Halle 1968, Goldsmith 1979, Scobbie 1993, Prince & Smolensky 1993). Regularities within and between languages are described using either symbolic strings subjected to rewrite rules, or structural constraints on formal representations comprised of discrete linguistic elements. Traditionally, a symbol based formalism is the only possible linguistic formalism. I argue that a discrete symbol system cannot capture significant phonotactic regularities which are traditionally a component of phonological theory. These regularities, originally subsumed as part of the Obligatory Contour Principle (OCP) applied to place of articulation tiers (McCarthy 1986, 1988, 1994), are in fact based on similarity (Pierrehumbert 1993), an inherently gradient property. The proper formulation of this constraint requires a fuzzy or probabilistic phonology capable of modeling continuous data. In this thesis, I adopt the STOCHASTIC CONSTRAINT MODEL, originally proposed in Frisch, Broe, & Pierrehumbert (1995). In addition, I develop a model of gradient constraint combination which is equivalent to fuzzy logic set intersection. Thus, I implement a phonological model which contains both gradient constraints and gradient constraint combination.

The necessity for reanalyzing OCP-Place as a phonotactic constraint based on similarity supports a model of linguistic competence where linguistic knowledge are grounded in the general cognitive functions. In cognitive psychology and psycholinguistics, similarity is used as an explanation for categorization (see Goldstone 1994a), reasoning by analogy (Gentner 1989), and confusability in processing (e.g. in speech errors and perceptual confusability). I adopt the model of similarity developed in Frisch, Broe, & Pierrehumbert (1995). Similarity is computed based on the representation of segments in STRUCTURED SPECIFICATION (Broe 1993). Structured specification is a hierarchical representation of feature dependencies which structurally encodes the distinctiveness or redundancy of features. Structured specification explicitly encodes the natural classes of a segment inventory. The Frisch, Broe, & Pierrehumbert (1995) similarity model computes similarity as function of shared and non-shared natural classes. This model is an extension of the contrast model of similarity (Tversky 1977, Tversky & Gati 1982) that also incorporates a synergistic effect of feature matching on similarity (Hayes-Roth & Hayes-Roth 1977, Gluck & Bower 1988, Goldstone 1994b).

The representations in structured specification encode, rather than eliminate, redundancy. Thus structured specification is incompatible with UNDERSPECIFICATION THEORY (Kiparsky 1982, Archangeli 1984). In underspecification theory, underspecified features are left out of the representation, and thus they have no effect on similarity (Stemberger 1991b). In structured specification, traditionally underspecified features have a reduced effect on similarity, due to the natural class structure. I demonstrate that the model of similarity using structured specification is empirically superior to a similarity model using underspecification theory.

In modeling gradient linguistic constraints using the stochastic constraint model of Frisch, Broe, & Pierrehumbert (1995), frequency plays a crucial role. In this model, the goodness of a form is reflected directly in its frequency of occurrence: the best forms are the most frequent; the worst forms are infrequent or non-occurring. Thus, this model makes a second connection with general cognition by incorporating frequency effects within the phonology. Given the existence of frequency effects in all other human cognitive functions this is clearly a natural and necessary step (Pierrehumbert 1996, cf. Bod 1995).

A model of phonology which admits gradient generalizations grounded in continuous variables like similarity and frequency is computationally very different from the traditional symbolic approach. Consequently, the implementation of such a phonology in a natural language system requires very different mechanisms from those of formal language theory. Similarity and frequency are intrinsic aspects of connectionist models, which are inherently gradient and quantitative in nature. Connectionism is a plausible basis for the computational implementation of gradient constraints and gradient constraint satisfaction (Dell 1996). Thus, contra Pinker & Prince (1988), connectionism may be preferable to a discrete symbol system for implementing phonological theory.

In addition to being gradient and quantitative, connectionism differs from standard symbolic approaches in incorporating the notion of time within the model. Connectionist models have an inherent temporal component, as processing occurs via the spreading of activation between nodes in the system. Dell (1986) and Sevald & Dell (1994) employ a connectionist model of phonological encoding to account for temporal effects in speech production. I adopt this model and present evidence that temporal effects of phonological encoding are observed not only in phonological processing, but also in the phonotactic constraint, OCP-Place. I thus make a third connection between phonology and general cognition by showing that the process of phonological encoding impacts the abstract phonology.

Parallel to the case of phonological universals based on articulatory or auditory constraints (Lindblom 1983, 1990; Browman & Goldstein 1986, 1990; Sagey 1986; Silverman 1995; Flemming 1995), I propose that language may be constrained by cognitive universals. Further, the influence of cognitive processing constraints over time may result in the grammaticization of soft cognitive constraints into language particular constraints. Language particular patterns may thus be grounded in cognitive processes much like, for example, assimilation is grounded in soft physical constraints like articulator overlap and minimization of effort (Lindblom 1983, Browman & Goldstein 1986).

### *1.1 Rules versus Constraints in Linguistic Theory*

Formal linguistic theory has undergone a major paradigm shift since the early conception of a generative grammar (e.g. Chomsky 1957). The relatively recent adoption of Unification-based Grammars (Gazdar, Klein, Pullum, & Sag 1985; Scobbie 1993; Pollard & Sag 1994), the Minimalist program (Chomsky 1993), and Optimality Theory (Prince & Smolensky 1993) marks a shift from derivational, rule-based approaches to non-derivational constraints. While these frameworks contain residual pockets of derivationalism to different degrees, their common goal is to maximize the relevance of the shape of the output in determining well-formedness. This is a necessary first step in incorporating gradient phenomena within linguistic theory.

### *1.1.1 Gradient constraints.*

A major goal of this thesis is to implement a model of a gradient linguistic constraint and gradient constraint combination in a rigorous manner. This thesis contributes to a growing body of evidence which indicates that statistical patterns in the lexicon are relevant linguistic generalizations (Greenberg 1950; Davis 1991; Pierrehumbert 1993; Berkley 1994a, b; Buckley 1995) to which native speakers are sensitive (McCarthy 1986; Pierrehumbert 1994; Beckman & Edwards 1996; Kessler & Treiman 1996; Treiman, Kessler, Knewasser, & Tinkoff 1996). Phonological theory should be able to model and explain this knowledge.

The model of a gradient linguistic constraint, called the stochastic constraint model (Frisch, Broe, & Pierrehumbert 1995), is a fuzzy logic version of a categorical constraint. In a categorical constraint, a form either violates the constraint or it does not; it is acceptable or it is not. A fuzzy logic constraint can be violated to a degree. I assume, following Frisch, Broe, & Pierrehumbert (1995), that the degree of acceptability of a form is correlated with its relative frequency. Thus, the acceptability of a form is dependent on how frequent that form is in comparison to the other forms in the language.

Traditional derivational phonology is wholly concerned with the mapping from underlying form to surface form. The presence, absence, and distribution of other forms in the language is irrelevant. Statistical patterns in the lexicon could be modeled in such a system by applying constraints probabilistically. If a constraint applied only to a certain percentage of underlying forms, then only a certain percentage of surface forms would obey the constraint. However, this approach assumes that whether a probabilistic constraint applies or does not, the resulting output is equally valid. In the stochastic constraint model, infrequent forms are less acceptable than frequent ones, implying that infrequent forms are poorer examples of the words of a language than frequent forms. Experimental evidence supports the stochastic constraint model: native speakers rate words containing infrequent phonotactic patterns as less word-like than words containing frequent phonotactic patterns (Pierrehumbert 1994, Kessler & Treiman 1996, Treiman et al. 1996).

Basing gradient acceptability on frequency is a fuzzy logic extension of the implicit assumption that phonotactic patterns which never appear are unacceptable.[1] In standard linguistic analyses, language learners are assumed to be able to conclude that forms which would plausibly exist, but appear to be absent from their linguistic input, are unacceptable. The fuzzy logic extension assumes that language learners are able to conclude that forms which rarely appear, or which appear much less frequently than expected, are poor instances of words and are not representative of the words of their language.

[1] A related challenge to categorical models of language is the problem of principled versus accidental gaps. If statistical patterns are admitted as relevant, then a solution presents itself. A gap is principled if the absence of forms of a particular type is statistically significant, when the lexicon is considered as a sample of the combinatorial possibilities available in the language.

### *1.1.2 Underspecification.*

Underspecification refers to the practice of leaving blanks in the representation of a segment to indicate the default or redundant status of certain features (Kiparsky 1982, Archangeli 1984, Steriade 1987). Traditionally, underspecified features are left blank in underlying representations and become filled in during the course of the phonological derivation. Thus, underspecification is crucially dependent on the existence of abstract underlying forms and a procedure where these forms are mapped into fully specified surface forms. Underspecification is used to account for the transparency of certain segments containing default or redundant features to phonological processes. Broe (1993) raises a number of logical objections to underspecification theory. He offers structured specification as an alternative, and proposes that the redundant or default status of features instead be encoded explicitly in a feature dependency hierarchy and a markedness hierarchy. In order to capture transparency effects, phonological constraints are sensitive to the redundancy and markedness relations in the hierarchies.

Stemberger (1991a, b) utilizes underspecification theory in the analysis of certain patterns in both experimentally elicited and naturally occurring phonological speech errors. Stemberger's analysis provides indirect evidence against current constraint based models of phonology. Stemberger (1991b) demonstrates that underspecified features influence error rate, but have less of an effect than specified features. He claims that, since underspecified features are not present underlyingly, they do not affect similarity at early stages in the derivation. At later stages, all features are specified, and both specified and underspecified features influence error rate. This model of phonological speech errors can be compared directly with one based on similarity using structured specification (Frisch, Broe, & Pierrehumbert 1995). I show that the structured specification model provides a far better description of the error data than underspecification theory. In structured specification, default and redundant features have a lesser effect on similarity due to the explicit use of natural classes in the computation of similarity. Thus, structured specification captures Stemberger's result without a derivational model of phonology that employs underlying feature blanks.

## *1.2 Similarity*

The use of similarity in cognitive psychology is quite widespread. Many theories of categorization depend on similarity comparisons between objects and category prototypes (e.g. Rosch & Mervis 1975) or individual category members (e.g. Medin & Schaffer 1978). These theories of categorization contrast with standard categories in phonology, like natural classes, which are defined absolutely by rule. For example, the class of coronal consonants are all and only those segments with the feature [+coronal]. These are referred to as classic categories in cognitive psychology.

In its original form (McCarthy 1988), the OCP-Place constraint bans roots with more than one consonant that share place of articulation. In McCarthy's formalization, consonants are effectively divided into classic categories by place of articulation, and roots containing any two homorganic consonants are unacceptable. Pierrehumbert (1993) showed that the effect is gradient, and based on the perceived similarity of homorganic consonants within a root. Thus, while the prohibition against homorganic stops is nearly absolute, roots containing homorganic stops and fricatives are more frequent, and roots containing homorganic obstruents and sonorants are commonly found. The categories of homorganic consonants are not uniform, classical categories. Rather, the degree to which two consonants are judged to be from the same category depends on their similarity. The category of homorganic consonants, to which the OCP-Place constraint applies, is non-uniform, suggesting instead that it is a prototype or exemplar based category.

To appropriately formalize a gradient version of the OCP-Place constraint, a metric of similarity between segments is required. A metric of similarity based on structured specification has a number of advantages over more standard, feature counting metrics of similarity (e.g. Shattuck-Hufnagel & Klatt 1979, van den Broeke & Goldstein 1980, Pierrehumbert 1993). First, by computing similarity based on the natural classes, rather than the individual features, similarity is influenced only by contrastive features. Non-contrastive features do not create additional natural classes. Differences in the relative importance of contrastive and non-contrastive features is found not only in Stemberger's (1991b) speech error data, as mentioned above, but is also found in OCP-Place effects (Pierrehumbert 1993). Second, when similarity is based on natural classes, similarity is increased by both individual feature matches (since a single feature defines a natural class) but also by conjunctions of features (which also define natural classes). The influence of both simple and conjunctive feature matches on similarity has been demonstrated in similarity judgement and categorization tasks (Hayes-Roth & Hayes-Roth 1977, Gluck & Bower 1988, Goldstone 1994b). Note that there is an interaction between the contrastiveness of a feature and conjunctive feature matching: Conjunctions of features only contribute to increasing similarity to the extent that the conjunction of properties contrasts with other conjunctions of properties.

The dependence of the similarity metric on contrast offers a principled solution to one of the basic problems with similarity discussed in the cognitive psychology literature (e.g. Goodman 1972). Tversky (1977) argued that the influence of any particular feature (property) on similarity is based in part on the classificatory significance of features, and that classificatory significance is dependent on the particular object set under study. Similarity is often criticized as being too flexible and unconstrained, in particular in determining the set of features which are considered relevant to the similarity comparison (Goodman 1972; see Goldstone 1994a). When the set of natural classes containing two objects is used to determine their similarity, only a subset of the objects' features actually contribute to forming distinct shared and non-shared natural classes between the two objects. Further, the effect of a particular feature is dependent on the item to which it is compared. Similarity based on structured specification thus provides an objective means of determining which features have classificatory significance.

### *1.3 Frequency and Connectionism*

Accepting the notion of gradient constraints into phonology raises an immediate question: What are the consequences to a form which violates a gradient constraint? In a categorical phonology, violating a constraint results in an unacceptable form. Forms which do not violate a constraint are acceptable. In the binary logic upon which categorical phonology is implicitly based, these extremes can be represented as 0 (unacceptable) and 1 (acceptable). The phonology of a language can be conceived of as a mapping from the space of possible phonological forms to {0,1}. In the stochastic constraint model of a gradient constraint, forms can be more or less acceptable. In a gradient phonology, the space between 0 and 1 is filled in. The phonology of a language can be conceived of as a mapping from the space of possible phonological forms to the interval [0,1]. This is a straightforward extension from the binary logic of a categorical phonology to a gradient phonology based on fuzzy logic (Zadeh 1965). Our question becomes: What does it mean for a form to have acceptability in (0,1)?

Frisch, Broe, & Pierrehumbert (1995) propose that the relevant scale in natural language is frequency. In a categorical phonology, unacceptable items are non-occurring (frequency = 0) and acceptable items are occurring (frequency > 0, modulo the accidental gaps, which can be attributed to sampling as noted in footnote 1 above). In a gradient phonology, not all acceptable forms are equivalent. Forms which are frequent, for example phonotactic patterns which are found in many lexical items, have high acceptability. Infrequent forms have low acceptability.

With the introduction of a fuzzy logic version of acceptability, a formal mechanism for simultaneous constraint combination is immediately available. In a categorical phonology, a form must simultaneously satisfy all constraints in order to be acceptable.[2] Constraints apply conjunctively, which is equivalent to multiplication in the algebra of the binary set {0,1}.Constraint conjunction as multiplication can straightforwardly be extended into the fuzzy set [0,1] (Zadeh 1965, Kosko 1991). The acceptability of a form subject to two or more constraints is equal to the product of the acceptability of the form with respect to each constraint independently.

The approach advocated in this thesis shares fundamental conceptual ground with the theory of connectionism. Connectionist models are inherently gradient, and sensitive to similarity and frequency. In this thesis, I primarily present algebraic models of gradient linguistic behavior. However, to the extent to which these gradient linguistic patterns can be captured by a connectionist model, we need not posit that native speakers have implicit knowledge of algebraic functions and their parameterizations. I present evidence that there is a regular relationship between the algebraic similarity model and a connectionist similarity model, and thus that the results in this thesis are amenable to connectionist modeling. Connectionism provides a psycholinguistically plausible basis for a gradient phonological theory (Dell 1996).

Pinker & Prince (1988) argue against connectionist models of linguistic competence, based on traditional categorical linguistic formalism. The results of this thesis undermine Pinker & Prince’s arguments in two ways. First, I demonstrate predictable gradient effects which cannot be captured in current categorical formalisms. Second, I demonstrate that processing constraints on lexical access impacts the abstract phonology. This interaction undermines the traditional competence/performance distinction and instead suggests that the phonology is constrained by

---

[2] Optimality Theory (Prince & Smolensky 1993) offers an alternative to constraint conjunction which allows categorical constraints with conflicting structural requirements to coexist within a single grammar, by arranging constraints in a strict dominance hierarchy. I discuss extending constraint ranking to gradient constraints in chapter 7.

limits on cognitive processing. Combined with the growing body of evidence that the phonology is also constrained by limits on the articulatory and auditory systems (Lindblom 1983, 1990; Browman & Goldstein 1986, 1990; Clements 1985; Silverman 1995; Flemming 1995) I conclude, following many others, that the formal phonology is not an abstract, arbitrary system, but is instead a reflection of the biological system which must store and manipulate it.

Finally, note that explaining phonological patterns with functional principles need not result in a loss of formal rigor and precision. In this thesis, I apply models which are algebraically well defined, and subject to statistical standards, much like the models of Lindblom (1983, 1990). Similarly, connectionist models of linguistic processing like Dell (1986) and mathematical models of articulatory and acoustic behavior like Lindblom & Sundberg (1978), Task Dynamics (Kelso, Saltzman, & Tuller 1986, Saltzman & Munhall 1989), Perturbation theory (e.g. Mrayati, Carre, & Guerin 1988), and Quantal Theory (Stevens 1978, 1989) are equally precise. The use of continuous variables does not imply imprecise reasoning and untestable conclusions. Rather, models can be compared statistically and be judged by an objective standard of goodness of fit.

### *1.4 Outline of the Thesis*

The thesis is organized as follows. Chapter 2 presents structured specification (Broe 1993), the representation of the segment inventory which I adopt. Chapter 3 introduces the similarity metric of Frisch, Broe, & Pierrehumbert (1995). In chapter 4, I model data from phonological speech errors in English. I show that computing similarity using the representations in structured specification provides the best model of the speech error data. I turn to the OCP-Place constraint in chapter 5. I present the stochastic constraint model (Frisch, Broe, & Pierrehumbert 1995) which is superior to traditional categorical models of OCP-Place effects. Chapter 6 presents new evidence for the special status of word onsets in speech production. In chapter 7, I show that the word onset has special status in the OCP-Place constraint as well. Together, chapters 6 and 7 suggest that the temporal order of a lexical item is a factor in both phonological processing and abstract phonological constraints. I present evidence against underspecification models of speech errors in chapters 8 and 9, demonstrating that structured specification provides a superior alternative. In addition, I suggest that cognitive factors (specifically, frequency and lexical neighborhood density) account for some effects attributed to underspecification. In chapter 10, I show that evidence for underspecification in phonotactics can be given a more satisfactory account using frequency. I also discuss a number of additional frequency effects in phonotactics, which provide evidence that the grammar is frequency sensitive. Chapters 8 through 10 provide an alternative to underspecification that is cognitively motivated and formally coherent. In chapter 11, I discuss the results of the thesis in light of current phonological theories and sketch out an alternative based on connectionist formalism (Dell 1996). Present theories are found wanting as they do not incorporate gradient data into the phonology. The alternative presented in this thesis is a first step toward modeling and formalizing a continuous phonology.

## CHAPTER 2
## Representation of the Phoneme Inventory

In this chapter, I present Broe's (1993) theory of featural specification, STRUCTURED SPECIFICATION, which explicitly represents the segment in the context of the set of natural classes in the segment inventory. This representation depends on the use of distinctive features (Jakobson, Fant, & Halle 1952). Like Halle's BRANCHING DIAGRAMS (Halle 1959) or FEATURE GEOMETRY (Clements 1985, Sagey 1986), structured specification invokes a hierarchy of phonological features. Unlike previous proposals, the redundancy hierarchy in structured specification is unambiguously determined by the set of segments and their feature specifications. The hierarchy represents in a formally coherent way the language specific relations of redundancy and contrast among features. In structured specification, redundancy and defaults are structurally encoded, and the use of feature blanks to encode default or redundant features is avoided. By avoiding underspecification, structured specification is fully compatible with non-derivational, constraint based phonologies.

In the redundancy hierarchy, the segment inventory is represented as a hierarchy of natural classes. A natural class is a set of segments which share a feature or a conjunction of features. Natural classes are at the core of traditional phonological analysis, as phonological processes are assumed to apply only to natural classes. The hierarchy among natural classes arises from the partial ordering of natural classes based on set containment. The redundancy hierarchy provides crucial structural information that reveals the distinctness or redundancy of features. Unlike feature specifications which minimize the number of underlying specifications in some way (Halle 1959, Kiparsky 1982, Archangeli 1984, Steriade 1987), structured specification allows redundancy to be both context dependent and gradient. These are shown to be desirable properties when similarity is calculated for segments in chapter 3.

In using this type of representation for the segment inventory, linguistic classification need not be viewed as a special component of an autonomous linguistic mechanism. Rather, linguistic knowledge of phonological categories is one example of the more general cognitive ability to classify. The study of phonological classification in this thesis is therefore a case study from the general cognitive perspective.[3] In addition, the study of general cognitive categories can be informed by the treatment of contrast and redundancy which is applied here in the phonological domain.

---

[3] Note that this position does not imply that people do not have specific knowledge about language. Rather, the way in which that linguistic knowledge is learned, organized, and applied is consistent with knowledge of other categories. The localized storage of specific linguistic knowledge in the brain can be damaged, leading to language deficits, even though the general mechanism of categorization is unaffected.

*2.1 Features and Natural Classes*

The segment is traditionally represented as a bundle of features (Jakobson, Fant, & Halle 1952; Chomsky & Halle 1968). The segments /a/, /i/, and /u/ for example, can be represented as in (1).

(1)

$$/a/ = \begin{bmatrix} -\text{high} \\ +\text{back} \end{bmatrix} \quad /i/ = \begin{bmatrix} +\text{high} \\ -\text{back} \end{bmatrix} \quad /u/ = \begin{bmatrix} +\text{high} \\ +\text{back} \end{bmatrix}$$

The purpose of featural representations is to capture the NATURAL CLASS behavior of segments (Chomsky & Halle 1968). It is often the case that /i/ and /u/ pattern together, phonologically, but rarely the case that /i/ and /a/ do. The shared [+high] feature between /i/ and /u/ formally captures their grouping as a natural class. The segments /i/ and /a/ share no feature, and are thus not predicted to be a natural class. The purpose of the featural representation is to describe the natural classes of a language as sets of segments having a certain feature or conjunction of features. The natural class is represented here using set notation.

(2) {[+high]} = {i, u}

The natural classes represent both conjunctions of features and sets of segments. A natural class can be a set containing only a single segment. The difference in the representation of /i/ in (1) and (3) is merely notational. Structured specification explicitly encodes the dualism between features and natural classes using a single representation.

(3) {[+high]&[-back]} = {i}

In addition to describing natural classes, features also encode CONTRAST between segments. For example, /i/ and /u/ are distinct segments based on their contrasting values for the feature [back]: /i/ is [-back] and /u/ is [+back]. Note that a feature in and of itself is not contrastive. A feature is only contrastive if it makes a distinction between segments or classes of segments.

Finally, phonological features are traditionally grounded in phonetics. Features have been proposed on acoustic (e.g. Jakobson, Fant, & Halle 1952; Stevens & Keyser 1989; Flemming 1995) and articulatory (e.g. Chomsky & Halle 1968, Sagey 1986) grounds. The properties by which segments are grouped must be observable to be learned. Chomsky & Halle (1968) point out that phonological representation must be grounded in phonetic reality given the “crucial fact that items which have similar phonetic shapes are subject to many of the same rules” (p. 295). To the extent that features are based on gross acoustic and articulatory generalizations, categories can be formed based on articulatory and acoustic similarity. Many non-linguistic categories are grounded in superficial similarity in an analogous manner (Goldstone 1994b).

The formal description of segment inventories using distinctive features raises a number of issues. The discussion which follows makes many of the same points as Broe (1992, 1993:

136-147), though the exposition has been rearranged somewhat. In addition, I adopt particular assumptions about the nature of feature specifications and the classification of features in this thesis which Broe leaves as open questions. In section 2.1.1, I discuss problems with the traditional assumption of bivalent features reviewed in Broe (1992, 1993) and argue for a classification scheme which uses only monovalent features. In section 2.1.2, I consider the notion of a contrast more carefully, and introduce the use of second order feature classes, adopted for classificatory purposes by Broe (1993). Additionally, I represent the notion of an articulatory or acoustic dimension of contrast using feature classes. In section 2.1.3, I present Broe's discussion of the formal inadequacies of the previous generative approaches to redundancy.

### *2.1.1 Feature valency.*

In *The Sound Pattern of English* (Chomsky & Halle 1968, henceforth *SPE*) the featural representations of segments at the level in which they are input to the phonology are fully specified feature matrices with binary feature values. For example, a typical three vowel inventory {a, i, u} would be represented with the matrix in (4), which presents the same feature specifications as used above in an abbreviated format.

(4)

| | /a/ | /i/ | /u/ |
|---|---|---|---|
| [high] | - | + | + |
| [back] | + | - | + |

Each feature encodes a bivalent opposition: [+high] versus [-high]; [+back] versus [-back]. In a binary feature system of this kind, there is an implicit assumption of complementation (Broe 1993). The two specifications for [high] create complementary natural classes: {[+high]} = {i, u}; {[-high]} = {a}. In addition, each feature is assumed to be specified (to have a plus or minus value) for each segment.

The constraints of bivalency and full specification are in some cases problematic. For example, consider the *SPE* classifications for the English places of articulation (p. 177). The four way contrast in place is represented by a pair of binary features.

(5)

| | labial | dental | palato-alveolar | velar |
|---|---|---|---|---|
| [coronal] | - | + | + | - |
| [anterior] | + | + | - | - |

Fant (1969/1973) points out that these specifications depart from the original goal of the feature description as a natural phonetic classification of phonemes. He writes:

> I find the encoding of the class of labial consonants as [+anterior] and [-coronal] to constitute a clear departure from the unifying principles. One single phonetic dimension, 'labiality', which has distinctive function has here lost its identity on the phonological level. (p. 173)

Yip (1989) notes that the natural classes defined by the *SPE* place of articulation features have been a long standing problem for phonology. They are unattested in phonological phenomena. One problem arises from the assumption that set {[-coronal]}, the complement of the natural class {[+coronal]}, is necessarily a natural class. While phonological processes frequently act on the natural class of [+coronal] segments {dental, palato-alveolar}, the natural class of {[-coronal]} = {labial, velar} is never found.[4] Sagey (1986) proposes instead that place of articulation be classified by a set of monovalent features, which have no minus value, corresponding to the active articulators [labial], [coronal], and [dorsal].

It has long been known that the feature [anterior] also defines unattested natural classes. Neither the class of [+anterior] consonants {labial, dental} nor the set of [-anterior] consonants {palato-alveolar, velar} is a natural class. In this case, the problem arises because the feature [anterior] properly sub-classifies the coronals, but does not cross classify all places of articulation. Utilizing feature geometry, Sagey (1986) proposes that [anterior] be made a structural dependent of [coronal], restricting the use of anterior specifications to the coronals. The feature geometric representation of place of articulation features is shown in (6).

| | | | | |
|---|---|---|---|---|
| (6) | labial | dental | palato-alveolar | velar |
| | [labial] | [coronal] | [coronal] | [dorsal] |
| | | \| | \| | |
| | | [+anterior] | [-anterior] | |

In section 2.1.2, I show that Sagey's proposal introduces a confusion of orders into the feature geometry formalism (Broe 1993). Following Broe, I assume that features can be defined over sub-classes without requiring any feature geometric dominance relation between the features. Anteriority is simply left undefined for labials and velars. Similarly, we can think of the monovalent feature [coronal] as undefined for labials and velars.

From the preceding discussion, I conclude that the *SPE* system which used binary cross-classificatory features is too constrained. First, automatic complementation, which is implicitly present in bivalent classification, is an undesirable property of phonological feature matrices. The use of monovalent features, which are either present or absent for any particular segment, avoids this problem in the general case (assuming we cannot refer to the lack of a feature, which would effectively increase the valency of the feature by one to bivalent (Stanley 1967)). Second, some feature values are not defined for all segments. Once again, the general use of monovalent features allows feature specifications to be naturally restricted to subclasses of segments.

In this thesis, therefore, I use all monovalent feature specifications. Note that, given monovalent features, complementation and cross-classification is not necessarily lost. Consider, for example, the following representation of the three vowel inventory {a, i, u} using monovalent features.

---

[4] However, see Lombardi (1996) for recent arguments that [-coronal] is attested. This does not detract from the general point, as there is no reason to assume in advance that the complement of every natural class is a natural class.

(7)

| | /a/ | /i/ | /u/ |
|---|---|---|---|
| [high] | | + | + |
| [low] | + | | |
| [front] | | + | |
| [back] | + | | + |

The two specifications for vowel height, [high] and [low], create complementary natural classes: {[high]} = {i, u}, {[low]} = {a}; as do the specifications of [front] and [back]. Complementation is still possible, but as seen in the case of monovalent place features, it is not a necessary consequence of feature assignment. In addition, the pairs of monovalent features in the three vowel inventory cross-classify the system. In effect, the binary system is reconstructed with a monovalent system which is just a notational variant.

The use of monovalent features allows contrasts of any valency to be represented, without introducing the spurious natural classes required by binary parsing. The three way laryngeal contrast in the Thai stop consonants, for example {t, $t^h$, d}, can be represented with three monovalent features. In a bivalent classification scheme (Halle & Stevens 1971), the pairs {t, $t^h$} and {t, d} are natural classes of [-voice] and [-spread glottis], respectively. Rules involving [-spread glottis] are rare (Kenstowicz 1994). Also, neutralization of laryngeal contrasts tends to leave a single option (e.g. only /t/ in Thai syllable codas) rather than neutralizing [±voice] and [±spread glottis] separately (Lombardi 1991). Groupings of sets of segments along a multivalent contrast can be made, for example {[coronal]} selects the same set of objects as the union of {[interdental]}, {[alveolar]}, and {[palatal]} in English, but such groupings are not required by the formalism. The monovalent system has more descriptive power than the *SPE* system, but the use of the more general system is necessary, given the evidence presented above.[5]

Finally, the use of monovalent features allows all contrasts to be implemented in computationally equivalent ways in the similarity metric introduced in chapter 3. Each feature in common increases similarity and each difference decreases similarity. In a specification system which mixes monovalent and bivalent features, a choice must be made as to whether a difference between bivalent specifications, for example [+voice] and [-voice], is equivalent to a difference in monovalent specifications, for example [labial] and [coronal] (Pierrehumbert 1993). If they are not equivalent, then the notational difference is elevated to a substantive difference, which should be empirically supported. The null hypothesis is that all differences are equivalent (Frisch, Broe, & Pierrehumbert 1995). This hypothesis is adopted here by the use of all monovalent specifications.

---

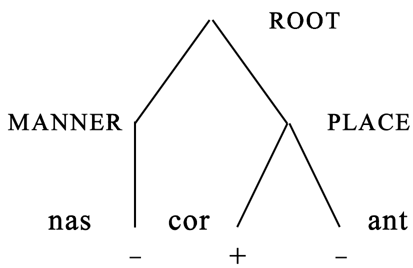
[5] Note that I adopt the total use of monovalent features only for consistency. The distinction between [+voice] and [-voice] as opposed to [voiced] and [voiceless], for purposes of classification, is purely notational. Equivalently, [1 place], [2 place], and [3 place] are notational variants to [labial], [coronal], and [dorsal]. Distinctions between these notations could be made relevant if the phonological system is sensitized to them. For example, alpha-notation (x → [α voice] / _ [α voice]) adds content to the notational distinction. Alpha-notation has been replaced in generative phonology by feature geometry (see Clements 1985 for details), so the values of feature specifications are no longer necessary to the formalism.

### *2.1.2 Dimensions of phonological contrast.*

What is implicit in the bivalent classification systems of Jakobson, Fant, & Halle (1952) and Chomsky & Halle (1968) is that particular specifications lie along an articulatory or acoustic dimension of contrast, like ‘vowel height’ and ‘place of articulation’. However, in cases where the contrast involves a multivalent distinction, additional bivalent features are added and the unidimensionality of the contrast is lost, as remarked by Fant (1969/1973) cited above. The use of CLASS NODES in feature geometry (Clements 1985) or equivalently ATTRIBUTE-VALUE MATRICES in unification grammar (see Broe 1992, 1993: 141-142) provide a formal representation of the notion of a dimension of contrast.

In feature geometry, features are grouped into classes or TIERS, and arranged hierarchically. The grouping is motivated by a desire to capture the fact that certain phonological processes typically act on intuitively well-defined subsets of features rather than arbitrary subsets of features. The class nodes correspond to familiar descriptive groupings like ‘place of articulation’ features and ‘manner’ features. Clements (1985) makes a careful distinction between class nodes and feature nodes. Class nodes can dominate other class nodes or feature nodes, but feature nodes are always terminals in the hierarchy. A partial feature hierarchy, including the superordinate ROOT node, is shown in (8). Dominance can be interpreted as an IS-A relationship. The features [coronal] and [anterior] are PLACE features, [nasal] is a MANNER feature.

(8)




The feature geometry is a graph theoretic representation of an attribute-value matrix (Broe 1992). In an attribute-value matrix, there are first-order features, the standard features of phonological theory, and there are second-order or CATEGORY-VALUED features. While first-order features are primes and cannot be decomposed, category valued features take a feature matrix as a value. This feature matrix may include other category-valued features, parallel to the organization of class nodes in feature geometry. The attribute-value matrix equivalent to (8) is given in (9).

(9)

$$\left[ \text{ROOT} = \left[ \begin{array}{l} \text{PLACE} = \left[ \begin{array}{l} \text{cor} = + \\ \text{ant} = - \end{array} \right] \\ \\ \text{MANNER} = \left[ \text{nas} = - \right] \end{array} \right] \right]$$

The use of category-valued features allows monovalent feature specifications to represent a multivalent dimension of contrast. The previously discussed examples of the three vowel inventory and a three way laryngeal contrast are repeated here, with appropriate class information. The features [high] and [low] are HEIGHT features. The features [front] and [back] are FRONTING features. The features [voiced], [voiceless], and [spread glottis] are LARYNX features. Note that the content features and category-valued features could equivalently be acoustic properties and acoustic dimensions of contrast (cf. Flemming 1995). There is no part of the formalism that specifies the nature of the contrasts as articulatory.

| (10) | | | /a/ | /i/ | /u/ |
|---|---|---|---|---|---|
| | HEIGHT: | [high] | | + | + |
| | | [low] | + | | |
| | FRONTING: | [front] | | + | |
| | | [back] | + | | + |

| (11) | | | /d/ | /t/ | /tʰ/ |
|---|---|---|---|---|---|
| | LARYNX: | [voiced] | + | | |
| | | [voiceless] | | + | |
| | | [spread glot] | | + | |

Using category-valued features, we can refer to *sets* of featural primes, more recently referred to as FEATURE CLASSES (Padgett 1995b, c). For example, Padgett uses the class COLOR to account for harmony which determines both back and round specifications in Turkish. The OCP-Place constraint, which is examined in detail in this thesis, makes reference to the similarity of consonants which share place of articulation features. The use of classes replaces the problematic use of so-called ‘alpha notation’ in *SPE* (Clements 1985). The elimination of alpha specifications is compatible with the use of monovalent features adopted in this thesis.

In combination with the articulatory/acoustic definition of a class as a dimension of contrast, harmony and neutralization have a natural interpretation as processes applying to classes. Given the assumption of all monovalent features, and representations which do not employ underspecification (see section 2.1.3), the formalization of these processes must also take place at the level of the class. The spreading or neutralization of a single feature becomes an incoherent notion.

Sagey (1986) removes the formal distinction between class nodes and feature nodes in feature geometry, by proposing that [±anterior] be a dependent of the [coronal] node. Unlike the original feature geometry model, or attribute-value matrices, a feature node dominates another feature node. This is a confusion of the relative properties of classes and features (Broe 1993). Broe writes:

> A characterization like PLACE is fundamentally different in kind from a characterization like [coronal]. [coronal] is an attribute of objects (segments); PLACE is an attribute of features, not segments. This fact is respected in the everyday speech of phonology, where we talk of ‘coronal segments’, but ‘place

> features'. Just as features like [coronal] allow us to refer to natural classes of segments, terms like PLACE allow us to refer to 'natural classes of features' (Clements 1987), features which behave as a unit in phonological rules. (p. 142) [formatting of feature names adjusted for consistency, SF]

In addition to a conceptual confusion, Sagey's proposal introduces a formal confusion. The relations '[coronal] is a PLACE feature' and 'the value of [anterior] is defined only for [coronal] segments' are both represented through immediate dominance in the feature geometric hierarchy (Broe 1993). Broe notes that:

> This distinction, a fundamental one concerning the domain of classification — objects or features — is not reflected in the *phonological formalism* itself. In current feature geometry, there is no way to tell from the representation what 'type' the relation between dominant and dominated features is. (p. 145)

I show in section 2.2 that the results of Sagey's proposal can be achieved while the distinction between features and classes is maintained, using structured specification (Broe 1993). Thus, Sagey's problematic proposal can be avoided, as the classification of features, on the one hand, and segments, on the other, are independent components of the phonology. The classification of features is based on the category valued features. The classification of segments is based on structured specification.

There is a second use of bivalent feature contrasts in early generative work, to denote distinctness of representations. Roughly speaking, two representations are distinct if they contain opposite values for at least one feature (see Stanley 1967 for the technical details adopted in *SPE*). In the feature specifications I adopt, contrast is defined over feature classes and all features are monovalent. Further, I adopt the hierarchical representation of structured specification. Under these assumptions, two representations are distinct if they contain different monovalent features within at least one class. Informally, two segments are different if they contrast along at least one articulatory or acoustic dimension. I return to this point in the next section, where I discuss the use of feature blanks in phonological theory.

### *2.1.3 Redundancy.*

Traditional approaches to feature specification have distinguished distinctive features from redundant features. Broe (1993, chapter 4) provides an extensive review of the evolution in the treatment of redundancy in generative phonology. We can draw from his discussion the following two conclusions. First, some phonological processes have been shown to be sensitive to the status of a feature as distinctive or redundant. Thus, redundancy must be encoded in the representation. Second, redundant features are traditionally encoded by omission of the feature specifications in underlying representations (so-called UNDERSPECIFICATION). Broe accepts that redundancy must be encoded in phonological representations, but rejects the method of encoding redundancy using feature blanks. The use of feature blanks to encode redundancy creates a number of formal problems for phonology (Broe 1993: 193-209). The problems I discuss here can be grossly characterized as the confusion of representational types.

The confusion of representational types emerges because feature blanks are commonly used to denote other relations apart from redundancy. Blanks have been used in the representations above to indicate that a segment is undefined for a particular feature. Monovalent features, like place of articulation, which are either present or absent are blank for segments where they are absent. Bivalent features which subclassify a set of segments are also blank for all segments outside of that set. In both cases, the absence of a feature does not indicate the redundant presence of a predictable property. If it did, we would generate the same unattested natural classes which promoted the use of monovalent features and subclassification in the first place. The use of blanks to encode 'undefined for [F]' is adopted in structured specification, and thus is employed in this thesis.

A second use of blank is to indicate an ARCHISEGMENT, a segment which is yet to be defined for a particular feature. Broe (1993) summarizes the problem succinctly.

> [An archisegment is] a segment which, in underlying representation, has a *choice* of completions; the particular choice is determined by context, and is in that sense predictable. In the case of redundant blanks, however, the blank indicates a segment which, in underlying representation has its specification for F 'already decided', as it were. In the former case, involving an underdetermined specification, the accompanying features are insufficient to determine what value it should assume: that is why it is left blank. In the case of a predetermined specification, on the other hand, the accompanying features are *totally* sufficient to predict what value it will assume; and *that* is why it is left blank. (p. 196)

In the case of an archisegment, the representation is underdetermined, and in the case of the redundant feature specification it is not. There is no principled way to distinguish between the two based on the representation. Structured specification explicitly represents the difference, through its use of natural classes (which correspond to archisegments) in the representation of the segment inventory. In structured specification, the use of feature blanks to mean 'undefined for F' is thus interpreted equivalently for segments and for natural classes. The use of a feature blank maintains a consistent interpretation and does not conflate distinct relations.

A third use of blanks is introduced in the theory of RADICAL UNDERSPECIFICATION (Archangeli 1984). In radical underspecification, default feature values are left blank, and filled in by derivational rule. Given the assumption that there is a default for every features, there is one default vowel in every language. This is the vowel which has the default value for every vocalic feature. It is represented underlyingly as featureless. Consider a second featureless vowel, the archi-vowel, whose surface form is entirely determined by context. The archi-vowel, too, must be underlyingly featureless (Kisseberth 1971). Broe (1993: 198-202) reviews evidence that default segments must be distinguished representationally from archi-segments, by showing that Klamath and Basque treat the completely unspecified archi-vowel in a distinct manner from the underspecified default vowel. In radical underspecification, both vowels are represented as featureless. Broe instead proposes, parallel to the case of redundant specifications, that the default status of features be encoded hierarchically, in a markedness hierarchy.

I have more to say about radical underspecification in chapters 3, 8, 9 and 10. I would like to note here that the idea of encoding a default segment with no feature specifications is computationally problematic for a similarity metric based on features and natural classes. First note that the only sensible interpretation of two segments being mutually undefined for a feature is that this is neither a feature match nor a feature mismatch. If the fact that two segments are mutually undefined for a feature is counted as a feature match, then we implicitly assume feature complementation and the unattested natural classes which are generated by complements. The default segment has no feature matches with any other segment, and thus is equally similar to every other segment. In addition, it has no feature mismatches with itself, and its self similarity is undefined. I show in chapter 8 that the default status of features is not directly relevant for determining similarity, and that any effect attributed to the default status of a specification can instead be attributed to the higher frequency of the default versus the non-default specifications.

There an additional formal and conceptual problem with omitting feature specifications based on their predictability. This is the well known problem of reciprocal dependency. The choice of which feature to omit as predictable is often arbitrary (Halle 1959, Stanley 1967, Broe 1993). Consider the matrix of the three vowel inventory with both backness and rounding features.

(12)

| | /a/ | /i/ | /u/ |
|---|---|---|---|
| [high] | | + | + |
| [low] | + | | |
| [front] | | + | |
| [back] | + | | + |
| [round] | | | + |
| [non-round] | + | + | |

In this case, every feature is predictable from other features or combination of features, so any feature could in principle be eliminated. For example, [+back]&[+non-round] ⇒ [+low], [+round] or [+low] ⇒ [+back]. Eliminating features due to their predictability requires an arbitrary choice. Further note that the choice has consequences for the set of natural classes which are available. Eliminating [+non-round] eliminates the natural class {a, i}. What is a redundant feature from the perspective of individual segments is in fact not redundant in the description of natural classes. In general, proposals concerning redundancy and underspecification have been overly concerned with the description of individual segments and less concerned with the consequences for natural classes. I now turn to the theory of structured specification, which represents both the individual segments and the natural classes explicitly, providing an elegant solution to the representation of redundancy and reciprocal dependencies.

## *2.2 Structured Specification*

The problems introduced by multiple interpretations of representational blanks and reciprocal dependencies is solved in the theory of structured specification. In structured specification, feature blanks are used solely to represent ‘undefined for [F]’. Redundancy is encoded in the REDUNDANCY HIERARCHY. There is a second hierarchy of specification in Broe (1993), the DEFAULT HIERARCHY, which indicates the marked and unmarked status of feature specifications. Properties which were formerly encoded by blanks are differentiated formally. In this thesis, I am concerned primarily with the redundancy hierarchy, as it is the representation of the segment inventory upon which similarity is computed. For more information on the default hierarchy, see Broe (1993).

Given a segment inventory and a set of features, the redundancy hierarchy for that set of segments given that feature set can be unambiguously determined. The hierarchy is based on the partial ordering of natural classes of segments given a featural representation. The natural classes are ordered by set containment: larger natural classes contain smaller ones. I exemplify the algorithm with a simple case, the three vowel inventory. The interested reader should consult Broe (1993) for a more rigorous treatment of the set theoretic and graph theoretic ideas employed here.

Recall the feature specifications for the three vowel inventory presented above.

(13)

| | /a/ | /i/ | /u/ |
|---|---|---|---|
| [high] | | + | + |
| [low] | + | | |
| [front] | | + | |
| [back] | + | | + |

To construct a redundancy hierarchy for the three vowel inventory, we first consider the set of natural classes which can be denoted by the feature matrix in (13). There are 4 feature values, so there are $2^4 = 16$ possible conjunctions of features. Each feature conjunction with its corresponding natural class is given in (14). The symbol ∅ denotes the empty set, which is the natural class created by an incompatible conjunction of features.

(14)
{[]} = {a, i, u}
{[+high]} = {i, u}
{[+low]} = {a}
{[+front]} = {i}
{[+back]} = {a, u}
{[+high]&[+low]} = ∅
{[+high]&[+front]} = {i}
{[+high]&[+back]} = {u}
{[+low]&[+front]} = ∅
{[+low]&[+back]} = {a}
{[+front]&[+back]} = ∅
{[+high]&[+low]&[+front]} = ∅
{[+high]&[+low]&[+back]} = ∅
{[+high]&[+front]&[+back]} = ∅
{[+low]&[+front]&[+back]} = ∅
{[+high]&[+low]&[+front]&[+back]} = ∅

Out of the 16 possible conjunctions of features, there are 7 distinct sets of segments. These are the natural classes. They are {a, i, u}, {a, u}, {i, u}, {a}, {i}, {u}, ∅. These 7 sets are partially ordered by the following set containment relationships.

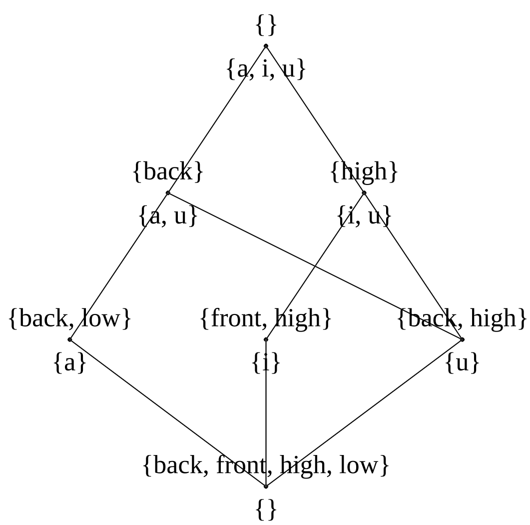
(15)
{a, i, u} ⊇ {a, u}, {i, u};
{a, u} ⊇ {a}, {u};
{i, u} ⊇ {i}, {u};

{a}, {i}, {u} ⊇ ∅

Note that not all set containment relationships are given in (15). Relationships which can be deduced by transitivity are omitted. For example, {a, i, u} ⊇ {u}, which can be deduced from {a, i, u} ⊇ {a, u} and {a, u} ⊇ {u}.

A LATTICE is a partial ordering of the natural classes of segments which are possible given a featural representation. Figure 2.1 shows the lattice of the three vowel inventory graphically. Each node in the lattice is a natural class, and the set of features and segments which the node denotes are shown above and below the node, respectively. Nodes are ordered from top to bottom by size. The top node of the lattice represents the entire inventory, and the bottom node is the empty set. The row of nodes just above the bottom are the natural classes containing the individual segments. The features which denote these nodes are the features of the segments. Lines connecting nodes indicate set containment. Note that, as in (15), not all set containment relationships are indicated by lines, those which can be deduced by transitivity are excluded.

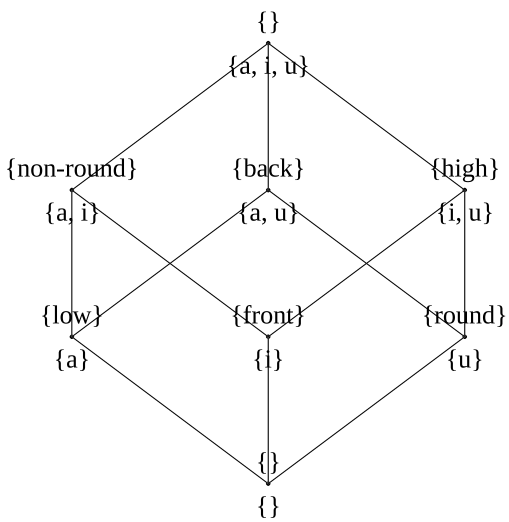
**Figure 2.1**: Lattice of the three vowel inventory.




There is a dualism between sets of segments (natural classes) and sets of features in the lattice. The hierarchical set containment relationship between the natural classes corresponds to an inheritance relation for the features that define those natural classes. For example, the natural class {u} is {[+back]&[+high]}. It is contained by the natural class {i, u}, which is {[+high]} and by the natural class {a, u} which is {[+back]}. The natural class {u} inherits the feature [+back] from the natural class {a, u} and it inherits [+high] from {i, u}. Through set containment and feature inheritance, the lattice represents redundancy structurally. For example, {[+front]} = {i} is contained by {[+high]} = {i, u}. Thus, every segment which is [+front] is a member of {[+high]}, in other words [+front] ⇒ [+high]. Featural redundancy can be ‘read off of’ the lattice.

The redundancy hierarchy explicitly represents the natural class structure of the segment inventory, given the feature assignments. Naturally, given different feature assignments, different natural classes may result. For example, consider the effect of characterizing the three vowel inventory using rounding in addition to backness. Adding the feature [+round] in this example does not alter the hierarchy. In this case, {[+round]} = {u} is not a new natural class, so no additional nodes or containment relations are created by the additional feature. However, if we were to add the feature [+non-round], we would define a new natural class: {a, i}. In effect, [+round]/[+non-round] adds a new dimension of contrast in the phonology. The lattice of the three vowel inventory which includes [+non-round] is given in Figure 2.2.[6]

**Figure 2.2**: Lattice of the three vowel inventory, with [+non-round].



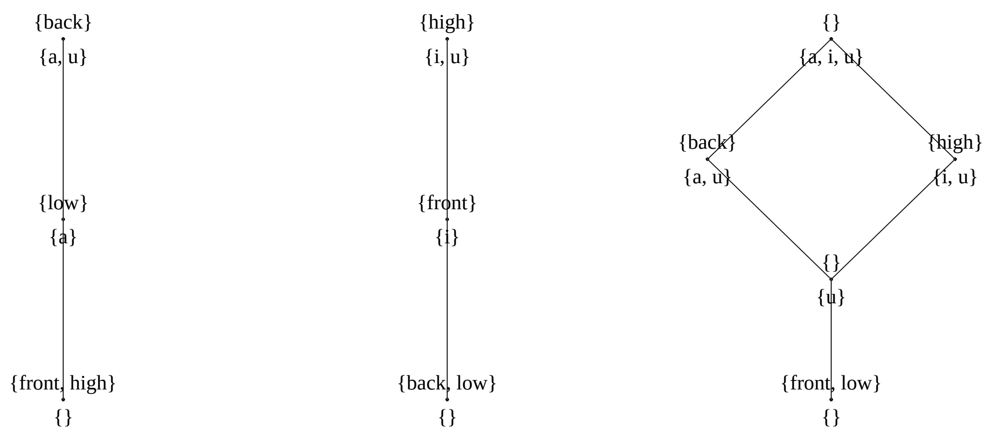

With addition of [+non-round], the redundancy hierarchy contains all of the logically possible natural classes for {a, i, u}. There is thus an upper bound on the number of features which can affect the natural class structure of the language, where adding additional features to the lattice does not change it. This property of the redundancy hierarchy is desirable in the creation of a similarity metric for segments, as it puts a limit on the number of features which are relevant to similarity based on the segment inventory of the language. In the general case, not all sets of segments are natural classes, so the set of features is further constrained.

The redundancy hierarchy represents the system of natural classes. The representation of a segment in structured specification is the sub-lattice representing all natural classes which contain that segment. Figure 2.3 shows the segment lattices for the three vowel inventory from Figure 2.1. Note that the redundancy relationships are obvious when the segment lattices are

[6] To reduce visual clutter, inherited features have been removed from the display of the lattice. For example, the natural class {a} has features [+low], [+back], and [+non-round]. However, [+non-round] and [+back] are inherited from {a, i} and {a, u} respectively, and so are not displayed at the {a} node, though they are present. In general, lattices will be displayed in the thesis this way for clarity.

examined: [+low] ⇒ [+back] and [+front] ⇒ [+high].

**Figure 2.3**: Segment lattices of the three vowel inventory.




The redundancy hierarchy also explicitly represents the archisegments. The natural classes which contain more than one segment are archisegments. They are nodes which are not sufficiently specified to individuate a particular segment.

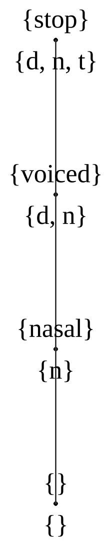
The redundancy hierarchy reveals a serious inadequacy of current proposals to use monovalent feature assignments to represent the 'inertness' of certain features, so-called TRIVIAL UNDERSPECIFICATION or INHERENT UNDERSPECIFICATION (e.g. Archangeli & Pulleyblank 1994, Lombardi 1991, Steriade 1995). These proposals contrast from radical and contrastive underspecification (Archangeli 1984, Steriade 1987) in that the inert features are *never present* in the representation. In radical and contrastive underspecification, feature blanks are filled in by rule, leading to fully specified surface forms.

Consider, for example, the use of privative [nasal] and [voiced] features, as in the feature matrix in (16), to capture the lack of phonological processes which refer to [-nasal] and [-voice]. Such feature assignments fail to render all segments distinct (Broe 1995).

| (16) | | /t/ | /d/ | /n/ |
|---|---|---|---|---|
| | [stop] | + | + | + |
| | [voiced] | | + | + |
| | [nasal] | | | + |

Figure 2.4 shows the lattice for the corresponding redundancy hierarchy. The lattice reveals that there is no natural class containing /t/ distinct from /d/ or /n/, and similarly that /d/ has no natural classes distinct from /n/. There is no way based on these feature assignments to distinguish /t/ from the archi-stop. The problem is that the set of features of /t/ are a proper subset of the sets of features of /d/ and /n/, and their feature sets are thus non-distinct.

**Figure 2.4**: Hierarchy of non-distinct segments.




The lack of a feature specification does not create a contrasting property, since I have assumed monovalent features and cannot refer to blanks. In order to avoid confounding completely specified segments with archisegments, the representation must contain a sufficient number of contrastive features to INDIVIDUATE all of the segments (Broe 1995; Frisch, Broe, & Pierrehumbert 1995). I here define a segment to be individuated in a segment inventory by a feature matrix if it is present as the sole member of a natural class in the redundancy hierarchy generated by that feature matrix. Equivalently, given the dualism between features and natural classes, a segment is individuated in a particular segment inventory by a feature matrix if the features of that segment are *not* a subset of the features of any other segment.

In order to individuate /t/, /d/, and /n/, there must be a feature possessed by /t/ that is not possessed by /d/ and /n/, and a feature possessed by /d/ that is not possessed by /n/. The simplest solution, shown in (17), reconstructs the bivalent distinctions of traditional feature specifications, using notationally equivalent monovalent features. Recall that in the general case the problem is not equivalent to the use of bivalent oppositions. While any bivalent specification can be given an equivalent monovalent specification, the reverse is not true. Note that many contrasts are indeed bivalent, which is one reason why bivalent features were so attractive in the first place, but that we do not want bivalency imposed on us by the representational system.

(17)

| | t | d | n |
|---|---|---|---|
| [stop] | + | + | + |
| [voiced] | | + | + |
| [voiceless] | + | | |
| [nasal] | | | + |
| [oral] | + | + | |

The representation of the phonological inventory using features and natural classes does not treat the knowledge required to classify phonological categories in a special way. Any domain where objects can be defined on the basis of distinctive properties could also be represented using the redundancy hierarchy and lattices. In general, work in cognitive psychology has focussed on constructed feature sets which are orthogonal, so that there is no redundancy. It is hoped that applying the formally coherent representation of redundancy using structured specification to create a redundancy sensitive metric of similarity in this thesis will encourage work which examines the effects of redundant versus distinctive information in general cognition.

### *2.3 Representation of the English Consonant Inventory*

The feature assignments I use for the English consonant segments are given in (18). As discussed above, all features are monovalent features. The feature assignments were made with three goals in mind. First, I attempted to represent the natural classes of English consonants. Many of the features are thus familiar ones from *SPE* and Jakobson, Fant, & Halle (1952). Second, the features are based on articulatory or acoustic properties. While articulatorily based features have been the standard, Stevens & Keyser (1989) discuss a number of acoustic correlates of many of the features used here. Third, all segments had to be individuated. The feature inventory is thus rich enough to distinctively identify every segment. The classifications presented in Ladefoged & Maddieson (1996) were used as an overall guide.

The features in (18) have been divided into dimensions of contrast. These are second order features which classify the features into ARTICULATOR, PLACE, STRICTURE, MANNER, and LARYNGEAL features. These classes could equally be represented hierarchically using feature geometry, and the second order classes themselves can be grouped into SUPRALARYNGEAL and ROOT classes. Within the classes, some features cross-classify the entire inventory and other features sub-classify subsets of segments (Stevens & Keyser 1989). Many features adopted from Stevens & Keyser (1989) are used over smaller subsets of segments than originally proposed, taking advantage of the monovalent classification system which is not restricted to bivalency and allows sub-classification.

The ARTICULATOR features correspond to the active (supralaryngeal) articulators for the segment (Sagey 1986). The PLACE features are the points of contact for the active articulators. They can be thought of either as sub-classifying the ARTICULATOR features or as independently cross-classifying the entire inventory.[7] The STRICTURE features roughly characterize the degree of constriction of the vocal tract. The [obstruent] and [sonorant] features cross-classify the inventory and are sub-classified by the [stop] and [continuant] features and [glide] and

---

[7] The division of 'place of articulation' features into articulator features and features representing point of contact has not been proposed to my knowledge, though Keating (1991) does note the difference between the passive and active articulatory status of [coronal] and [anterior]. This is basically the classification system used by Ladefoged & Maddieson (1996). Such a division might be useful in contrasting phonological assimilation processes, which cause wholesale changes in both articulator and point of contact, with coarticulation processes, which are gradiently influenced by points of contact but generally do not affect the active articulator.

[consonantal] features, respectively. The MANNER features further sub-classify the stops, continuants, and consonantal sonorants. Finally, the LARYNGEAL features contain the traditional voicing contrast and the additional [spread glottis] feature for /h/ (Halle & Stevens 1971). Interestingly, the [spread glottis] feature is necessary to individuate /h/, since /h/ has no (supralaryngeal) ARTICULATOR or PLACE features.

(18) a. ARTICULATOR features:

| | p | b | f | v | m | t | d | θ | ð | s | z | ʃ | ʒ | tʃ | dʒ | k | g | ŋ | l | r | n | w | y | h |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| labial | + | + | + | + | + | | | | | | | | | | | | | | | + | | + | | |
| coronal | | | | | | + | + | + | + | + | + | + | + | + | + | | | | + | + | + | | + | |
| dorsal | | | | | | | | | | | | | | | | + | + | + | | | | | | |

b. PLACE features:

| | p | b | f | v | m | t | d | θ | ð | s | z | ʃ | ʒ | tʃ | dʒ | k | g | ŋ | l | r | n | w | y | h |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| bilabial | + | + | | | + | | | | | | | | | | | | | | | + | | + | | |
| dental | | | + | + | | | | + | + | | | | | | | | | | | | | | | |
| alveolar | | | | | | + | + | | | + | + | | | | | | | | + | + | + | | | |
| palatal | | | | | | | | | | | | + | + | + | + | | | | | | | | + | |
| velar | | | | | | | | | | | | | | | | + | + | + | | | | | | |

c. STRICTURE features:

| | p | b | f | v | m | t | d | θ | ð | s | z | ʃ | ʒ | tʃ | dʒ | k | g | ŋ | l | r | n | w | y | h |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| obstruent | + | + | + | + | | + | + | + | + | + | + | + | + | + | + | + | + | | | | | | | + |
| sonorant | | | | | + | | | | | | | | | | | | | + | + | + | + | + | + | |
| stop | + | + | | | + | + | + | | | | | | | + | + | + | + | + | | | | | | |
| continuant | | | + | + | | | | + | + | + | + | + | + | | | | | | | | | | | + |
| glide | | | | | | | | | | | | | | | | | | | | | | + | + | |
| consonantal | | | | | + | | | | | | | | | | | | | + | + | + | + | | | |

d. MANNER features:

| | p | b | f | v | m | t | d | θ | ð | s | z | ʃ | ʒ | tʃ | dʒ | k | g | ŋ | l | r | n | w | y | h |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| oral | + | + | | | | + | + | | | | | | | | | + | + | | | | | | | |
| affricate | | | | | | | | | | | | | | + | + | | | | | | | | | |
| strident | | | | | | | | | | + | + | + | + | + | + | | | | | | | | | |
| distributed | | | + | + | | | | + | + | | | | | | | | | | | | | | | + |
| lateral | | | | | | | | | | | | | | | | | | | + | | | | | |
| rhotic | | | | | | | | | | | | | | | | | | | | + | | | | |
| nasal | | | | | + | | | | | | | | | | | | | + | | | + | | | |

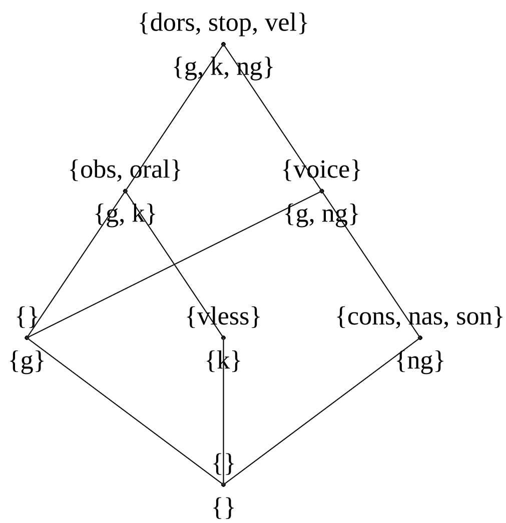
e. LARYNGEAL features:

| | p | b | f | v | m | t | d | θ | ð | s | z | ʃ | ʒ | tʃ | dʒ | k | g | ŋ | l | r | n | w | y | h |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| voice | | + | | + | + | | + | | + | | + | | + | | + | | + | + | + | + | + | + | + | |
| voiceless | + | | + | | | + | | + | | + | | + | | + | | + | | | | | | | | + |
| spread glottis | | | | | | | | | | | | | | | | | | | | | | | | + |

There is a well defined redundancy hierarchy for the English consonants. Displaying such a hierarchy is impractical, however, as it is extremely large and complex. Instead, I present here three separate lattices, one for each active articulator. Many of the redundancy relationships are present in the articulator lattices, so their examination is useful. These lattices provide the reader with additional examples of redundancy hierarchies, which are at the core of the theory of structured specification. Finally, lattices analogous to these are used in chapter 5 to compute the similarity of homorganic consonants in Arabic for the OCP-Place constraint.

The simplest articulator sub-lattice for English is the dorsal lattice, shown in Figure 2.5. Due to font limitations of the graphics routines, the velar nasal /ŋ/ is 'ng' in the lattice. Recall that inherited features are omitted from the display to reduce clutter, and that set containment relationships which can be deduced from transitivity are also implicit. For example, the {g} node is not featureless, it inherits the features [+obstruent] and [+oral] from the {g, k} node, the feature [+voice] from the {g, ng} node, as well as the features [+dorsal], [+velar], and [+stop] from the {g, k, ng} node.

**Figure 2.5**: Lattice of the English dorsal consonants.




The dorsal lattice reveals the well-known redundancy relationship [+sonorant] ⇒ [+voice], and the equally valid but less often mentioned relationship [+voiceless] ⇒ [+obstruent]. This lattice shows there are also a number of predictable relationships given [+dorsal]. Given [+dorsal], we can predict [+stop] and [+velar]. Also, given [+dorsal], [+sonorant] ⇔ [+consonantal] ⇔ [+nasal] and [+obstruent] ⇔ [+oral]. Thus, the use of sub-lattices like the dorsal lattice explicitly represents that the contrastiveness of a feature is context dependent. These relationships do not hold in the other articulator sub-lattices. Contrastive underspecification (Steriade 1987) would omit some of these features under just these circumstances. I show in chapter 5 that this facet of structured specification captures crucial differences in similarity effects between place of articulation classes in the OCP-Place constraint (Frisch, Broe, & Pierrehumbert 1995; cf. Pierrehumbert 1993).

The labial articulator sub-lattice is more complex than the dorsal lattice, due to the larger inventory of labials in English. The labial lattice is shown in Figure 2.6, on the following page. This lattice reveals a reciprocal dependency between PLACE features and both STRICTURE and MANNER features among the labials: the stops are bilabial and the fricatives are labio-dental.

Also, the [+sonorant] ⇒ [+voice] and [+voiceless] ⇒ [+obstruent] relationships are repeated here.

The lattice of coronal consonants is the most complex. This lattice is shown in Figure 2.7, two pages below. The segments /θ, ð, ʃ, ʒ, tʃ, dʒ/ are shown as 'th, dh, sh, zh, tsh, dzh'. In addition, the segments corresponding to each natural class, except for the bottom most row, have been left out. They can be deduced from the set containment relationship. Redundancy relationships between features are still represented. Once again, [+sonorant] ⇒ [+voice] and [+voiceless] ⇒ [+obstruent] can be found. The other redundancy relationships in the lattice primarily reflect the sub-classification. In this lattice, as with the labials, [+distributed] ⇔ [+dental], indicating that this particular feature pair represents a clear case where articulatory and acoustic features have a one-to-one relationship.

**Figure 2.6**: Lattice of the English labial consonants.

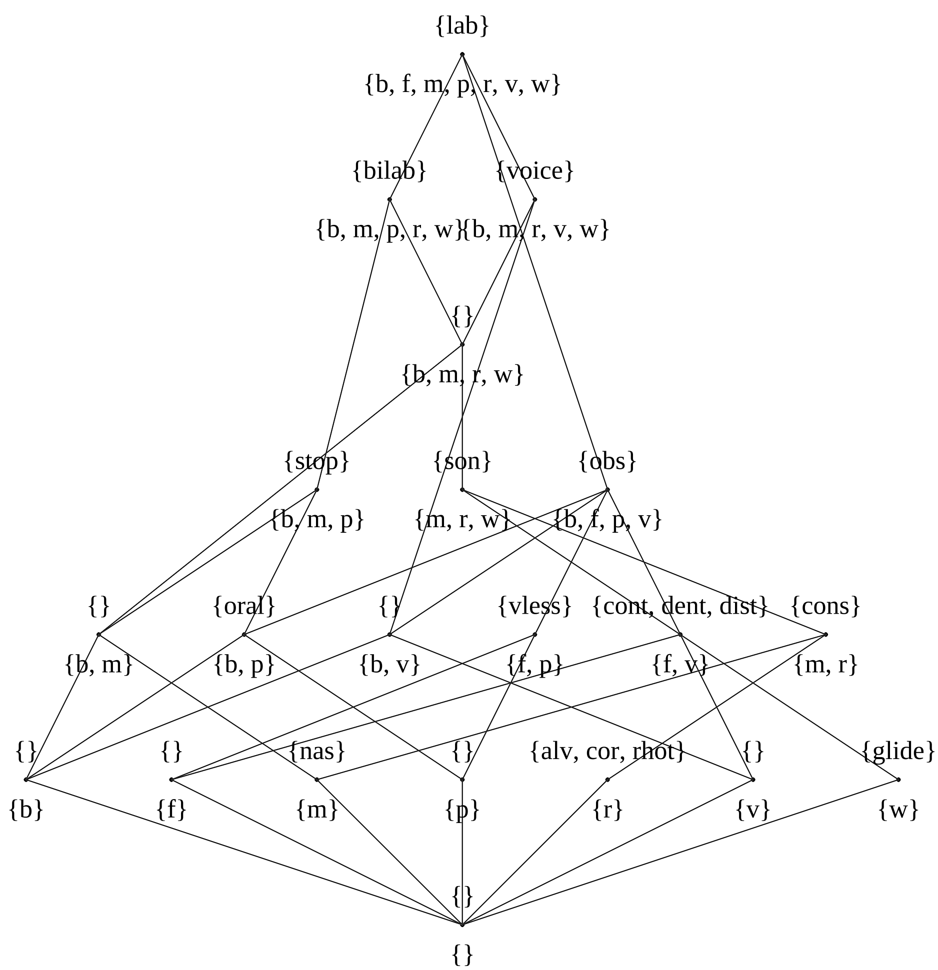
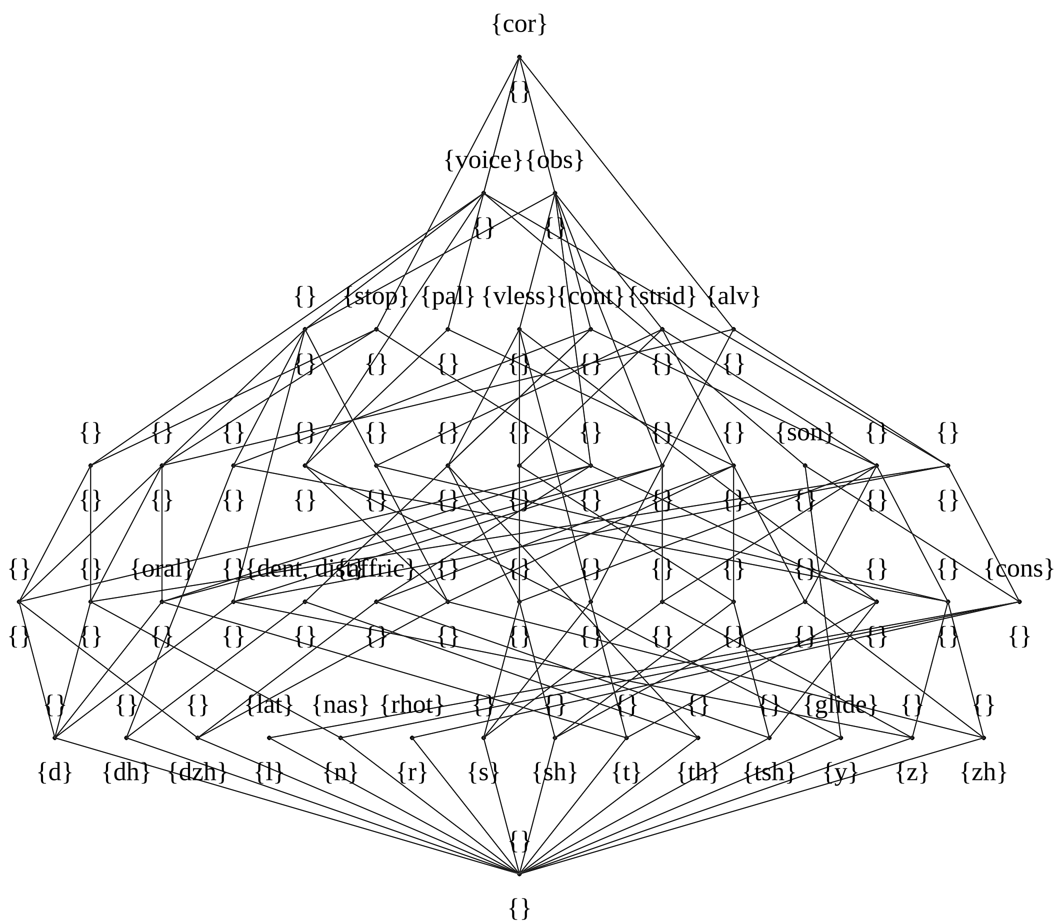


**Figure 2.7**: Lattice of the English coronal consonants.



# CHAPTER 3
# Similarity of Phonological Segments

Features represent the degree of similarity between two segments. If two segments share a feature, they pattern together for any phonological phenomena that depend on that feature. Further, if features are grounded in articulatory or auditory contrast, then there is a degree of 'superficial' similarity between segments that share a feature. Psycholinguists typically use simple feature counting in quantitative arguments for similarity effects (e.g. Shattuck-Hufnagel & Klatt 1979) and have compared different feature representations by comparing their predictions for similarity (e.g. van den Broeke & Goldstein 1980, Stemberger 1991b). In this chapter, I present a model of similarity which has significant empirical and conceptual advantages over feature counting models.

There is a large body of work on similarity in the cognitive psychology literature. One very influential model which computes similarity based on features is the FEATURE-CONTRAST MODEL (Tversky 1977, Tversky & Gati 1982). I adopt a variant of the feature-contrast model to compute similarity in this thesis. The model I adopt computes similarity using the natural classes of the redundancy hierarchy (Frisch, Broe, & Pierrehumbert 1995). Thus, features still play a role in determining similarity, but *relations* between features influence similarity as well. Computing similarity using the redundancy hierarchy takes into account the distinctive or redundant status of a feature. Redundant features have less of an influence on similarity than distinctive ones (Frisch, Broe, & Pierrehumbert 1995).

In computing similarity over the redundancy hierarchy, conjunctions of features in addition to individual features contribute to the determination of similarity. Conjunctions of features have been shown to influence similarity judgements (Hayes-Roth & Hayes-Roth 1977; Gluck & Bower 1988; Goldstone, Medin, & Gentner 1991; see Goldstone 1994a). Connectionist or spreading activation models of similarity can capture the influence of conjunctions of features (Gluck & Bower 1988, Goldstone 1994a). The metric of similarity I adopt is a closed form alternative to connectionist models for computing the similarity of segments.

Section 3.1 introduces the feature-contrast model (Tversky 1977, Tversky & Gati 1982) and the model of similarity based on the natural classes of the redundancy hierarchy (Frisch, Broe, & Pierrehumbert 1995), showing that they are closely related in general form. In section 3.2, I present some advantages of the natural classes model by demonstrating that it is sensitive to the distinctive/redundant status of features, and that it models the effect that feature conjunctions have on similarity. I also review some general arguments against the use of similarity as a basis for categorization and show that the natural classes model handles these problems in a natural and non-arbitrary way. In combination with the general applicability of the redundancy hierarchy to broad cognitive categories, like natural kind classes, the natural classes model of similarity is a genuine alternative for cognitive psychology. In section 3.3, I present the similarity of English consonant pairs, based on the feature assignments from chapter 2. Section 3.4 compares similarity in the natural classes model with the activation level of a spreading activation network of segments and features, showing a close relationship between the two.

### *3.1 The Feature-Contrast Model and Natural Classes Model*

Intuitively, determining similarity between objects (segments) based on a set of properties (features) involves comparing how many properties are shared by the objects and how many properties are not. Computational models of similarity use analogous principles, computing similarity as a function of shared and non-shared features. In this section, I present the Frisch, Broe, & Pierrehumbert (1995) model of similarity between segments. This model is an adaptation of the model of similarity used in Pierrehumbert (1993) that uses natural classes instead of features. Before I proceed, I present a frequently used model of similarity in the cognitive psychology literature. This model is more general than the model I adopt, and thus highlights some of the assumptions in the natural classes model.

#### *3.1.1 Modeling similarity using features.*

Tversky (Tversky 1977, Tversky & Gati 1978) introduced a general feature counting metric for similarity called the feature-contrast model. The feature-contrast model computes similarity as a function of shared and non-shared features between two objects. The general equation for the feature-contrast model is given in (19).

(19) similarity$(X, Y) = F[\theta\, f(X \cap Y) - \alpha\, f(X - Y) - \beta\, f(Y - X)]$,
where $F$ is an increasing function,
$\theta, \alpha, \beta$ are positive constants,
$f$ is a measure function of the features,
$X \cap Y$ denotes the features shared by $X$ and $Y$,
$X - Y$ denotes the features in $X$ but not in $Y$, and
$Y - X$ denotes the features in $Y$ but not in $X$.

The function $F$ allows flexibility in the relationship of the model to the data (linear, exponential, logarithmic, etc.). The constants $\theta, \alpha, \beta$ are used to model asymmetries in similarity data, and to influence the relative weight of feature matches against feature mismatches. The function $f$ converts a set of features to a quantitative value. It is generally just the size of the set of features, and so is merely a feature count.

Pierrehumbert (1993) computed the similarity of segments using a normalized feature model. In her model similarity is defined to be:

(20)
$$\text{similarity} = \frac{\text{shared features}}{\text{shared features} + \text{non-shared features}}$$

This model shares many basic properties with the Frisch, Broe, & Pierrehumbert (1995) model. The Pierrehumbert (1993) model is symmetric, and the self-similarity of all segments is 1. In addition, the range of similarity values is [0,1]. The Pierrehumbert (1993) model is a special case of the feature-contrast model which is discussed by Tversky (1977). We can derive the

Pierrehumbert (1993) model by setting the weights to $\theta = 1$, $\alpha = \beta = 0$, letting $F(x) = x$, and dividing by a normalizing factor.

(21)
$$\text{similarity}(X,Y) = \frac{f(X \cap Y)}{f(X \cap Y) + f(X - Y) + f(Y - X)}$$

*3.1.2 Similarity in the redundancy hierarchy.*

In comparison to the feature-contrast model, the similarity model of Frisch, Broe, & Pierrehumbert (1995) is relatively simple. This model computes similarity over the redundancy hierarchy, and thus determines similarity based on natural classes rather than features. There is, of course, a close relation between the two, as single features do pick out natural classes, and the other natural classes are the result of conjunctions of features. The natural classes model is identical to the similarity model of Pierrehumbert (1993), but it counts natural classes instead of features. I show in section 3.2 that there are certain advantages of the natural classes model due to the use of the redundancy hierarchy. The natural classes similarity model I employ is:

(22)
$$\text{similarity} = \frac{\text{shared natural classes}}{\text{shared natural classes} + \text{non-shared natural classes}}$$

There are two properties of the feature-contrast model which the similarity model of Frisch, Broe, & Pierrehumbert (1995) does not possess. First, the feature-contrast model allows the 'self-similarity' of objects to be different. The feature-contrast model subtracts the measures of feature mismatches from the measure of feature matches. The model is not normalized, and so the size of the feature set can directly influence similarity. For example, Tversky (1977) presented schematic-face stimuli to subjects for direct similarity judgements. He found that when subjects were comparing identical faces, the more complex stimuli that had additional features were given higher self-similarity. More complex schematic-faces had higher similarity to themselves than simple schematic-faces did. In the natural classes model, the self-similarity of every segment is one, by assumption, and similarity ranges over [0,1].

The second property of the feature-contrast model which is not possessed by the Frisch, Broe, & Pierrehumbert (1995) model is that the feature-contrast model can account for asymmetries in similarity. Asymmetries can be modeled by changing the weighting coefficients $\alpha$ and $\beta$. If $\alpha > \beta$, then the properties possessed by $X$ but not by $Y$ are weighted more heavily than the properties possessed by $Y$ but not by $X$. According to Tversky (1977), when people rate the similarity of item $X$ to item $Y$ directly (e.g. by answering the question "How similar is $X$ to $Y$?"), item $X$ serves as the subject of the comparison, and its properties become more salient. This increased salience is modeled by weighting. Asymmetries of this kind were found in the schematic-face stimuli (Tversky 1977). In some of the stimuli, subjects judged the similarity of faces with many features ($f(X)$ large) to analogous faces with fewer features ($f(Y)$ small). Subjects rated the complex faces as less similar to the simple faces than the reverse. With $\alpha > \beta$, similarity($X$, $Y$) is reduced more by $f(X - Y)$ than similarity($Y$, $X$). In an analogous task with more natural stimuli, subjects rated the similarity of countries. Subjects rated less prominent countries (e.g. North Korea) as more similar to prominent countries (e.g. China) than the reverse (Tversky & Gati 1978).

In this thesis, I primarily analyze data from two sources: phonological speech errors between consonants, and the OCP-Place constraint, which restricts the number of similar homorganic consonants in a word. In the case of phonological speech errors, the question of self-similarity is moot. If errors do occur between pairs of identical consonants, there is no clear evidence for it in the phonemic output. Differences in self-similarity may be an issue in the OCP-Place constraint, as the constraint is strongest for consonants which are maximally similar. However, as we will see in chapter 5, since there are very few occurrences of identical consonant pairs of any kind, the crucial data are sparse. Thus, I feel that the natural classes model, which is normalized with respect to the total number of natural classes, is sufficient to model these data.

By contrast, there are almost certainly principled asymmetries in the speech error and OCP-Place data which are not be captured by the natural classes model. In both cases, the currently available data are not likely to reveal statistically significant asymmetries. The feature-contrast model can be used as the basis for a natural classes model of similarity which can capture asymmetries. We could count natural classes instead of features in the feature-contrast model. In equation (19), let $f(X \cap Y)$ be the natural classes shared by $X$ and $Y$, let $f(X - Y)$ be the natural classes of $X$ but not $Y$, and similarly for $f(Y - X)$. Obviously, this metric could model asymmetric similarity and differences in self-similarity. In this thesis, I do not argue for the exact form of the metric. Rather, I argue for the use of the redundancy hierarchy as the basis of similarity.

In order to analyze such asymmetries, controlled experiments are needed to generate additional data. For example, Stemberger (1991a), discussed in chapter 9, found only two significant asymmetries in speech error rates between consonants in his corpus. He presents a series of experiments investigating asymmetries in speech error rates. In order to test for asymmetries in OCP-Place effects, experimental work on implicit phonological knowledge such as Pierrehumbert (1994) and Treiman et al. (1996) could be applied. A thorough investigation of asymmetries in similarity effects is beyond the scope of this thesis. The data I analyze are sufficiently symmetric that significant generalizations can be made based on a symmetric similarity model. The study of a natural classes version of the feature contrast model is also left as an open research topic. Such a study would require a larger data set than what is considered in this thesis in order to estimate the additional parameters accurately.

I next compute sample similarity values for the three vowel inventory. The basic feature matrix for the three vowel inventory used in chapter 2 is repeated in (23a). The additional rounding features, also used in chapter 2, are given in (23b). The lattices based on the features without and with rounding from chapter 2 are repeated in Figure 3.1. Note that the lattice without rounding features displays all features, while the lattice with rounding features displays only non-inherited features. Similarity values for the three vowels with and without rounding features for the two lattices are given in Table 3.1. Since in this model, similarity is symmetric, only the lower triangular half of the table is filled.

(23)

| | | | |
|---|---|---|---|
| a. | /a/ | /i/ | /u/ |
| [high] | | + | + |
| [low] | + | | |
| [front] | | + | |
| [back] | + | | + |
| b. | | | |
| [round] | | | + |
| [non-round] | + | + | |

**Figure 3.1**: Lattices of the three vowel inventory, without and with rounding features.

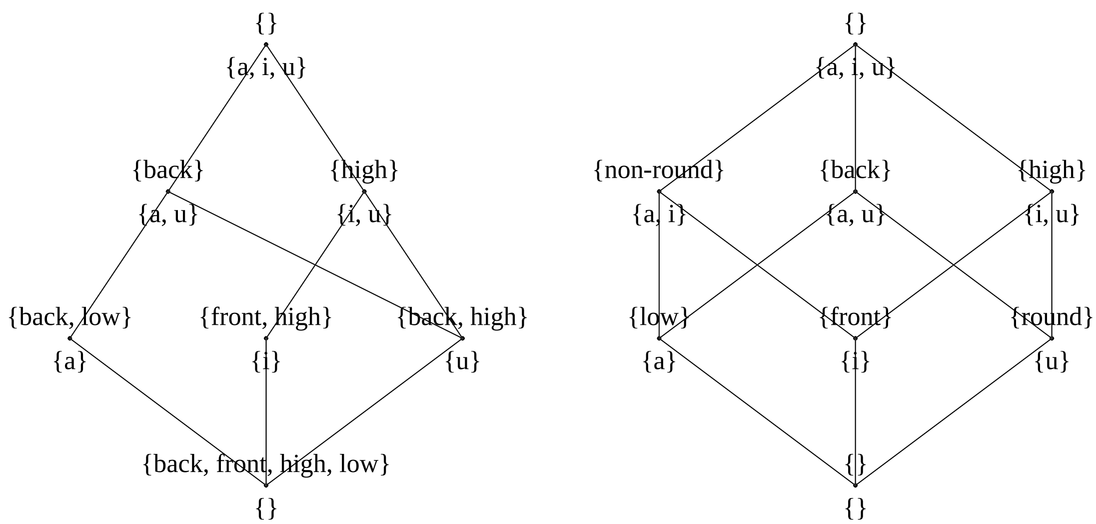



**Table 3.1**: Similarity of {a, i, u} using natural classes, without and with rounding.

| Without rounding | | | | | With rounding | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | /a/ | /i/ | /u/ | | | /a/ | /i/ | /u/ |
| /a/ | shared | 3 | | | /a/ | shared | 4 | | |
| | non-shared | 0 | | | | non-shared | 0 | | |
| | **similarity** | **1** | | | | **similarity** | **1** | | |
| /i/ | shared | 1 | 3 | | /i/ | shared | 2 | 4 | |
| | non-shared | 4 | 0 | | | non-shared | 6 | 0 | |
| | **similarity** | **0.2** | **1** | | | **similarity** | **0.33** | **1** | |
| /u/ | shared | 2 | 2 | 3 | /u/ | shared | 2 | 2 | 4 |
| | non-shared | 3 | 3 | 0 | | non-shared | 6 | 6 | 0 |
| | **similarity** | **0.4** | **0.4** | **1** | | **similarity** | **0.33** | **0.33** | **1** |

We can compare these similarity values to what would be computed by the Pierrehumbert (1993) model, and a simple parameterization of the feature-contrast model, given in (24). Values are shown in Table 3.2.

(24) similarity($X$,$Y$) = $f(X \cap Y) - f(X - Y) - f(Y - X)$

**Table 3.2**: Similarity of {a, i, u} using features, without and with rounding features.

| | Without rounding | | | | With rounding | | | |
|---|---|---|---|---|---|---|---|---|
| Pierrehumbert model | | /a/ | /i/ | /u/ | | /a/ | /i/ | /u/ |
| | /a/ | 1 | | | /a/ | 1 | | |
| | /i/ | 0 | 1 | | /i/ | 0.2 | 1 | |
| | /u/ | 0.33 | 0.33 | 1 | /u/ | 0.2 | 0.2 | 1 |
| Feature-contrast model | | /a/ | /i/ | /u/ | | /a/ | /i/ | /u/ |
| | /a/ | 2 | | | /a/ | 3 | | |
| | /i/ | -4 | 2 | | /i/ | -3 | 3 | |
| | /u/ | -1 | -1 | 2 | /u/ | -3 | -3 | 3 |

Comparing the tables, we find (apart from differences in magnitude) that the three metrics of similarity make nearly identical predictions. In the case without rounding, the similarity of /a/ to /i/ is less than the similarity of /a/ to /u/ and /i/ to /u/. In the case with rounding, the similarity of all non-identical pairs is the same. Comparing across rounding cases, we see that the similarity of /a/ to /i/ with rounding is between the similarity of /a/ to /i/ without rounding and the similarity of /a/ to /u/ and /i/ to /u/ without rounding for all three metrics. There is one difference. The 'without rounding' and 'with rounding' cases differ in the total number of features each segment possesses. From Tversky's (1977) perspective, the 'with rounding' vowels are more complex than the 'without rounding' vowels, and have higher self-similarity. The Pierrehumbert (1993) model and the natural classes model of Frisch, Broe, & Pierrehumbert (1995) have self-similarity of 1 in all cases by assumption.

## *3.2 Advantages of the Natural Classes Model*

A simple case, like the three vowel inventory, is not sufficient to highlight the differences between the natural classes model and metrics of similarity based directly on features. In this section, I present two advantages of the natural classes model. First, the natural classes model incorporates the distinctive or redundant status of a feature into the similarity computation. Redundant features influence similarity less than non-redundant features. The effect of redundancy is also gradient and context dependent. Second, the natural classes model incorporates a 'synergistic' effect of multiple feature matches into the similarity computation that has been found in experiments involving similarity judgments and categorization.

### *3.2.1 The problem of features and redundancy.*

It has been argued that the cognitive notion of similarity is only well defined if the computation of similarity can be based in a principled manner on a reasonable and relevant set of features (Goodman 1972; Tversky 1977; Medin, Goldstone, & Gentner 1993; see Goldstone 1994b for a review). If the computation is only based on counting individual features, the type and number of features used has a great deal of influence on the computation. This was evident in the similarity computations for {a, i, u} using the feature-contrast model above. In the case where rounding features are included, self-similarity is greater than when they are not.

Recall that the system of natural classes is only altered by adding features which create new natural classes, in other words features which are contrastive within the set of segments. The set of relevant features is thus based on attributes across the entire inventory and not just for one individual segment. There are, at most, $2^n$ natural classes that can be created out of a set of *n* objects, so there is a strict upper bound on the number of features that can affect similarity in the system. In phonological classification, as the number of features increases, the level of redundancy in the feature matrix also increases.

Tversky (1977) demonstrates experimentally that 'diagnostic factors' influence the effect a particular feature has on similarity. He writes:

> The diagnostic factors are highly sensitive to the particular object set under study. For example, the feature "real" has no diagnostic value in the set of actual animals, since it is shared by all actual animals and hence cannot be used to classify them. This feature, however, acquires considerable diagnostic value if the object set is extended to include legendary animals, such as a centaur, a mermaid, or a phoenix. (p. 342)

Tversky's diagnostic and non-diagnostic features are equivalent to the distinctive and redundant features in linguistic theory. Computing similarity over the redundancy hierarchy gives differential weight to features based on redundancy, providing a model of Tversky's diagnostic factors. Contrastive features have more weight than totally redundant ones.

By computing similarity over the natural class structure, three degrees of redundancy are differentiated. The first case is a TOTALLY REDUNDANT feature. A totally redundant feature adds no new natural classes to the redundancy hierarchy. The feature [+round] in the three vowel inventory has this property. The addition of a totally redundant features does not affect similarity. Totally redundant features are not independently contrastive. Another pair of totally redundant features are the ARTICULATOR feature [+dorsal] and the PLACE feature [+velar] in English. In English, all consonants articulated with the tongue dorsum make a constriction in the velar region. The two features [+dorsal] and [+velar] influence similarity like a single feature. In other words, the two features are a single contrastive unit, and they are present in the exact same natural classes in the lattice. More formally, for features [x] and [y], let *n*([x]) be the set of natural classes which are [+x]. Let *n*([+y]) be the set of natural classes which are [+y]. Then [x] and [y] are totally redundant if and only if *n*([x]) = *n*([y]).

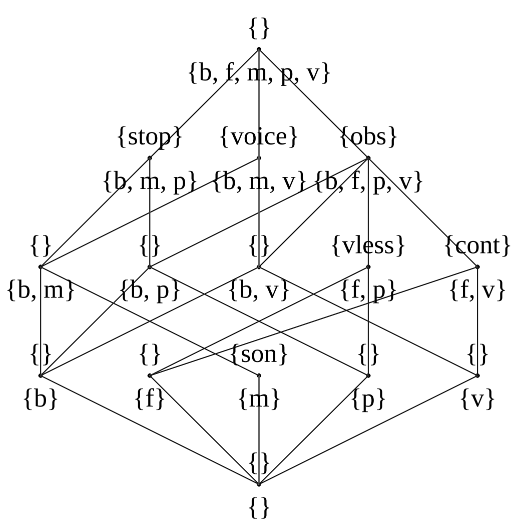
A feature can also be PARTIALLY REDUNDANT. A classic example of partial redundancy is found in the voicing of sonorants. Consider the feature specifications in (25). The feature [+voice] is redundant for sonorants, but [+voice] is contrastive for obstruents. Similarly [+obstruent] is redundant for voiceless consonants, but contrastive among voiced ones.

(25)

| | /p/ | /b/ | /f/ | /v/ | /m/ |
|---|---|---|---|---|---|
| [sonorant] | | | | | + |
| [obstruent] | + | + | + | + | |
| [stop] | + | + | | | + |
| [continuant] | | | + | + | |
| [voiced] | | + | | + | + |
| [voiceless] | + | | + | | |

Partially redundant features have a reduced effect on similarity. Consider the redundancy hierarchy in Figure 3.2. While /f/ is a member of five natural classes, /v/ is a member of six natural classes. Since [+voiceless] consonants are always obstruents, the set {[+obstruent]}⊃ {[+voiceless]}. When determining similarity between voiceless obstruents, they will have a shared natural class due to the [+voiceless] feature, and a shared natural class due to the [+obstruent] feature. By contrast, when determining similarity between voiced obstruents, they will have a shared natural class for [+voice] and a shared natural class for [+obstruent], as well as a shared natural class for [+voice]&[+obstruent]. Thus, all other things being equal, the voiceless obstruents are less similar to one another than the voiced obstruents. In chapter 4, I show that this is the correct prediction based on speech error data, which are known to depend on similarity. The features [+low] and [+front] in the three vowel inventory are another example of partially redundant features.

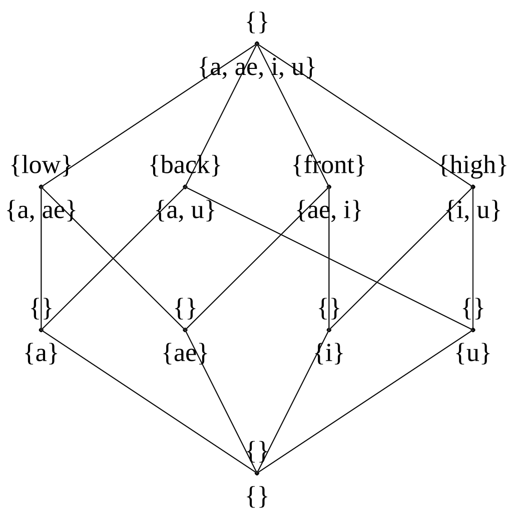
**Figure 3.2**: Redundancy hierarchy including partially redundant features.




For partially redundant features, the presence of one feature implies the other, so the set of natural classes which contain one feature is a subset of the set of natural classes containing the other. If features [x] and [y] are partially redundant with respect to one another, then either $n([x]) \subset n([y])$ or $n([y]) \subset n([x])$.

Features like [+voice] and [+obstruent] in the previous example, or [+high] and [+back] in the three vowel inventory, are NON-REDUNDANT with respect to one another. Note that redundancy can only be determined with respect to a particular segment inventory and feature matrix. The form of the redundancy hierarchy, and consequently the similarity values, change on a context dependent basis (see below). Some languages employ voiceless sonorants, and in those languages, [+sonorant] is a non-redundant feature.[8] In a four vowel inventory {a, æ, i, u}, [+low] and [+front] are non-redundant. The lattice of the four vowel inventory is shown in Figure 3.3.

**Figure 3.3**: Redundancy hierarchy with no redundant features.




Note that every feature in this lattice defines a natural class of two segments, and no feature is ordered with respect to any other by set containment. Non-redundant features have no hierarchical relationship with respect to one another. If features [x] and [y] are non-redundant with respect to one another, then neither $n([x]) \subseteq n([y])$ or $n([y]) \subseteq n([x])$. As we saw above, non-redundant features have the greatest effect on similarity, as they contribute natural classes based on their individual features, as well as conjunctions with other features.

Goodman (1972) remarks that the context dependence of similarity weakens the

[8] Voiceless sonorants are very rare (Ladefoged & Maddieson 1996), and thus presumably highly marked. Recall that markedness is not encoded in the redundancy hierarchy, it is encoded in a separate markedness hierarchy, and thus markedness has no direct effect on similarity as it is computed in this thesis. However, see chapter 9 for some indirect effects of markedness via frequency. I leave open the possibility that the markedness hierarchy is also a factor in determining similarity.

explanatory value of similarity so much that it is no longer needed. For example, when we say *'X* is similar to *Y* with respect to *Z*', he claims that the 'with respect to *Z*' has all of the explanatory power. The natural classes model of similarity, along with the classification of features using category-valued features, provides a principled basis upon which to reject Goodman's argument. For example, we can compute the similarity of consonants with respect to manner of articulation by creating a sub-lattice of the inventory using the STRICTURE and MANNER features. The 'with respect to' part fixes the relevant feature classes, and similarity within the sub-lattice is computed normally. The contrastiveness and redundancy relationships between features is different in different domains, but the means of computing similarity is independent of the domain. The general notion of similarity still has explanatory power, even though it is context dependent.

### *3.2.2 Synergistic effects in similarity.*

We saw above that non-redundant features increase similarity in a more than linear manner. When segments share two non-redundant features, they share three natural classes. When segments share three non-redundant features, they share seven natural classes: {[+F1]}, {[+F2]}, {[+F3]}, {[+F1]&[+F2]}, {[+F1]&[+F3]}, {[+F2]&[+F3]}, {[+F1]&[+F2]&[+F3]}. When segments share *n* non-redundant features, they share $2^n$-1 natural classes.

Synergistic effects of multiple feature matches on similarity have been found in experiments on categorical cue learning (Hayes-Roth & Hayes-Roth 1977, Gluck & Bower 1988) and in direct similarity judgments (Goldstone, Medin, & Gentner 1991). The synergy of multiple feature matches has been modeled using a SIMPLE AND CONJUNCTIVE FEATURES MODEL (see Goldstone 1994a). This model counts features and conjunctions of features toward similarity. This is identical to the natural classes model in the case of non-redundant features. The natural classes model has an advantage over the simple and conjunctive features model when redundancy is encountered among the features.

This synergistic property has also been modeled using network of spreading activation models of similarity (Gluck & Bower 1988, Goldstone 1994a). The lattice representation provides a close match to the implementation of similarity by spreading activation. An explicit comparison of similarities and differences between the natural classes similarity model and a spreading activation model is made in section 3.4.

## *3.3 Similarity of English Consonants*

In this section, I present computed similarity values based on the natural classes model for the English consonants. These computations are based on the feature specifications given in chapter 2. The first example, as it was in chapter 2, is the dorsal consonants. Note that the similarity of the dorsal consonants computed over this sub-lattice is different from the similarity of the dorsal consonants when the redundancy hierarchy for the entire inventory is used (see below). Computing similarity using the natural classes model over sub-lattices is a *context dependent* measure of similarity, as noted above. Some features which are totally redundant in the dorsal sub-lattice are either partially redundant or non-redundant in the hierarchy of the entire inventory. For example, the features [+oral], [+sonorant], and [+consonantal] are not contrastive within the dorsal inventory but are contrastive in the coronals. The similarity of consonants over the dorsal sub-lattice is given in Table 3.3.

**Table 3.3**: Similarity over the dorsal sub-lattice.

| | k | g | ŋ |
|---|---|---|---|
| k | 1 | | |
| g | 0.4 | 1 | |
| ŋ | 0.2 | 0.4 | 1 |

Analogous calculations can be made for the more complex labial and coronal sub-lattices. Similarity for these sub-lattices is given in Tables 3.4 and 3.5. Once again, similarity over the sub-lattice is not identical to similarity over the entire inventory.

**Table 3.4**: Similarity over the labial sub-lattice.

| | p | b | f | v | m | r | w |
|---|---|---|---|---|---|---|---|
| p | 1 | | | | | | |
| b | 0.42 | 1 | | | | | |
| f | 0.33 | 0.15 | 1 | | | | |
| v | 0.18 | 0.33 | 0.38 | 1 | | | |
| m | 0.23 | 0.46 | 0.08 | 0.15 | 1 | | |
| r | 0.17 | 0.31 | 0.09 | 0.18 | 0.6 | 1 | |
| w | 0.18 | 0.33 | 0.1 | 0.2 | 0.5 | 0.63 | 1 |

**Table 3.5**: Similarity over the coronal sub-lattice.

| | t | d | θ | ð | s | z | ʃ | ʒ | tʃ | dʒ | l | r | n | y |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| t | 1 | | | | | | | | | | | | | |
| d | 0.4 | 1 | | | | | | | | | | | | |
| θ | 0.19 | 0.10 | 1 | | | | | | | | | | | |
| ð | 0.11 | 0.2 | 0.36 | 1 | | | | | | | | | | |
| s | 0.3 | 0.15 | 0.31 | 0.16 | 1 | | | | | | | | | |
| z | 0.17 | 0.33 | 0.15 | 0.33 | 0.36 | 1 | | | | | | | | |
| ʃ | 0.13 | 0.07 | 0.31 | 0.16 | 0.47 | 0.2 | 1 | | | | | | | |
| ʒ | 0.08 | 0.14 | 0.15 | 0.33 | 0.2 | 0.45 | 0.36 | 1 | | | | | | |
| tʃ | 0.32 | 0.16 | 0.18 | 0.11 | 0.23 | 0.12 | 0.42 | 0.21 | 1 | | | | | |
| dʒ | 0.17 | 0.33 | 0.10 | 0.2 | 0.11 | 0.23 | 0.2 | 0.45 | 0.38 | 1 | | | | |
| l | 0.12 | 0.21 | 0.08 | 0.15 | 0.11 | 0.21 | 0.05 | 0.1 | 0.05 | 0.1 | 1 | | | |
| r | 0.12 | 0.21 | 0.08 | 0.15 | 0.11 | 0.21 | 0.05 | 0.1 | 0.05 | 0.1 | 0.75 | 1 | | |
| n | 0.21 | 0.42 | 0.06 | 0.12 | 0.09 | 0.17 | 0.04 | 0.08 | 0.09 | 0.17 | 0.5 | 0.5 | 1 | |
| y | 0.06 | 0.1 | 0.08 | 0.17 | 0.05 | 0.1 | 0.11 | 0.22 | 0.12 | 0.22 | 0.3 | 0.3 | 0.21 | 1 |

Computing similarity over place of articulation sub-lattices is relevant for modeling OCP-Place effects (Frisch, Broe, & Pierrehumbert 1995). In the case of Arabic, which I discuss in detail in chapter 5, the contrastiveness of features has a crucial effect (Pierrehumbert 1993). For example, /f/ and /m/ have a very strong cooccurrence constraint, while /θ/ and /n/ have a very weak one. The labial inventory consists only of {b, f, m}, while the coronal inventory contains 14 consonants. Among the coronals, [+stop], [+continuant], [+dental], [+alveolar], [+voice], and [+voiceless] are all non-redundant. Among the labials [+stop], [+bilabial] and [+voice], as well as [+continuant], [+dental], and [+voiceless], are mutually redundant. In the case of the coronals, the cross-classification creates a large number of non-shared natural classes which reduce the similarity of /θ/ to /n/ relative to the similarity of /f/ to /m/.

Using the redundancy hierarchy derived from the feature specifications given in chapter 2 (which was not displayed as a lattice due to its size and complexity) and the natural classes

| | p | b | f | v | m | t | d | θ | ð | s | z | ʃ | ʒ | tʃ | dʒ | k | g | ŋ | l | r | n | w | y | h |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| p | 1 | | | | | | | | | | | | | | | | | | | | | | | |
| b | 0.4 | 1 | | | | | | | | | | | | | | | | | | | | | | |
| f | 0.26 | 0.13 | 1 | | | | | | | | | | | | | | | | | | | | | |
| v | 0.15 | 0.3 | 0.38 | 1 | | | | | | | | | | | | | | | | | | | | |
| m | 0.19 | 0.39 | 0.07 | 0.15 | 1 | | | | | | | | | | | | | | | | | | | |
| t | 0.3 | 0.14 | 0.1 | 0.06 | 0.06 | 1 | | | | | | | | | | | | | | | | | | |
| d | 0.14 | 0.28 | 0.05 | 0.11 | 0.11 | 0.39 | 1 | | | | | | | | | | | | | | | | | |
| θ | 0.11 | 0.06 | 0.43 | 0.19 | 0.03 | 0.2 | 0.11 | 1 | | | | | | | | | | | | | | | | |
| ð | 0.07 | 0.12 | 0.19 | 0.39 | 0.06 | 0.12 | 0.23 | 0.38 | 1 | | | | | | | | | | | | | | | |
| s | 0.1 | 0.05 | 0.18 | 0.1 | 0.03 | 0.3 | 0.15 | 0.4 | 0.2 | 1 | | | | | | | | | | | | | | |
| z | 0.06 | 0.11 | 0.09 | 0.19 | 0.06 | 0.17 | 0.33 | 0.19 | 0.44 | 0.37 | 1 | | | | | | | | | | | | | |
| ʃ | 0.1 | 0.05 | 0.18 | 0.1 | 0.03 | 0.18 | 0.1 | 0.4 | 0.2 | 0.58 | 0.24 | 1 | | | | | | | | | | | | |
| ʒ | 0.06 | 0.11 | 0.09 | 0.19 | 0.06 | 0.11 | 0.2 | 0.19 | 0.44 | 0.24 | 0.57 | 0.37 | 1 | | | | | | | | | | | |
| tʃ | 0.21 | 0.11 | 0.1 | 0.06 | 0.06 | 0.44 | 0.22 | 0.21 | 0.13 | 0.27 | 0.14 | 0.41 | 0.21 | 1 | | | | | | | | | | |
| dʒ | 0.11 | 0.22 | 0.06 | 0.11 | 0.11 | 0.22 | 0.47 | 0.11 | 0.24 | 0.13 | 0.28 | 0.19 | 0.44 | 0.39 | 1 | | | | | | | | | |
| k | 0.44 | 0.19 | 0.14 | 0.08 | 0.08 | 0.35 | 0.16 | 0.13 | 0.08 | 0.11 | 0.06 | 0.11 | 0.06 | 0.25 | 0.13 | 1 | | | | | | | | |
| g | 0.21 | 0.42 | 0.08 | 0.16 | 0.15 | 0.17 | 0.33 | 0.07 | 0.15 | 0.06 | 0.13 | 0.06 | 0.13 | 0.14 | 0.27 | 0.39 | 1 | | | | | | | |
| ŋ | 0.09 | 0.15 | 0.04 | 0.09 | 0.37 | 0.07 | 0.13 | 0.04 | 0.08 | 0.04 | 0.07 | 0.04 | 0.07 | 0.07 | 0.13 | 0.17 | 0.33 | 1 | | | | | | |
| l | 0.04 | 0.07 | 0.04 | 0.08 | 0.17 | 0.11 | 0.19 | 0.08 | 0.17 | 0.11 | 0.22 | 0.07 | 0.14 | 0.07 | 0.13 | 0.05 | 0.09 | 0.24 | 1 | | | | | |
| r | 0.1 | 0.19 | 0.07 | 0.14 | 0.44 | 0.09 | 0.16 | 0.06 | 0.13 | 0.09 | 0.18 | 0.06 | 0.11 | 0.06 | 0.11 | 0.04 | 0.07 | 0.17 | 0.56 | 1 | | | | |
| n | 0.06 | 0.12 | 0.03 | 0.06 | 0.26 | 0.19 | 0.38 | 0.06 | 0.13 | 0.09 | 0.18 | 0.06 | 0.11 | 0.12 | 0.24 | 0.07 | 0.14 | 0.33 | 0.53 | 0.4 | 1 | | | |
| w | 0.14 | 0.25 | 0.09 | 0.19 | 0.44 | 0.03 | 0.06 | 0.04 | 0.08 | 0.04 | 0.07 | 0.04 | 0.07 | 0.04 | 0.06 | 0.05 | 0.09 | 0.18 | 0.17 | 0.42 | 0.12 | 1 | | |
| y | 0.04 | 0.07 | 0.04 | 0.09 | 0.13 | 0.07 | 0.13 | 0.08 | 0.17 | 0.07 | 0.14 | 0.12 | 0.23 | 0.12 | 0.21 | 0.05 | 0.09 | 0.18 | 0.40 | 0.29 | 0.27 | 0.25 | 1 | |
| h | 0.15 | 0.08 | 0.47 | 0.21 | 0.04 | 0.12 | 0.06 | 0.41 | 0.19 | 0.23 | 0.11 | 0.23 | 0.11 | 0.13 | 0.07 | 0.19 | 0.1 | 0.06 | 0.06 | 0.04 | 0.04 | 0.06 | 0.06 | 1 |

**Table 3.6**: Similarity of English consonant pairs using the natural classes model.

similarity model, I computed the similarity of all English consonant pairs. This is a 'context free' similarity with no restriction to a sub-lattice.[9] Table 3.6 presents the pairwise similarity of the segments of English that I will be using in the analysis of speech errors in chapter 4. I believe that roughly comparable similarity values would result from different feature assignments, but I present evidence in chapter 4 that these feature assignments make a very good prediction of speech error rates, which are a function of similarity (Nooteboom 1969, MacKay 1970, Fromkin 1971).

### *3.4 The Natural Classes Model and Spreading Activation*

One of the ways in which synergistic effects of multiple feature matches on similarity have been modeled is by the simple and conjunctive features model (see Goldstone 1994a). This model is derived from a model of exemplar classification called the PROPERTY-SET MODEL (Hayes-Roth & Hayes-Roth 1977). In the property set model, exemplars are encoded as the set of component properties and combinations of properties of the exemplar. Thus, each match in a single feature or in a conjunction of features increases the likelihood of classifying two exemplars as belonging to the same category. Similarity provides a natural basis for classification in this way (Goldstone 1994b). More recent models of similarity and classification have employed spreading activation or connectionist methods (Gluck & Bower 1988, Goldstone 1994a). In these models, activation (and in some models, inhibition) spreads bidirectionally between nodes representing objects and properties over time. The degree of mutual activation between objects indicates the degree of similarity. Once again, synergistic effects of feature matches are found, as properties which activate the same object reinforce one another indirectly via that object. In this section, I compare the predictions of the simple and conjunctive features model, a simple spreading activation model, and the natural classes model.

It is clear that there is at least one difference between the natural classes model and the models of spreading activation and the simple and conjunctive features model. The natural classes model constrains the influence of features based on their relative contrastiveness. As discussed above, there is a natural limit to the number of features which affect similarity in the natural classes model and not all features weigh equally. I present evidence in chapters 4 and 5 that the natural classes model makes the correct prediction.

Consider the five hypothetical inventories of objects and features given in (26), chosen to test other possible differences between the models. Inventories a, b, c, and d have a gradual progression from having one natural class of segments to the point where every combination of segments is a natural class. Inventories a, b, and e show a development from one shared natural class between objects A and B, to three shared natural classes between A and B. Inventories a, b, and e have increasing numbers of contrastive features shared by A and B.

---

[9] Of course, this could be considered a sub-lattice of the lattice of the entire English segment inventory, including the vowels. The degree to which the vowel sub-lattice and the consonant sub-lattice are independent depends on whether or not the same features are used for vowels and consonants. This is an issue beyond the scope of this thesis.

(26)

| a. | F0 | F1 | F2 | F3 |
|---|---|---|---|---|
| A | + | + | | |
| B | + | | + | |
| C | + | | | + |

| b. | F0 | F1 | F2 | F3 | F4 |
|---|---|---|---|---|---|
| A | + | + | | + | |
| B | + | + | | | + |
| C | + | | + | | |

| c. | F0 | F1 | F2 | F3 | F4 |
|---|---|---|---|---|---|
| A | + | + | | + | |
| B | + | + | + | | |
| C | + | | + | | + |

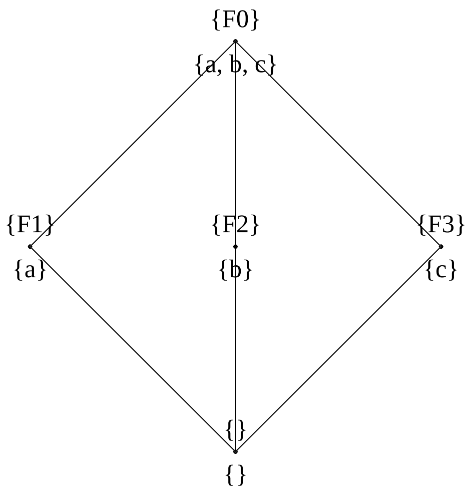
| d. | F0 | F1 | F2 | F3 |
|---|---|---|---|---|
| A | + | + | | + |
| B | + | + | + | |
| C | + | | + | + |

| e. | F0 | F1 | F2 | F3 | F4 | F5 | F6 |
|---|---|---|---|---|---|---|---|
| A | + | + | + | + | | | |
| B | + | + | + | | + | | |
| C | + | + | | | | + | |
| D | + | | | | | | + |

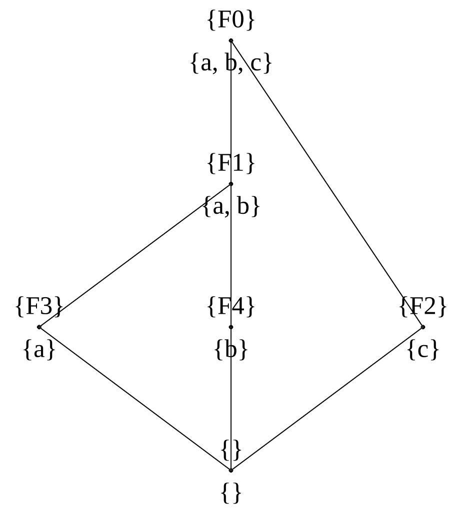
Lattices corresponding to the redundancy hierarchies for these five inventories are shown in figure 3.4. The lattices clearly reflect the changes in natural class structure. Lattices for Inventories a, b, c, and d reveal increasing cross-classification. The lattices for Inventories a, b, and e contain progressively more shared classes between A and B.

**Figure 3.4**: Lattices for the five hypothetical Inventories a-e.

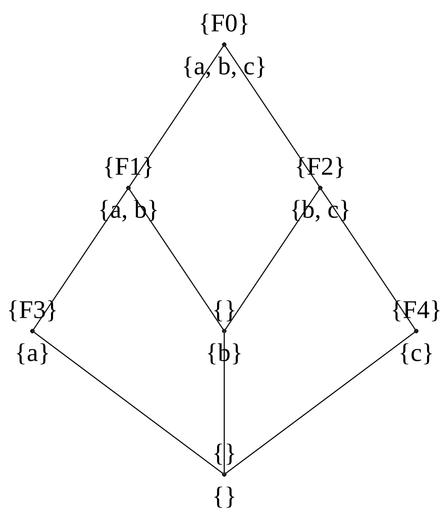
a.



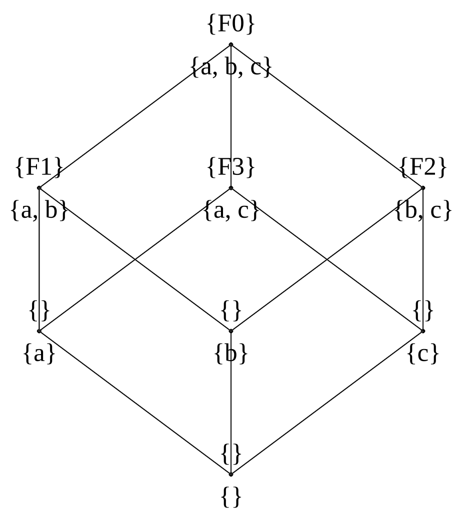

b.



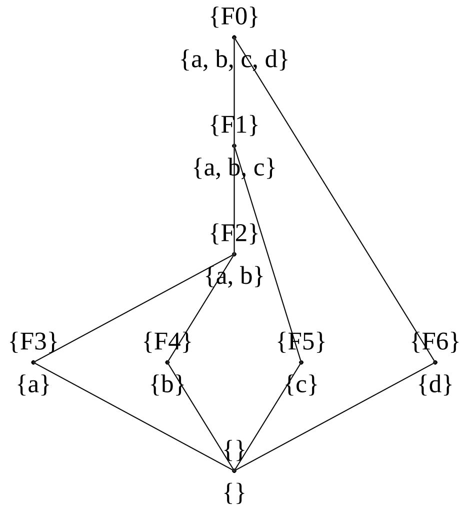

c.




d.

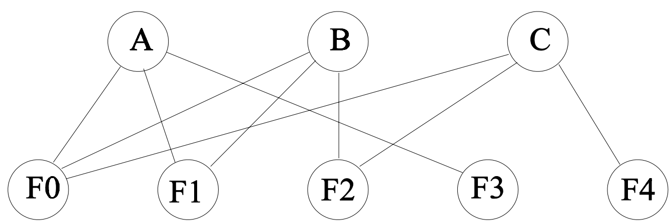


e.




The network of spreading activation models of these inventories are based on the Dell (1986) spreading activation model of phonological encoding. Each network has a set of nodes for the objects A-D, and a set of nodes for the features F0-F6. If an object possesses a feature, the nodes for the object and feature are bidirectionally connected. A constant proportion of the activation of each node spreads to every other node to which it is connected in each time step. In addition, activation decays by a constant proportion in each time step. I introduce the Dell (1986) model of phonological encoding in chapter 4, and I defer the details of the spreading activation model until then. Figure 3.5 shows the activation network for Inventory c.

**Figure 3.5**: Network of activation model of Inventory c.




Similarity in the network model is determined by assigning an arbitrary amount of activation to one object (e.g. to A) and then letting the system spread activation for a fixed number of time units. The activation of the other objects (e.g. B, C, and D) is a measure of how similar they are to the originally activated object. In my simulations, I use 100 units of initial activation, and let activation spread for six time steps, following the general characteristics of the Dell (1986) model.

In the simple and conjunctive features model, similarity is computed by the following formula:[10]

(27) Similarity = (number of single feature matches) + (number of feature pair matches) + (number of feature triplet matches) + (number of feature quadruplet matches)

In Inventories a-e, there are never more than a quadruplet of feature matches, so no higher order terms are given. For example, consider the similarity of A to B in Inventory e. A has features [+F1], [+F2], [+F3], [+F4]. B has features [+F1], [+F2], [+F3], [+F5]. There are three single feature matches: [+F1], [+F2], [+F3]. There are three feature pair matches: [+F1]&[+F2], [+F1]&[+F3], [+F2]&[+F3]. There is one feature triplet match: [+F1]&[+F2]&[+F3]. The similarity of A to B is 3 + 3 + 1 = 7.

Table 3.7 compares the similarity of A to A, B, C, and D in the natural classes model, spreading activation model, and simple and conjunctive features model. Each column gives values for a different inventory. Each row shows a different similarity model.

[10] In the implementation of the simple and conjunctive features model in Goldstone (1994a), each level of feature matching was given a coefficient so that the model could be used in a regression model of experimental data. Since I am comparing gross characteristics of the model, I treat all levels of feature matching equivalently.

**Table 3.7**: Comparison of similarity of Inventories a-e in three models.

| Model | Inventory | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Lattice | a | | b | | c | | d | | e | |
| A to A | 1.00 | | 1.00 | | 1.00 | | 1.00 | | 1.00 | |
| A to B | 0.33 | | 0.50 | | 0.40 | | 0.33 | | 0.60 | |
| A to C | 0.33 | | 0.25 | | 0.20 | | 0.33 | | 0.40 | |
| A to D | | | | | | | | | 0.20 | |
| Activation | a | norm a | b | norm b | c | norm c | d | norm d | e | norm e |
| A to A | 3.12 | 1.00 | 4.37 | 1.00 | 4.37 | 1.00 | 4.49 | 1.00 | 5.95 | 1.00 |
| A to B | 1.13 | 0.36 | 2.38 | 0.55 | 2.42 | 0.55 | 2.51 | 0.56 | 3.97 | 0.67 |
| A to C | 1.13 | 0.36 | 1.21 | 0.28 | 1.34 | 0.31 | 2.51 | 0.56 | 2.63 | 0.44 |
| A to D | | | | | | | | | 1.34 | 0.22 |
| SCFM | a | norm a | b | norm b | c | norm c | d | norm d | e | norm e |
| A to A | 3 | 1.00 | 7 | 1.00 | 7 | 1.00 | 7 | 1.00 | 15 | 1.00 |
| A to B | 1 | 0.33 | 3 | 0.43 | 3 | 0.43 | 3 | 0.43 | 7 | 0.47 |
| A to C | 1 | 0.33 | 1 | 0.14 | 1 | 0.14 | 3 | 0.43 | 3 | 0.20 |
| A to D | | | | | | | | | 1 | 0.07 |

The upper row shows natural class similarity. The middle row shows the amount of mutual activation in the left side of the columns, and a normalized activation on the right. Activation was normalized for each series by dividing all values by the activation from A to A for that inventory. The bottom row shows similarity and normalized similarity in the simple and conjunctive features model.

Examining the Inventories a-d, where the structure of the inventory develops from no natural classes to all possible natural classes, we see that all models predict the same relative levels of similarity between objects for Inventories a and b. In Inventories c and d, which have a greater amount of cross-classification, the natural classes model gives relatively lower similarity of A to B and A to C (compared to the similarity of A to A) than the activation model. The simple and conjunctive features model agrees better with the natural classes model for Inventory c, but it agrees better with the activation model for Inventory d. The natural classes model diverges from the other models in the case where there is a large amount of cross-classification in the inventory. Examining Inventories a, b, and e reveals that the natural classes model and the spreading activation model are in close agreement, and neither is too distant from the simple and conjunctive features model.

Overall, the three models generally agree. The major difference arose in the case where there was a great deal of cross-classification in the inventory. The difference arises because the natural classes model, unlike the spreading activation model or the simple and conjunctive features model, takes into account the *differences* as well as the similarities. When there are many cross-classifying features, there are many non-shared natural classes, which reduces similarity. There is evidence in favor of the way the natural classes model handles cross-classification, mentioned above. Pierrehumbert (1993) noted that the cooccurrence restriction in Arabic, OCP-Place, is stronger for {b, f, m} than it is for {d, s, n}. Presumably all the features involved in contrasting {b, f, m} and {d, s, n} are the same. The difference is that there are many more coronals in the Arabic inventory, but {b, f, m} are the only labials. The existence of additional cross-classification in the coronals reduces the similarity among {d, s, n}, and thus the OCP-Place constraint is weaker among the coronals. I discuss the Arabic case in detail in chapter 5.

I present additional data from phonological speech errors in English in chapter 4 which show that the effect of redundancy on the natural classes model of similarity is supported. The spreading activation model and simple and conjunctive features model do not differentiate the effect of features based on redundancy, and are thus unable to model these effects. To the extent that the models in this thesis can be represented by a spreading activation network, we do not need to posit that native speakers know the equations used in the mathematical models. Rather, the models are closed form approximations of a connectionist model of competence. It is thus desirable to find a connectionist model with behavior similar to the natural classes model.

I discuss in several places in this thesis how the quantitative models might be implemented in a connectionist/spreading activation network. Here I address the differences between the natural classes model and the spreading activation model found so far. We saw above that the models diverged when there was a great deal of cross-classification. I believe that a network which has inhibition, in addition to activation, will behave very much like the natural classes model. In a model with inhibition, the excessive mutual activation that was found in the network with a great deal of cross-classification is damped. If there is no cross-classification, then inhibition only plays a small role, and the revised model behaves much as it did in the simulation above.

Though I did not explicitly test it here, I am certain that the natural classes model and the spreading activation model would also differ in their treatment of redundancy. The natural classes model gives lesser weight to redundant features in determining similarity. There is no differentiation between features in the spreading activation model which would reduce the effect of redundant features. I believe that the redundancy effects can be represented in a more complex connectionist network. A true connectionist network, unlike a simple spreading activation model, is sensitive to the frequency of individual features and the frequency of their cooccurrence. To the extent that two features have a reciprocal dependency, such that they are predictably found together in many cases, the two features behave more like a single property than a conjunction of independent properties in the network. As a single property, they influence activation less than two independent properties.

As a second test of the relatedness between activation in a Dell (1986) spreading activation model of phonological encoding and the natural classes model, we can compare the similarity of consonant pairs in English (from Table 3.6) with the activation in a larger spreading activation model which includes nodes for all of the consonants in English, and all of the features from chapter 2. The spreading activation models were implemented exactly as above, just with more segment and feature nodes. To get a representative sample of consonant types, I ran four network simulations, one for each of the consonants {s, g, m, l}.

Figure 3.6 shows natural class similarity plotted against activation from the spreading activation model for each pairing of the consonants {s, g, m, l} with every other consonant. In all four cases, there is a clear logarithmic relationship between natural class similarity and activation. For each of the four simulations, I fit a linear regression of the form:

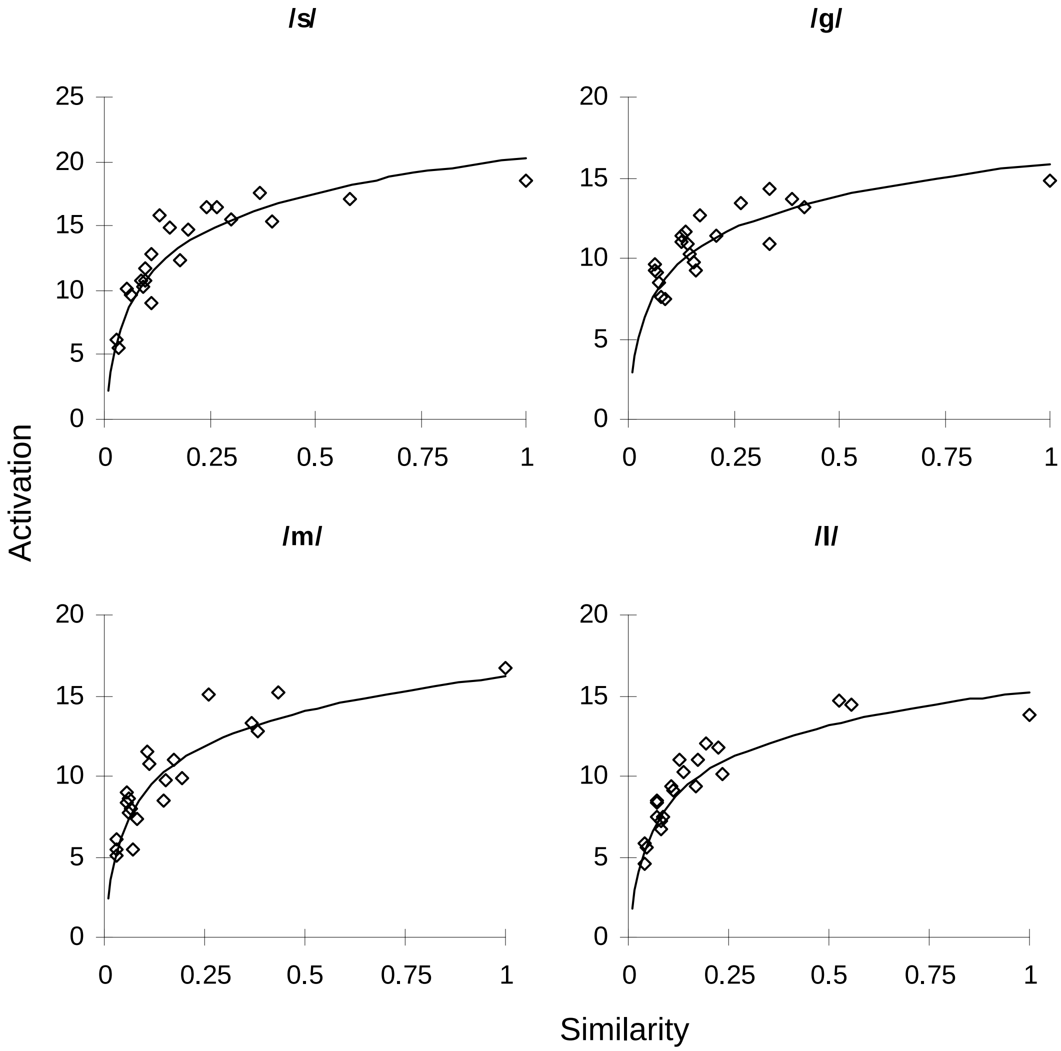
(28) Activation = A + B log(Similarity)

The regression curves are shown in the figures. Results of the regression analysis are shown in Table 3.8. The regression models predict similarity reasonably well. The best model, for /s/, fit with $R^2 = 0.82$, and the worst model, for /g/, fit with $R^2 = 0.64$. The $R^2$ statistic shows what fraction of the variation in the data is accounted for by each model. Models with a higher $R^2$ account for more of the variation in the data.

**Table 3.8**: Regression models of activation as a function of log(similarity).

| | /s/ | /g/ | /m/ | /l/ |
|---|---|---|---|---|
| $R^2$ | 0.82 | 0.64 | 0.76 | 0.77 |
| A | 20.3 | 15.9 | 16.2 | 15.2 |
| B | -9.4 | -6.7 | -7.2 | -7.0 |

**Figure 3.6**: The relation between similarity and activation for four consonants.




These simulations, based on a real redundancy hierarchy and spreading activation model confirm that there is a close relationship between natural class similarity and activation. Thus, keeping in mind the differences discussed above, I conclude that the natural classes model can serve as a closed form approximation of the behavior of a connectionist network. Implementing a complete connectionist model, which models effects of cross-classification and redundancy, is left as an open research problem.

# CHAPTER 4
# English Speech Errors as a Test of the Metric

In this chapter, I analyze two confusion matrices of segmental speech errors between consonants in English. I first present a measure of the relative degree of confusability of two segments, based on the ratio of the number of occurrences of an error between two segments divided by the number of occurrences which would be expected if errors were to occur between segments at random. I then compare the ratio of observed versus expected errors (O/E) between two segments to the similarity of the two segments, computed as in chapter 3. Similarity as computed in the natural classes model is a good predictor of the speech error rate. More similar segments show a greater error rate, and less similar segments show a reduced error rate.

I compare the natural class similarity model with feature based similarity models, and show that the natural classes model provides a better prediction of error rate. Further, the natural classes model predicts differences in error rate between segments which would be predicted to have the same error rate in a feature similarity model.

Section 4.1 contains an introduction to phonological speech error data and presents the relative measure of error rate I use in this thesis: the ratio of the number of observed errors to the number expected at random (O/E). In section 4.2, I analyze the confusion matrix of errors from Stemberger (1991a), and show that the natural classes similarity model is a very good predictor of error rate between consonant pairs. In section 4.3, I replicate the results of section 4.2 with a second corpus of phonological speech errors from the MIT-Arizona corpus. I present additional evidence for the natural classes model by demonstrating that partially redundant features have a lesser effect on similarity than non-redundant ones in section 4.4. In section 4.5, I describe the Dell (1986) spreading activation model of phonological encoding, and show that, assuming there is log-normally distributed variation in activation levels, the model predicts error rates much like those found in the natural error corpus. I finish the chapter with a review of the results achieved thus far, and a discussion of the relative merits of algebraic versus dynamic models of linguistic behavior.

## *4.1 Data and Measure*

A speech error is a spontaneous unintentional deviation from the intended utterance. Phonological speech errors are errors which are based on phonological shape. Examples of phonological speech errors are given in (29), the error is presented along with the intended target in parentheses (errors taken from Fromkin (1971)).

(29) a. correcting (collecting)
b. a hunk of jeep (a heap of junk)
c. plan the seats (plant the seeds)

In this thesis I will be examining single segment errors between two consonants. The example in (29a) is such an error. Each single segment error has a TARGET, the intended phoneme, and an INTRUSION, the erroneous phoneme which is actually produced. For example, in (29a) the target is /l/, and the intrusion is /r/.

Stemberger (1991a) presents a confusion matrix of single segment consonant errors caused by the interaction of one segment in the utterance plan with another. Examples in (30) are from Stemberger (1991a). Interactions can involve the ANTICIPATION of one segment for another (30a), the PERSEVERATION of a previously uttered segment (30b), or an EXCHANGE of positions by two segments (30c).

(30) a. setting ... getting such bad luck
b. about six seat (about six feet)
c. like box (bike locks)

The errors in (30) provide evidence that sentence production involves some degree of advance planning in phonological production. For an anticipation error to occur, the intruding segment must be accessible at the time of the error, even though that particular segment is not due to be immediately produced. See Levelt (1989) for a review of error evidence in a model of language production.

Table 4.1 shows the distribution of a total of 1273 single segment interaction errors published in Stemberger (1991a). The target segment is indicated in the left column. The intrusion segment is given across the top row. Informal inspection indicates that many errors occur between similar segments, and few occur between dissimilar segments. However, the absolute number of errors is deceiving, as some segments are much more frequent in speech, and in speech errors, than others. A measure of error rate which factors out frequency effects is needed.

Following Pierrehumbert (1993), I use a measure of error rate which compares the number of errors which are observed to the number which would be expected if consonants were to substitute for one another at random. Random chance is determined using the actual frequencies of segments as targets and intrusions in the error corpus being studied. For example, /p/ is a target in 84 errors, so the probability of /p/ as a target is 0.066 = 84/1273. Similarly, /f/ is an intrusion in 69 errors, so the probability of /f/ as an intrusion is 0.054 = 69/1273. The relative probability of a /p/-/f/ error is thus $p(\mathrm{p,f}) = 0.00358 = 0.066 \times 0.054$. The expected number of errors for each pair is:

(31) $$\text{Expected}(x,y) = \frac{p(x,y)}{\sum p(x_i,y_j)} \cdot \text{Total errors}$$

Note that $\sum p(x_i,y_j)$ is less than one, as I assume that $p(x,x) = 0$. Errors between identical segments, if they do occur, cannot be detected, so a certain amount of the marginal probability is lost from the total. Expected errors are distributed by frequency over all non-identical pairs. In other words, expected counts of consonants interacting with themselves are set to zero and the other expected values are increased to insure that the total number of errors is correct.

| | p | b | f | v | m | t | d | θ | ð | s | z | ʃ | ʒ | tʃ | dʒ | k | g | ŋ | l | r | n | w | y | h | Total |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| p | | 5 | 25 | 0 | 4 | 22 | 0 | 1 | 0 | 3 | 0 | 0 | 0 | 0 | 0 | 21 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 3 | 84 |
| b | 9 | | 3 | 4 | 7 | 2 | 11 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 1 | 10 | 0 | 5 | 2 | 1 | 6 | 0 | 1 | 64 |
| f | 8 | 1 | | 1 | 1 | 3 | 0 | 5 | 0 | 22 | 0 | 1 | 0 | 3 | 0 | 4 | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 9 | 60 |
| v | 3 | 6 | 2 | | 2 | 2 | 1 | 0 | 1 | 0 | 2 | 0 | 0 | 0 | 1 | 0 | 3 | 0 | 2 | 1 | 0 | 0 | 0 | 0 | 26 |
| m | 4 | 7 | 0 | 4 | | 0 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 2 | 0 | 1 | 3 | 3 | 19 | 8 | 0 | 3 | 56 |
| t | 22 | 2 | 1 | 0 | 0 | | 6 | 7 | 1 | 13 | 3 | 2 | 0 | 15 | 1 | 42 | 0 | 0 | 3 | 1 | 6 | 0 | 0 | 8 | 133 |
| d | 0 | 5 | 2 | 1 | 0 | 11 | | 0 | 0 | 3 | 5 | 2 | 0 | 0 | 6 | 0 | 20 | 0 | 5 | 2 | 9 | 1 | 1 | 1 | 74 |
| θ | 0 | 1 | 3 | 0 | 0 | 2 | 0 | | 0 | 16 | 0 | 2 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 2 | 28 |
| ð | 0 | 0 | 0 | 1 | 0 | 0 | 4 | 0 | | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 2 | 1 | 0 | 1 | 0 | 1 | 12 |
| s | 3 | 0 | 16 | 0 | 1 | 11 | 1 | 29 | 0 | | 0 | 58 | 0 | 6 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 7 | 133 |
| z | 1 | 0 | 0 | 6 | 0 | 3 | 1 | 0 | 1 | 1 | | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 1 | 1 | 1 | 0 | 0 | 0 | 18 |
| ʃ | 0 | 0 | 1 | 0 | 0 | 2 | 0 | 1 | 1 | 33 | 0 | | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 41 |
| ʒ | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| tʃ | 1 | 1 | 0 | 0 | 0 | 9 | 0 | 0 | 0 | 7 | 1 | 1 | 0 | | 1 | 4 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 26 |
| dʒ | 1 | 1 | 1 | 0 | 0 | 1 | 9 | 0 | 0 | 3 | 1 | 0 | 0 | 0 | | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 18 |
| k | 21 | 1 | 7 | 1 | 0 | 28 | 1 | 2 | 0 | 11 | 0 | 0 | 0 | 9 | 0 | | 8 | 1 | 1 | 1 | 0 | 0 | 0 | 5 | 97 |
| g | 0 | 8 | 1 | 0 | 0 | 1 | 2 | 0 | 0 | 1 | 2 | 0 | 0 | 0 | 2 | 5 | | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 24 |
| ŋ | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | | 0 | 0 | 3 | 0 | 0 | 0 | 3 |
| l | 0 | 1 | 1 | 1 | 3 | 3 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 2 | 0 | 0 | | 55 | 8 | 11 | 12 | 2 | 101 |
| r | 0 | 1 | 0 | 0 | 3 | 0 | 4 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 67 | | 1 | 20 | 2 | 0 | 98 |
| n | 0 | 1 | 0 | 0 | 23 | 5 | 7 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 1 | 5 | 0 | 1 | 16 | 4 | | 0 | 0 | 4 | 69 |
| w | 1 | 5 | 1 | 0 | 10 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 9 | 27 | 1 | | 4 | 1 | 60 |
| y | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 2 | 0 | 1 | 0 | 0 | 0 | 0 | 11 | 2 | 0 | 1 | | 0 | 18 |
| h | 0 | 0 | 5 | 1 | 0 | 2 | 1 | 1 | 0 | 6 | 0 | 1 | 0 | 1 | 0 | 10 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | | 30 |
| Total | 74 | 46 | 69 | 20 | 55 | 107 | 51 | 46 | 4 | 122 | 16 | 69 | 1 | 37 | 15 | 98 | 43 | 4 | 126 | 102 | 51 | 50 | 19 | 48 | 1273 |

**Table 4.1:** Confusion matrix of interaction errors from Stemberger (1991a).

The ratio of the number of observed errors to the number of expected errors (O/E) provides a measure of error rate which factors out the frequencies of targets and intrusions. The measure of O/E is a measure of the error rate between consonants independent of their frequency.

(32) $$O/E = \frac{\text{Observed}(x,y)}{\text{Expected}(x,y)}$$

*4.2 Similarity and Interaction Errors*

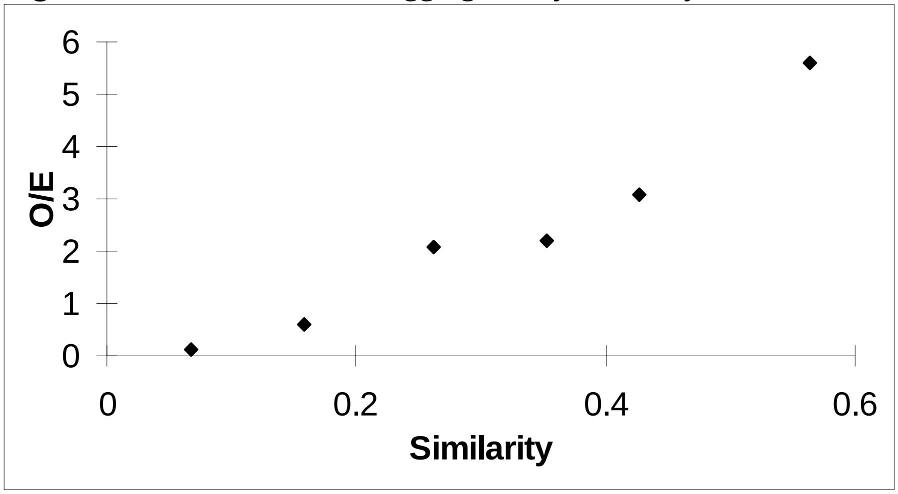
Table 4.2 presents aggregated total numbers of actual errors and expected errors for different levels of similarity in Stemberger's corpus of interaction errors. The data have been aggregated to more clearly reflect the relation between natural class similarity and error rate. The right-most column of Table 4.2 is the O/E measure of error rate based on the aggregate observed and expected totals.

**Table 4.2**: Interaction errors aggregated by natural class similarity.

| Similarity | Actual | Expected | O/E |
|---|---|---|---|
| 0-0.1 | 72 | 519.3 | 0.14 |
| 0.1-0.2 | 246 | 416.5 | 0.59 |
| 0.2-0.3 | 234 | 113.5 | 2.06 |
| 0.3-0.4 | 195 | 88.2 | 2.21 |
| 0.4-0.5 | 288 | 93.1 | 3.09 |
| 0.5-0.6 | 238 | 42.4 | 5.61 |

Figure 4.1 plots aggregate O/E against similarity. While the data have been aggregated, the groupings used in the aggregated similarity measure are still more sensitive than similarity measures previously used in speech error analyses. Most error analyses use three basic feature categories: place, manner, and voicing. Similarity is computed by counting the number of these basic features which are the same between two consonants (e.g. Nooteboom 1969, MacKay 1970, Shattuck-Hufnagel & Klatt 1979, van den Broeke & Goldstein 1980, Levitt & Healy 1985). For example, /p/ and /t/ share two features, manner and voicing, but differ by place. In this measure, there are only four relevant levels of similarity (0, 1, 2, or 3 shared features). The aggregated similarity in Table 4.2 differentiates seven levels of similarity, and the actual similarity measure is potentially continuous. The fact that error rate increases monotonically with the finer measure of similarity justifies the more precise measure of natural class similarity for modeling speech errors.

**Figure 4.1**: Interaction errors aggregated by similarity




The natural classes similarity metric provides a good prediction of error rate in Stemberger's corpus. However, the natural class model involves additional assumptions about the representation of segments, the effect of redundant features on similarity, and the synergistic effects of multiple feature matches. The natural classes model can be compared to three simpler models of speech errors which we might prefer to use for parsimony. The first I call the FREQUENCY MODEL, which assumes that similarity is not a factor in errors, and the predicted number of errors is equal to the number expected as computed above. This model has been shown to have a poor fit by many researchers, but it is included as a baseline to show how much of the error rate is accounted for solely by frequency. The second model is the SIMPLE FEATURE MODEL. In this model, similarity is based on simple place, manner, and voicing contrasts, as mentioned in the previous paragraph. This model has only four distinct similarity values. The third model is the COMPLEX FEATURE MODEL. This model is based on the same features used in the computation of similarity over natural classes, but instead similarity is computed based on shared and non-shared features (as in Pierrehumbert 1993, discussed in chapter 2). Each successive model requires additional assumptions about the nature of speech errors and of phonological representations: The similarity models assume that similarity is a factor in speech errors. The complex feature model and the natural classes model assume that detailed feature representations are needed. The natural classes model assumes that synergistic feature matching and redundancy are also relevant to similarity.

The similarity models can be compared based on their ability to predict the data in Stemberger's confusion matrix. For these models, I used the following nonlinear regression equation.

(33) Observed = Expected × (A + B × Similarity)

This equation is roughly equivalent to a linear regression on O/E. The regression was performed on unaggregated data. The models attempt to predict the actual error rate for each consonant pair as target and intrusion. Pairs which have high actual error rate or high expected error rate have the greatest influence on the model fit. Table 4.3 shows model fits and parameters for the frequency model and the three similarity models.

**Table 4.3**: Four models of Stemberger's (1991a) interaction errors.

| Model | $R^2$ | A | B |
|---|---|---|---|
| Frequency Model | 0.17 | - | - |
| Simple Feature Model | 0.57 | -0.61 | 1.37 |
| Complex Feature Model | 0.57 | -0.56 | 6.06 |
| Natural Classes Model | 0.72 | -0.69 | 9.88 |

The frequency model has very poor fit. As was concluded by many others, similarity is a factor in error rate. The simple feature model provides significant improvement in $R^2$. Surprisingly, the complex feature model does no better. This suggests that 'primary features' (Stevens & Keyser 1989) might be the only relevant features for determining similarity. But, the model based on similarity computed with natural classes provides a much better fit to the data than either feature model. Thus, it is not that the additional (secondary) features are not relevant, but that feature similarity does not properly differentiate features by contrastiveness and redundancy the way the natural classes model does. The additional assumptions of the natural classes model are supported by the data.

### *4.3 Similarity and Noncontextual Errors*

The natural class based model of similarity also makes a good prediction of error rate for a second corpus of single segment consonant errors. The corpus consists of the so-called noncontextual errors in the MIT-Arizona corpus of speech errors.[11] These errors differ from the interaction errors in that there was no apparent intrusion in the utterance, which makes these errors rare. This corpus provides independent replication of the results in the previous section.

These errors are spontaneous errors from naturally occurring speech which were collected opportunistically over the course of several years. The errors were recorded orthographically, with phonemic transcription included where necessary. In addition, the error recorders indicated the target word either by including the speaker's own correction or by noting their impression based on the discourse context. There are a total of 905 noncontextual errors in this corpus. Of these, 517 are single segment consonant errors used in the analysis in this section. Table 4.4 shows the confusion matrix for this error corpus.

Table 4.5 presents aggregated total numbers of actual errors and expected errors for different levels of similarity in the noncontextual error corpus, along with the aggregate O/E. Expected error rates were computed in the same manner as for the interaction errors in the

---

[11] I would like to take this opportunity to express my thanks to Stefanie Shattuck-Hufnagel for providing me with this corpus of noncontextual errors.

**Table 4.4**: Confusion matrix of noncontextual errors from the MIT-Arizona corpus.

| | p | b | f | v | m | t | d | θ | ð | s | z | ʃ | ʒ | tʃ | dʒ | k | g | ŋ | l | r | n | w | y | h | Total |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| p | | 4 | 9 | 0 | 3 | 8 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 4 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 28 |
| b | 6 | | 0 | 1 | 11 | 0 | 5 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 5 | 0 | 0 | 0 | 1 | 1 | 0 | 1 | 33 |
| f | 10 | 0 | | 7 | 1 | 2 | 0 | 1 | 0 | 18 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 1 | 42 |
| v | 1 | 6 | 5 | | 3 | 1 | 1 | 0 | 1 | 2 | 10 | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 33 |
| m | 2 | 1 | 0 | 0 | | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 12 | 3 | 0 | 0 | 21 |
| t | 8 | 0 | 0 | 0 | 0 | | 2 | 0 | 0 | 4 | 0 | 0 | 0 | 4 | 0 | 5 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 24 |
| d | 0 | 3 | 0 | 5 | 0 | 5 | | 0 | 0 | 0 | 2 | 0 | 0 | 1 | 11 | 0 | 3 | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 32 |
| θ | 0 | 0 | 3 | 0 | 0 | 2 | 0 | | 0 | 11 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 18 |
| ð | 0 | 0 | 0 | 2 | 0 | 0 | 0 | 0 | | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 3 |
| s | 0 | 0 | 8 | 0 | 1 | 5 | 1 | 8 | 2 | | 1 | 18 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 46 |
| z | 0 | 0 | 0 | 4 | 0 | 0 | 5 | 2 | 1 | 3 | | 0 | 7 | 0 | 5 | 1 | 0 | 0 | 1 | 0 | 2 | 0 | 0 | 0 | 31 |
| ʃ | 0 | 0 | 1 | 0 | 0 | 2 | 0 | 0 | 0 | 32 | 0 | | 0 | 4 | 0 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 41 |
| ʒ | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 4 | 0 | | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 7 |
| tʃ | 0 | 0 | 0 | 0 | 0 | 10 | 0 | 1 | 0 | 4 | 0 | 4 | 0 | | 1 | 3 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 23 |
| dʒ | 0 | 0 | 0 | 0 | 0 | 1 | 4 | 0 | 0 | 2 | 2 | 2 | 2 | 0 | | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 15 |
| k | 3 | 1 | 0 | 0 | 0 | 5 | 0 | 0 | 0 | 2 | 0 | 0 | 0 | 2 | 0 | | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 14 |
| g | 0 | 2 | 0 | 0 | 0 | 1 | 5 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 4 | | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 14 |
| ŋ | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| l | 0 | 1 | 1 | 0 | 1 | 0 | 5 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | | 13 | 1 | 2 | 8 | 0 | 32 |
| r | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 4 | | 1 | 6 | 0 | 0 | 13 |
| n | 0 | 0 | 0 | 1 | 7 | 1 | 3 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 3 | 0 | | 0 | 1 | 0 | 18 |
| w | 0 | 0 | 0 | 0 | 4 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 2 | 15 | 0 | | 0 | 0 | 21 |
| y | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 3 | 2 | 1 | 0 | | 0 | 7 |
| h | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | | 3 |
| Total | 30 | 18 | 28 | 20 | 33 | 45 | 32 | 12 | 5 | 82 | 22 | 24 | 9 | 11 | 19 | 22 | 12 | 1 | 15 | 34 | 20 | 13 | 10 | 2 | 519 |

previous section. These data show a very similar pattern to the interaction error data. Figure 4.2 shows similarity against O/E for the noncontextual segmental errors.

**Table 4.5**: Noncontextual errors aggregated by natural class similarity.

| Similarity | Actual | Expected | O/E |
|---|---|---|---|
| 0-0.1 | 26 | 197.1 | 0.13 |
| 0.1-0.2 | 98 | 178.0 | 0.55 |
| 0.2-0.3 | 100 | 57.1 | 1.75 |
| 0.3-0.4 | 82 | 40.6 | 2.02 |
| 0.4-0.5 | 131 | 31.8 | 4.12 |
| 0.5-0.6 | 82 | 14.4 | 5.68 |

**Figure 4.2**: Noncontextual errors by similarity

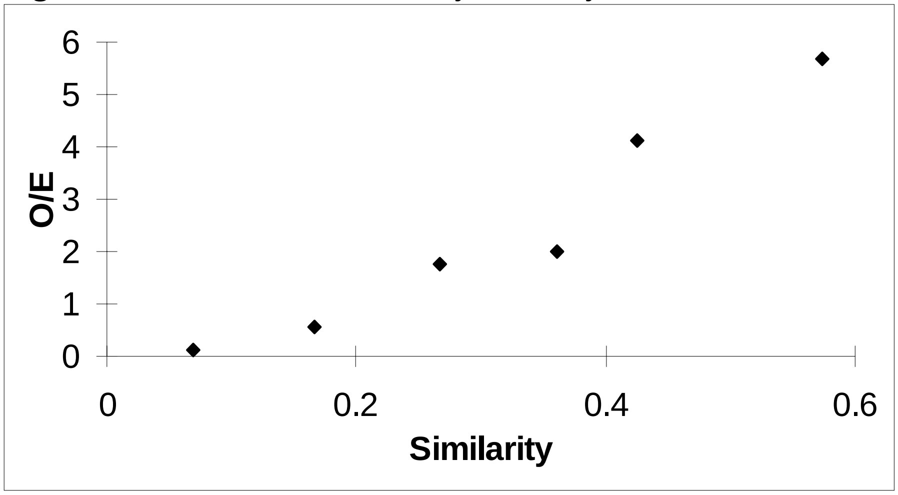



Table 4.6 compares the same four models of speech errors as in the previous section on the noncontextual error corpus. Once again, the evidence for a similarity effect is quite strong. In addition, the similarity model based on natural classes provides the best fit to the data. The noncontextual error data closely replicate the data from the previous section, providing additional evidence in favor of the natural classes model of similarity.

**Table 4.6**: Four models of the noncontextual error corpus.

| Model | $R^2$ | A | B |
|---|---|---|---|
| Frequency Model | 0.18 | - | - |
| Simple Feature Model | 0.54 | -0.71 | 1.36 |
| Complex Feature Model | 0.49 | -0.54 | 5.32 |
| Natural Classes Model | 0.62 | -0.57 | 8.96 |

### *4.4 Redundancy in Speech Errors*

Recall from chapter 3 that the natural classes similarity model predicts that redundant features have less of an effect on similarity than non-redundant features. Presumably, this fact accounts in part for why the natural classes similarity model performs much better than the complex feature model in predicting error rate. In this section, I present a specific case of redundancy effects.

Isolating specific cases where there is a minimal contrast between redundant and non-redundant features is difficult. For example, [+voice] is redundant in sonorants, but the inventory of sonorants in English has a very different natural class structure than the inventory of obstruents. There is a converse implication, not often discussed, which is that [+obstruent] is redundant in voiceless consonants. In order to compare a minimal difference based only on the redundancy of [+obstruent], I compare error rates among the voiced and voiceless continuants in English ({v, ð, z, ʒ} and {f, θ, s, ʃ}, respectively). The redundancy hierarchy for voiceless continuants is comparable to the hierarchy for voiced continuants except for one difference: The voiceless continuants have a shared natural class for [+voiceless] and a shared natural class for [+obstruent], while the voiced continuants have a shared natural class for [+voiced] and a shared natural class for [+obstruent], as well as a shared natural class for [+voiced]&[+obstruent].[12] The natural classes model thus predicts a higher similarity between voiceless continuants and voiced continuants, and hence a higher error rate among voiced continuants than among voiceless continuants.

Table 4.7 shows the aggregate error rates between continuants in the interaction error corpus and the no-source error corpus. In both corpora, the error rate between voiced continuants

[12] Note that the nasals are sonorant stops, so this comparison cannot be extended straightforwardly to the voiced and voiceless stops. The voiced stops have additional natural classes, like {[+stop]} and {[+stop]&[+voiced]} which contain both obstruents and sonorants. The effect of these additional natural classes on similarity is unclear, though an informal survey of the similarity of English stops in Table 3.6 shows that the natural classes model gives higher similarity between voiceless stops (e.g. p-t) than between comparable voiced stops (e.g. b-d). Presumably, additional cross-classification with the nasals provides a number of non-shared natural classes, which reduces similarity.

is higher than the error rate between voiceless continuants, as predicted. Note also that the error rate between voiced and voiceless continuants is much lower, and that these error rates are roughly equivalent regardless of whether the voiced obstruent is the target or the intrusion.

**Table 4.7**: Error rates among English continuants.

| | Interaction | | | No source | | |
|---|---|---|---|---|---|---|
| Target-Intrusion | Observed | Expected | O/E | Observed | Expected | O/E |
| voiceless-voiceless | 187 | 46.6 | 4.02 | 100 | 31.2 | 3.21 |
| voiced-voiced | 12 | 1.2 | 9.96 | 30 | 5.5 | 5.41 |
| voiceless-voiced | 2 | 9.0 | 0.22 | 11 | 16.8 | 0.66 |
| voiced-voiceless | 4 | 14.3 | 0.28 | 12 | 22.0 | 0.55 |
| Total | 205 | | | 153 | | |

It might be objected that the voiceless continuants {f, θ, s, ʃ} are much more frequent than the voiced continuants {v, ð, z, ʒ}, and that is why the error rate is lower for voiceless continuants. For example, it has been shown that high frequency words are produced with fewer errors than low frequency words (Dell 1988). However, this does not appear to be the cause of the difference in the error rates. Table 4.8 gives a breakdown of the error rate for each continuant as a target and as an intrusion in each corpus. The segments are presented in descending order of frequency. There is some indication that error rate is affected by frequency, but it is the higher frequency segments which appear to have higher error rate. It cannot be the case that low frequency predicts the high error rate between voiced continuants. Rather, it appears to be the differential effect of redundancy on similarity that accounts for the difference.[13]

[13] Note also that /s/ is by far the most frequent consonant among the continuants. Since /s/ has special behavior elsewhere, for example in sonority reversals in clusters, it may be special in speech errors in some way (cf. Newman, Sawusch, & Luce 1996). If /s/ is removed from the analysis, the results are unaffected.

**Table 4.8**: Error rate among continuants by consonant.

| | Interaction | | | Noncontextual | | |
|---|---|---|---|---|---|---|
| Target | Observed | Expected | O/E | Observed | Expected | O/E |
| /s/ | 103 | 25.0 | 4.12 | 37 | 11.2 | 3.29 |
| /ʃ/ | 36 | 9.5 | 3.78 | 33 | 14.9 | 2.22 |
| /f/ | 29 | 13.9 | 2.08 | 26 | 14.9 | 1.75 |
| /θ/ | 21 | 7.0 | 2.98 | 15 | 7.0 | 2.15 |
| /z/ | 9 | 5.0 | 1.81 | 17 | 11.4 | 1.50 |
| /v/ | 5 | 7.1 | 0.70 | 18 | 12.2 | 1.47 |
| /ð/ | 2 | 3.4 | 0.58 | 3 | 1.2 | 2.49 |
| /ʒ/ | 0 | 0 | | 4 | 2.8 | 1.45 |
| Total | 205 | | | 153 | | |
| Intrusion | Observed | Expected | O/E | Observed | Expected | O/E |
| /s/ | 73 | 18.9 | 3.87 | 66 | 29.2 | 2.26 |
| /ʃ/ | 61 | 16.0 | 3.82 | 18 | 8.8 | 2.05 |
| /f/ | 35 | 11.2 | 3.14 | 11 | 5.0 | 2.22 |
| /θ/ | 22 | 14.9 | 1.48 | 17 | 10.2 | 1.67 |
| /z/ | 8 | 4.9 | 1.64 | 13 | 7.7 | 1.70 |
| /v/ | 2 | 4.0 | 0.50 | 17 | 8.5 | 2.00 |
| /ð/ | 1 | 0.3 | 3.76 | 7 | 3.9 | 1.79 |
| /ʒ/ | 3 | 1.0 | 2.93 | 4 | 2.2 | 1.80 |
| Total | 205 | | | 153 | | |

### *4.5 A Spreading Activation Model of Speech Errors*

Dell (1986) proposes a processing model for phonological encoding using spreading activation. This model is one of a number of models of speech production which use a network of dynamically activated nodes to represent the lexicon, syllables, phonemes, and features (e.g. Dell & Reich 1980, 1981; Dell, Juliano, & Govindjee 1993; Levelt 1989; MacKay 1982; Stemberger 1982). The Dell (1986) model uses a slots-and-fillers design (Dell & Reich 1981; Fromkin 1971; MacKay 1972, 1982; Sevald & Dell 1994; Shattuck-Hufnagel 1979; Stemberger 1982) which posits that a phonological skeleton is constructed and then filled in with segments. The activation model determines which segments are selected to fill the slots in the skeleton. A schematic of the architecture of the Dell (1986) model of phonological processing is shown in

Figure 4.3.

The Dell (1986) model employs a number of phonological levels: morphemes, syllables, rimes, clusters, phonemes, and features. Phonological encoding is modeled as follows. A morpheme is marked to be encoded by allotting it a certain number of abstract activation units. Activation then spreads throughout the model, activating syllables, rimes, phonemes, and features associated with the morpheme. All connections are two-way, so that activation spreads both upwards and downwards. The network activates the appropriate phonemes needed to produce the morpheme, but other related morphemes, syllables, phonemes, and features are indirectly activated as well, through connections which are once removed from the original morpheme. The network could conceivably be extended to include semantic and syntactic encoding, with two way connections between the additional nodes containing syntactic and semantic units.

The system operates over discrete time units. The activation of a node $j$ at time unit $t_i$, denoted $A(j, t_i)$ is given by equation (34).

(34) $A(j, t_i) = [A(j, t_{i-1}) + \sum_{k=1 \text{ to } n}(p_k A(c_k, t_{i-1})](1-q)$

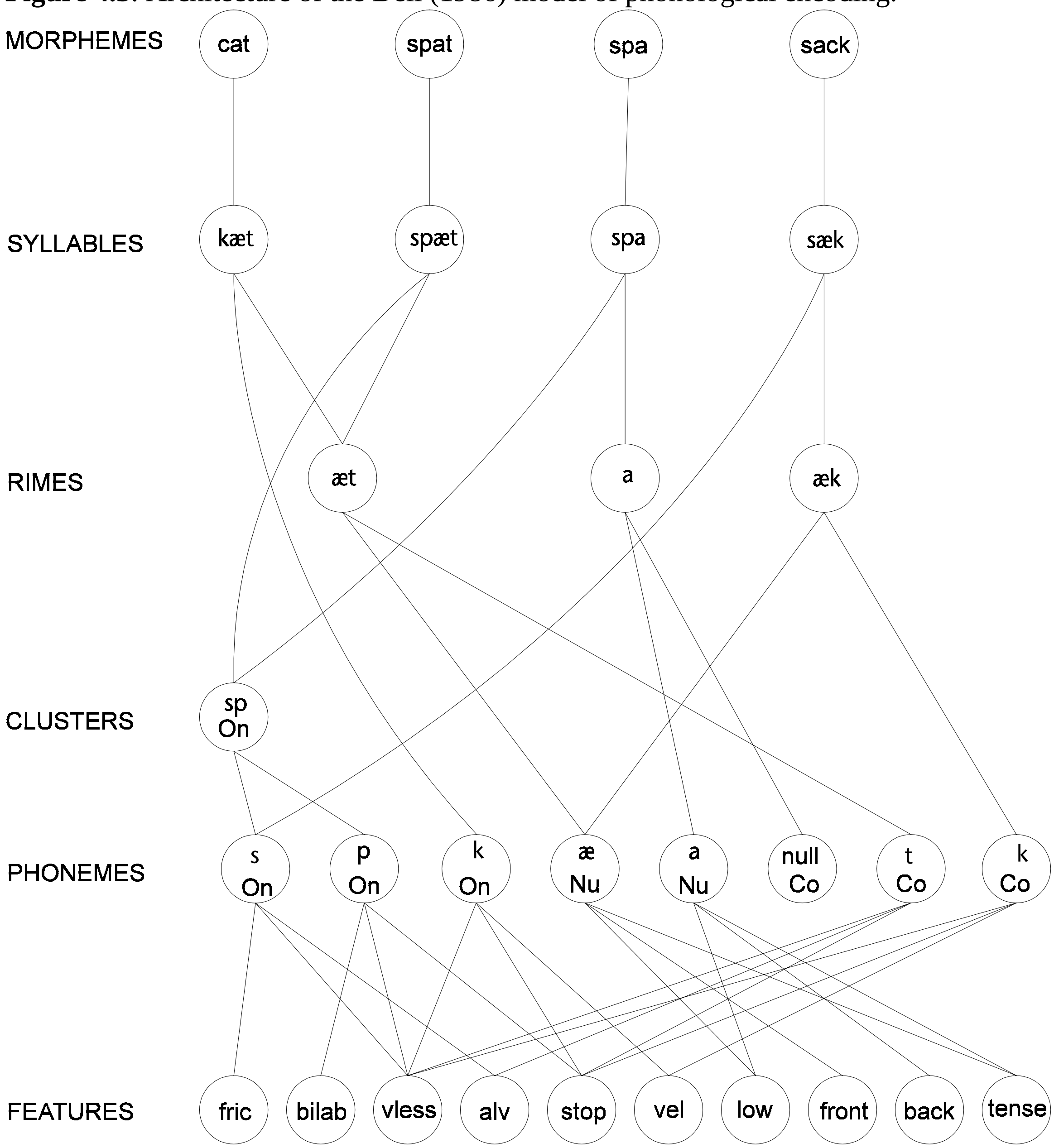
$A(j, t_{i-1})$ is the activation level of node $j$ at the previous time unit, $t_{i-1}$. The set $\{c_k, k = 1 \text{ to } n\}$ are all of the nodes which are directly connected to node $j$. The set $\{p_k, k = 1 \text{ to } n\}$ are weights attached to the connections between node $j$ and the other nodes. The value of $q$ controls the rate at which a node's activation decays over time.

The pattern of activation is transformed into a representation by assigning initial activation to a morpheme (theoretically due to higher level connections with semantic and syntactic nodes), waiting for a certain number of time units to pass, and then selecting the highest activated segments to fill the phonological frame for production. If a competing segment receives more activation than the target segment, it is selected and encoded, resulting in an error.

The Dell (1986) model of speech errors can account for a number of regularities found in speech error studies. In speech errors, there is a bias toward producing real word outcomes (Dell & Reich 1981). This model predicts the word bias by the two way connections between morpheme nodes and phonological content nodes. Combinations of phonemes which represent actual morphemes receive reinforcing activation from the morpheme nodes. Dell (1986) demonstrated that this bias is time dependent. Using an experimental technique to induce phonological speech errors, Dell (1986) found that subjects who had to respond under a 500 ms deadline had no tendency to produce word outcome errors in favor of non-word outcomes, but subjects with a 1000 ms deadline had a strong lexical effect. This model also accounts for the time bias. Interactive activation between morpheme nodes and phoneme nodes takes time units to travel up and down the network. In a short time deadline, fewer time units pass, and reinforcing activation does not have time to build up.

The model also has a bias toward producing phonotactically valid syllables, for the same reason. This is desirable, since speech errors rarely violate phonotactic constraints (Fromkin 1971). Finally, Dell (1984) demonstrates that there are segmental context effects on error rates. Syllables which share phonemes are more likely to slip with one another. This follows from the model's use of a single type node for each phonological unit. Two syllables which share vowels, for example /cæt/ and /sæk/ in Figure 4.3, are both connected to the same /æ/ node. Thus, when one is activated, activation spreads to the vowel, to the syllable, and then to the other phonemes within the syllable. As a result, a competing phoneme from a syllable with the same vowel receives more activation than it otherwise might.

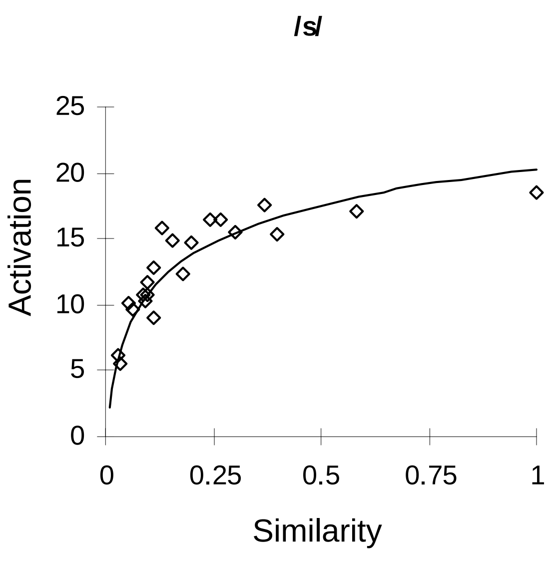
**Figure 4.3**: Architecture of the Dell (1986) model of phonological encoding.



The Dell (1986) model is comparable to the other network activation type models mentioned above for the encoding of a segment. Segments are connected to all of the feature nodes for the features of that segment. Since common feature nodes are shared, similar phonemes receive activation from the common feature nodes. As a result, similar segments should be more activated than dissimilar ones, and thus be more likely to be selected. In chapter 3, I implemented a sub-part of the Dell (1986) model. This model utilized only phoneme and feature nodes. The feature nodes are based on the same feature assignments used in the similarity computation, given in chapter 2. I used the same parameters for equation (34) that Dell (1986) did: I used a decay rate of $q = 0.6$, and $p = 0.3$ for all node connection weights. I started the network by assigning 100 units of abstract activation to a segment. Since this model has fewer interacting levels to spread activation, I ran the simulation for six time steps to allow a fair amount of interaction between the phonemes and features. Dell (1986) used three to five time units for his simulations.

Figure 4.4 shows the amount of abstract activation of each node in the network as a function of its similarity to /s/, repeated from chapter 3. As was noted in chapter 3, there is a clear correlation between similarity as computed with the natural classes model and spreading activation. Given the distribution of the points, the relation between similarity and activation is clearly logarithmic. The curve in Figure 4.4 is the best fit curve parameterized by a linear regression of log(Similarity) on activation (Activation = 20.3 + 9.4 log(Similarity), $R^2 = 0.82$).

**Figure 4.4**: Relation between natural class similarity and spreading activation for /s/.

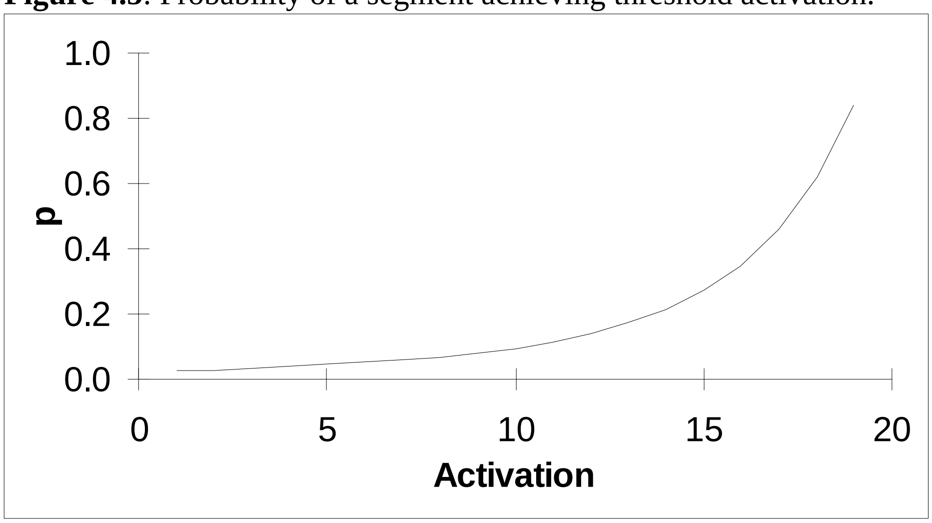



To determine the likelihood that one segment would be mis-selected for another in an activation network like this, we can assume that such mis-selections are the result of variation in the relative activation of nodes due to semantic, syntactic, and phonological context; physical factors like fatigue or drunkenness; and psychological factors such as anxiety. These effects can be abstractly modeled by applying a probability distribution to the activation model. Activation is a strictly positive random variable in the Dell (1986) model. For strictly positive random variables, it is standard to assume that variation is a log-normally distributed random variable (Devore 1977).

If an intrusion becomes more activated than the intended segment (/s/ in this case), an error occurs. Since there is also random variation in the activation level of the target, the system encounters a range of activation levels for the target segment. Suppose, for example, that variation in activation level is log-normally distributed with mean of one and standard deviation of one. The activation computed for /s/ by my mini-model after six time intervals was 18.6 units. With log-normal variation in activation, about 75% of instances of /s/ would have activation over 19. Setting the lower bound of the critical activation of an intruding segment arbitrarily at 19 units, consider how likely it is for a log-normally distributed level of activation to exceed this threshold. The probability that a segment of a particular base activation level is above this target due to random noise can be determined from the cumulative log-normal distribution.

Figure 4.5 shows the probability that random variation can increase the activation level of a segment over the critical level, as a function of the base activation of that segment due to featural similarity. If a segment's activation is very high, then it might be more activated than the intended /s/, and be produced instead. Thus, the function in Figure 4.5 should directly reflect the error rate between segments. Since activation is a function of log(Similarity), we can approximate the error rate of segments based on their relative activation via the error rate of segments based on their similarity.

**Figure 4.5**: Probability of a segment achieving threshold activation.




Using the regression model for the activation of /s/ based on similarity, we can convert between activation and similarity by the formula:

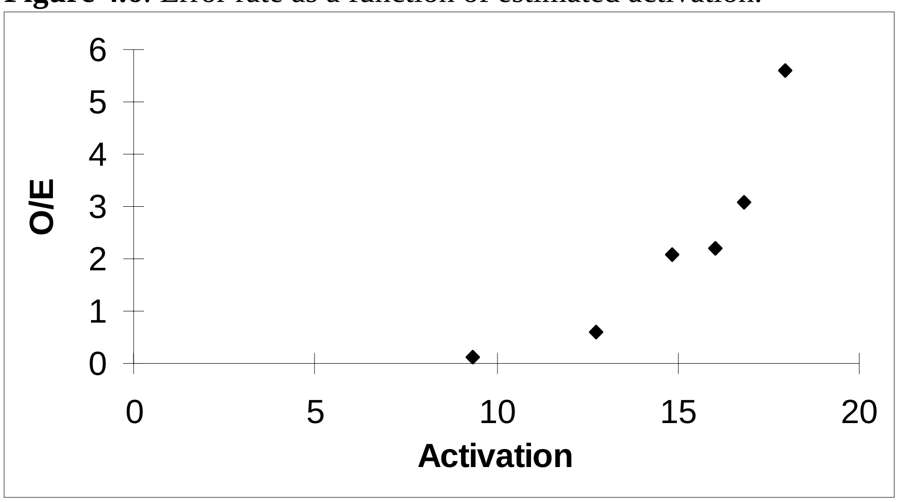
(35) Activation = 20.3 - 9.4 log(Similarity)

Table 4.9 shows mean similarity, log(Similarity), estimated activation from the regression model for /s/, and error rate for the aggregated interaction errors from section 3.2. Figure 4.6 plots O/E against activation for the data in Table 4.9. Figures 4.5 and 4.6 bear a striking resemblance, supporting the assumptions of the Dell (1986) model and log-normal variation in the system as an account for phonological speech error rate.

**Table 4.9**: Similarity, estimated activation, and error rate for interaction errors.

| Similarity | Log(sim) | Est. Activation | O/E |
|---|---|---|---|
| 0.07 | -1.17 | 9.35 | 0.14 |
| 0.16 | -0.80 | 12.77 | 0.59 |
| 0.26 | -0.58 | 14.84 | 2.06 |
| 0.35 | -0.45 | 16.04 | 2.21 |
| 0.43 | -0.37 | 16.82 | 3.09 |
| 0.56 | -0.25 | 17.96 | 5.61 |

**Figure 4.6**: Error rate as a function of estimated activation.




*4.6 Summary: Similarity, Activation, and Speech Errors*

The modeling here, in combination with the results of chapter 3, demonstrates that similarity computed using natural classes has a close connection to similarity as a function of activation in a connectionist network. They are related by a logarithmic transformation. The similarity model used in this thesis can thus substitute as a closed form estimate of the degree of interactive activation between segments. A static, closed form model has computational advantages over a spreading activation simulation. It is possible, for example, to apply algebraic techniques to a closed form equation, while an activation model can only be probed by repeated runs of the simulation.

Analogously, the spreading activation network provides a processing model for the similarity computed by the natural classes model. It thus can account for effects in speech errors, as well as in the phonotactic constraint I present in the next chapter. I extend the connection between similarity and activation further in chapters 6 and 7, where I argue that the effects of interference created by processing in a spreading activation model are found in both speech errors and in the phonotactic constraint OCP-Place. The computational model of similarity does not need to propose that the cognitive system selects objects, compares their sets of natural classes, and computes a ratio to determine similarity. Instead, similarity is emergent from the activation levels in the network of nodes and connections.

Earlier in this chapter, I showed that the similarity model of Frisch, Broe, & Pierrehumbert (1995), in which similarity is computed based on shared and non-shared natural classes, provides a very good model of consonantal segment error rates for both interaction and noncontextual speech errors. The examination of similarity effects in speech errors provides insight into the notion of similarity in general. Unlike previous treatments of similarity in cognitive psychology (e.g. Tversky 1977; Goldstone, Gentner, & Medin 1989; Nosofsky 1992) which use artificial features within controlled experimental situations, the model of speech errors presented here is based upon distinctive features which are physically based, naturalistic, and represent thoroughly entrenched knowledge. The set of phonological features is non-orthogonal, so redundancy relationships between features are relevant to the computation of similarity. The natural classes model predicts that languages with different segment inventories will have different error rates, due to differences in similarity over the redundancy hierarchy. This would be possible even between segments which are shared by both languages. This is clearly a testable hypothesis, and it is left here as a future research topic.

I found evidence for synergistic effects for multiple feature matches, and proposed that the correct computation of similarity is based not on individual features, but on the contrastive sets of objects that those features describe (the natural classes). This proposal incorporates redundancy relationships in a natural way. Totally redundant features define no new natural classes, and so do not contribute to increasing similarity. Partially redundant features increase similarity to a lesser degree, and non-redundant features are the most influential. These findings for similarity effects in phonology can also be applied in other cognitive domains, where naturalistic, non-orthogonal features are presumably commonplace.

# CHAPTER 5
# OCP-Place Effects in Arabic

In this chapter, I review the analysis of Arabic OCP-Place effects in Frisch, Broe, & Pierrehumbert (1995), henceforth FBP. FBP present a quantitative analysis of Arabic which shows a gradient dissimilarity constraint between homorganic consonants within the verbal roots. The FBP analysis of OCP-Place introduces the STOCHASTIC CONSTRAINT model of a gradient linguistic constraint. In the stochastic constraint model, 'acceptability' of a form is not an all-or-nothing relationship, as in current categorical phonological formalism. Acceptability is gradient, and reflected in the relative frequency of a form.

The Arabic data are presented in section 5.1. I review the standard autosegmental account of cooccurrence restrictions among the roots of Arabic in section 5.2. This account is based upon the claim that the Obligatory Contour Principle (OCP) prohibits adjacent identical elements in autosegmental representations (Leben 1973, Goldsmith 1979). The OCP has been applied to Arabic in two distinct ways using linguistic representations formalized in autosegmental phonology (Goldsmith 1979, 1994). It is applied as stated to mark as ill-formed any Arabic verbal root which contains a sequence repeated identical consonants (McCarthy 1979, 1986). In addition, it is applied within the feature geometric tier of place of articulation features (McCarthy 1988, 1994) to mark as ill-formed any root with adjacent identical place of articulation features.

In section 5.3, I review the arguments in Pierrehumbert (1993) and FBP which show that the standard categorical approach to OCP-Place effects in Arabic cannot account for the gradient nature of the data. There are systematic cases of consonant cooccurrence which reveal the non-categorical nature of OCP effects, and undermine the autosegmental model based on tier separation.

Pierrehumbert (1993) presented quantitative evidence of the influence of distance upon consonant cooccurrence, and pointed out that the autosegmental formulation is unable to account for these facts. She also proposed that OCP effects were based upon perceived similarity. In the similarity model, the OCP is seen as gradient rather than absolute. As such, it is capable of accounting for the interaction of similarity and distance effects. Pierrehumbert's model utilized contrastive underspecification. When examined more closely, FBP found contrastive underspecification to be both empirically and formally inadequate. FBP used the natural classes similarity metric to capture the effect of feature contrastiveness on similarity. The similarity model and the contrastive underspecification account are reviewed in detail in section 5.4. Additional evidence against underspecification in OCP effects is presented in chapter 10.

While FBP adopt the similarity-based approach of Pierrehumbert (1993), they improve on it by eliminating contrastive underspecification in favor of the similarity model based on natural classes discussed above. FBP present the first complete quantitative model of the OCP effects, called the stochastic constraint model, which predicts the rate of cooccurrence of different consonant combinations as a function of distance and similarity. This model, based upon the logistic function, provides a better fit to the Arabic data than the traditional account or the Pierrehumbert (1993) similarity model. In addition, the FBP model can be taken as a universal, grounded in the cognitive functions of similarity and categorization. The model can be parameterized as to the degree to which similarity and distance influence phonotactics. The stochastic constraint model provides an account of OCP-Place effects which has a basis in general cognitive functions, but allows for language specific variation on a limited set of dimensions. I discuss the stochastic constraint model in detail in section 5.6.

### *5.1 The Pattern of Root Morphemes in Arabic*

Arabic verbal root morphemes consist of a set of two to four consonants, with the canonical root containing three. Vowels are inserted between the consonants to make word forms, an example of a non-concatenative morphological system. For example, the verbal root *ktb* has among its word forms *katab* ‘to write’ and *kutib* ‘to be written’. In addition to non-concatenative vowel insertion, concatenative morphemes are used for many morphological processes. For example, *katab+a* ‘he wrote’ and *ma+katiib* ‘letters’.

Greenberg (1950) originally studied the overall statistical patterning of the triconsonantal verb root morphemes in Arabic, based on Lanes’s (1863) dictionary. The cooccurrence restrictions he describes apply only to the roots themselves, and not to any derived forms. Thus, the restrictions do not apply across concatenative morpheme boundaries. Greenberg begins his analysis with the observation that there are no roots which repeat the same consonant in first and second position. Thus, roots like **ddm* do not occur. Many roots are found with identical consonants in the second and third positions of the root. Examples include *mdd* ‘to stretch’ and *frr* ‘to flee’.

Greenberg notes that, more generally, Arabic consonants divide into groups of homorganic consonants that tend not to cooccur within the same root, apart from the pairs of identical consonants in the second and third position just mentioned. Pairs of homorganic but not identical consonants are not totally prohibited, but are found far less often than expected at random. Analogous cooccurrence constraints have been found in a number of Semitic languages (Bender & Fulass 1978, Buckley 1993, Greenberg 1950, Hayward & Hayward 1989, Koskinen 1964), as well as in English (Berkley 1994a, b), Javanese (Mester 1986), Ngbaka (Broe 1995), Russian (Padgett 1991), and other languages (Yip 1988).

Greenberg shows that the restriction against homorganic consonants applies not only to adjacent pairs of consonants, but also to nonadjacent consonants within the root. However, the cooccurrence constraint against homorganic consonants weakens over distance. The consonants of Arabic are shown in (36).

| (36) | Labial | Coronal | Emphatic | Velar | Uvular | Pharyngeal | Laryngeal |
|---|---|---|---|---|---|---|---|
| | | t | T | k | q | | ʔ |
| | b | d | D | g | | | |
| | f | θ, s | S | | χ | ħ | h |
| | | ð, z | Z | | ʁ | ʕ | |
| | | ʃ | | | | | |
| | | l, r | | | | | |
| | m | n | | | | | |

The complete set of distinctive features used to describe the Arabic consonants is given in section 5.3, where they are used to compute similarity. The particular features of the ‘guttural’ consonants (/χ/, /ʁ/, /ħ/, /ʕ/, /h/, /ʔ/) are based on McCarthy’s (1994) extensive review. McCarthy refers to these as the pharyngeal approximants. There is another set of consonants, called the ‘emphatics’, which deserve special mention. These are the coronal obstruents /T/, /D/, /S/, and /Z/. These consonants are similar to the familiar English consonants /t/, /d/, /s/, and /z/, but they contain a second vocal tract constriction in addition to one near the alveolar ridge, at the uvula (McCarthy 1994). Thus, these consonants have both [coronal] place of articulation and [dorso-guttural] place of articulation. These consonants constitute a special class when cooccurrence restrictions are discussed in detail, due to their two places of articulation.

The Arabic data in this thesis come from native and assimilated triliteral verbal roots from Wehr’s 1979 dictionary of modern Arabic (Cowan 1979). This is the same dictionary used in Pierrehumbert (1993) which is a later edition of the same dictionary used in McCarthy’s (1986, 1988, 1994) statistical analyses. Table 5.1 shows the distribution of adjacent consonant pairs in the Arabic triliteral roots ($C_1C_2$ and $C_2C_3$ consonant pairs). Table 5.2 shows the distribution of non-adjacent consonant pairs ($C_1C_3$ consonant pairs). There are 2676 triconsonantal roots, so there are 5352 adjacent pairs and 2676 non-adjacent pairs. Roots with repeated second and third consonants are excluded from these tables, because they are conceptually biliteral (McCarthy 1986), as discussed in section 5.2. In addition, roots with four surface consonants are excluded. Virtually all four consonant verbs are either reduplicated biliteral roots or conspicuously nonnative words (Pierrehumbert 1993). OCP-Place effects as discussed below have been demonstrated for native quadriliteral roots in Tigrinya (Buckley 1993), so the restriction of the analysis of Arabic to triliteral roots is a language particular idiosyncracy.

Examining Tables 5.1 and 5.2, the most obvious effect of the cooccurrence restriction divides the Arabic consonants into the following groups which have very low rates of cooccurrence (Greenberg 1950):

(37) a. Labials = {b, f, m}
b. Coronal Obstruents = {t, d, T, D, θ, ð, s, z, S, Z, ʃ}
c. Coronal Sonorants = {l, r, n}
d. Velars and Uvulars = {k, g, q, χ, ʁ}
e. Guttural Approximants = {χ, ʁ, ħ, ʕ, h, ʔ}

| | b | f | m | t | d | T | D | θ | ð | s | z | S | Z | ʃ | k | g | q | χ | ʁ | ħ | ʕ | h | ʔ | l | r | n | w | y |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| b | 0 | 0 | 0 | 10 | 13 | 19 | 4 | 7 | 6 | 15 | 6 | 4 | 0 | 8 | 11 | 8 | 13 | 7 | 7 | 11 | 12 | 14 | 9 | 24 | 31 | 9 | 19 | 10 |
| f | 0 | 0 | 0 | 16 | 10 | 6 | 8 | 3 | 1 | 11 | 5 | 5 | 3 | 8 | 5 | 5 | 20 | 5 | 2 | 10 | 11 | 5 | 5 | 19 | 31 | 11 | 20 | 12 |
| m | 0 | 0 | 0 | 8 | 13 | 9 | 6 | 3 | 3 | 18 | 10 | 7 | 1 | 13 | 6 | 9 | 10 | 6 | 4 | 14 | 19 | 6 | 6 | 27 | 30 | 15 | 12 | 19 |
| t | 6 | 7 | 8 | 0 | 3 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 2 | 4 | 3 | 5 | 1 | 2 | 4 | 5 | 1 | 2 | 11 | 17 | 6 | 7 | 8 |
| d | 10 | 11 | 21 | 0 | 1 | 0 | 0 | 2 | 0 | 8 | 0 | 0 | 0 | 4 | 0 | 6 | 5 | 4 | 5 | 10 | 13 | 9 | 4 | 16 | 22 | 14 | 21 | 12 |
| T | 11 | 14 | 13 | 0 | 1 | 0 | 0 | 0 | 0 | 5 | 0 | 0 | 0 | 5 | 0 | 0 | 4 | 1 | 1 | 6 | 7 | 4 | 3 | 15 | 19 | 8 | 16 | 7 |
| D | 10 | 3 | 7 | 0 | 3 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 3 | 0 | 2 | 4 | 6 | 5 | 2 | 2 | 5 | 12 | 4 | 8 | 9 |
| θ | 4 | 1 | 6 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 5 | 1 | 1 | 0 | 0 | 0 | 4 | 6 | 13 | 1 | 8 | 3 |
| ð | 7 | 4 | 3 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 3 | 0 | 2 | 2 | 0 | 0 | 5 | 2 | 3 | 7 | 17 | 2 | 8 | 7 |
| s | 16 | 14 | 21 | 2 | 13 | 8 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 7 | 9 | 9 | 12 | 1 | 10 | 9 | 7 | 6 | 21 | 22 | 9 | 21 | 13 |
| z | 11 | 7 | 14 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 4 | 6 | 10 | 2 | 5 | 8 | 13 | 6 | 5 | 11 | 15 | 9 | 11 | 14 |
| S | 14 | 14 | 10 | 1 | 14 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 5 | 1 | 3 | 6 | 6 | 4 | 1 | 17 | 13 | 5 | 14 | 10 |
| Z | 1 | 4 | 4 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 2 | 1 | 0 | 3 | 3 | 0 | 1 | 3 |
| ʃ | 11 | 11 | 16 | 4 | 9 | 9 | 0 | 0 | 2 | 0 | 1 | 0 | 2 | 0 | 8 | 7 | 9 | 5 | 4 | 10 | 11 | 5 | 4 | 6 | 23 | 9 | 17 | 18 |
| k | 13 | 11 | 13 | 8 | 10 | 0 | 1 | 4 | 1 | 15 | 5 | 1 | 1 | 6 | 0 | 0 | 0 | 0 | 0 | 2 | 5 | 5 | 5 | 14 | 22 | 11 | 10 | 11 |
| g | 12 | 4 | 15 | 0 | 10 | 0 | 0 | 2 | 6 | 10 | 12 | 0 | 0 | 3 | 1 | 0 | 0 | 0 | 0 | 7 | 12 | 8 | 6 | 21 | 24 | 18 | 19 | 4 |
| q | 18 | 10 | 19 | 5 | 12 | 12 | 5 | 0 | 6 | 6 | 2 | 14 | 2 | 9 | 0 | 0 | 0 | 0 | 0 | 8 | 13 | 4 | 2 | 17 | 21 | 11 | 15 | 18 |
| χ | 11 | 7 | 16 | 4 | 6 | 10 | 7 | 1 | 5 | 5 | 9 | 8 | 0 | 5 | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 13 | 22 | 8 | 11 | 9 |
| ʁ | 9 | 5 | 14 | 1 | 6 | 6 | 5 | 3 | 1 | 1 | 5 | 5 | 0 | 5 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 12 | 15 | 4 | 11 | 9 |
| ħ | 11 | 15 | 19 | 5 | 13 | 4 | 5 | 2 | 5 | 8 | 4 | 9 | 5 | 10 | 5 | 9 | 6 | 0 | 0 | 0 | 0 | 0 | 0 | 17 | 22 | 14 | 22 | 10 |
| ʕ | 17 | 8 | 12 | 10 | 11 | 12 | 5 | 5 | 5 | 10 | 9 | 8 | 4 | 7 | 6 | 10 | 12 | 0 | 0 | 0 | 0 | 2 | 0 | 15 | 21 | 12 | 15 | 14 |
| h | 14 | 6 | 18 | 4 | 19 | 4 | 4 | 1 | 3 | 1 | 9 | 2 | 1 | 5 | 4 | 12 | 5 | 0 | 0 | 0 | 0 | 0 | 0 | 16 | 23 | 7 | 17 | 20 |
| ʔ | 12 | 3 | 9 | 2 | 5 | 2 | 0 | 3 | 2 | 10 | 6 | 2 | 0 | 3 | 4 | 2 | 1 | 3 | 0 | 2 | 0 | 2 | 0 | 10 | 14 | 7 | 8 | 8 |
| l | 22 | 19 | 19 | 5 | 9 | 13 | 0 | 3 | 2 | 15 | 4 | 5 | 2 | 2 | 11 | 13 | 19 | 4 | 6 | 17 | 17 | 15 | 5 | 0 | 0 | 1 | 28 | 14 |
| r | 29 | 26 | 26 | 9 | 21 | 11 | 12 | 7 | 1 | 21 | 13 | 10 | 1 | 16 | 16 | 21 | 25 | 9 | 7 | 13 | 26 | 11 | 14 | 0 | 0 | 9 | 23 | 23 |
| n | 20 | 24 | 11 | 10 | 12 | 10 | 7 | 4 | 2 | 15 | 10 | 9 | 3 | 14 | 17 | 13 | 24 | 10 | 6 | 12 | 16 | 15 | 8 | 0 | 2 | 0 | 28 | 22 |
| w | 20 | 14 | 17 | 8 | 17 | 11 | 9 | 8 | 4 | 19 | 12 | 10 | 2 | 12 | 16 | 15 | 26 | 10 | 8 | 15 | 16 | 14 | 8 | 28 | 30 | 11 | 0 | 21 |
| y | 10 | 14 | 10 | 4 | 12 | 5 | 5 | 3 | 0 | 7 | 1 | 1 | 2 | 5 | 2 | 2 | 7 | 2 | 1 | 8 | 6 | 1 | 7 | 14 | 13 | 12 | 1 | 0 |

**Table 5.1**: Adjacent consonant pairs in the Arabic triliteral roots.

**Table 5.2**: Non-adjacent consonant pairs in the Arabic triliteral roots.

| | b | f | m | t | d | T | D | θ | ð | s | z | S | Z | ʃ | k | g | q | χ | ʁ | ħ | ʕ | h | ʔ | l | r | n | w | y |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| b | 1 | 0 | 7 | 5 | 4 | 2 | 4 | 2 | 0 | 3 | 3 | 2 | 1 | 2 | 3 | 4 | 7 | 2 | 3 | 4 | 7 | 2 | 7 | 13 | 16 | 4 | 3 | 6 |
| f | 0 | 0 | 10 | 4 | 6 | 3 | 3 | 0 | 0 | 4 | 2 | 3 | 0 | 5 | 3 | 2 | 6 | 5 | 1 | 9 | 6 | 3 | 5 | 8 | 11 | 3 | 4 | 7 |
| m | 0 | 0 | 0 | 2 | 7 | 7 | 4 | 2 | 0 | 5 | 1 | 3 | 0 | 0 | 4 | 4 | 8 | 3 | 2 | 9 | 6 | 1 | 3 | 9 | 11 | 11 | 7 | 6 |
| t | 3 | 3 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 3 | 0 | 0 | 0 | 0 | 1 | 1 | 2 | 0 | 0 | 2 | 3 | 4 | 0 | 3 | 2 | 1 | 2 | 2 |
| d | 3 | 3 | 5 | 0 | 2 | 0 | 1 | 1 | 0 | 8 | 0 | 1 | 0 | 3 | 6 | 4 | 3 | 1 | 3 | 1 | 5 | 1 | 3 | 4 | 10 | 9 | 7 | 7 |
| T | 5 | 3 | 2 | 0 | 2 | 0 | 0 | 1 | 0 | 4 | 1 | 0 | 0 | 4 | 0 | 0 | 6 | 1 | 0 | 6 | 4 | 0 | 2 | 3 | 5 | 5 | 5 | 5 |
| D | 1 | 2 | 3 | 0 | 2 | 3 | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 2 | 2 | 1 | 1 | 0 | 1 | 5 | 0 | 1 | 1 | 6 | 3 | 3 | 4 |
| θ | 5 | 1 | 2 | 1 | 1 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 4 | 4 | 2 | 1 | 3 |
| ð | 5 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 3 | 0 | 0 | 1 | 2 | 0 | 1 | 3 | 4 | 1 | 2 | 4 |
| s | 7 | 9 | 10 | 4 | 7 | 8 | 0 | 0 | 0 | 4 | 0 | 0 | 0 | 0 | 5 | 4 | 7 | 4 | 2 | 8 | 9 | 2 | 3 | 12 | 15 | 5 | 9 | 10 |
| z | 3 | 5 | 5 | 3 | 6 | 5 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 3 | 8 | 1 | 2 | 1 | 2 | 0 | 1 | 8 | 12 | 3 | 4 | 4 |
| ʃ | 7 | 6 | 3 | 3 | 6 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 3 | 2 | 3 | 7 | 8 | 0 | 2 | 4 | 7 | 3 | 5 | 7 |
| Z | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 2 | 1 | 0 | 0 |
| S | 9 | 6 | 5 | 3 | 3 | 7 | 0 | 2 | 1 | 3 | 0 | 1 | 0 | 2 | 3 | 1 | 7 | 4 | 0 | 5 | 7 | 6 | 2 | 9 | 15 | 6 | 8 | 7 |
| k | 7 | 7 | 8 | 5 | 6 | 1 | 0 | 1 | 0 | 7 | 2 | 0 | 0 | 5 | 0 | 0 | 0 | 2 | 0 | 7 | 3 | 2 | 1 | 7 | 7 | 6 | 4 | 6 |
| g | 7 | 8 | 7 | 0 | 9 | 1 | 2 | 0 | 0 | 5 | 3 | 0 | 1 | 5 | 0 | 0 | 0 | 1 | 0 | 5 | 6 | 1 | 4 | 9 | 12 | 1 | 5 | 6 |
| q | 6 | 9 | 11 | 2 | 6 | 7 | 4 | 0 | 0 | 8 | 2 | 5 | 2 | 4 | 0 | 0 | 2 | 0 | 0 | 6 | 10 | 1 | 3 | 10 | 14 | 5 | 8 | 9 |
| χ | 7 | 7 | 8 | 3 | 3 | 4 | 2 | 2 | 0 | 5 | 2 | 3 | 0 | 2 | 0 | 3 | 5 | 1 | 0 | 0 | 8 | 0 | 5 | 10 | 13 | 6 | 5 | 6 |
| ʁ | 5 | 2 | 5 | 0 | 3 | 4 | 2 | 2 | 0 | 3 | 2 | 2 | 2 | 1 | 0 | 1 | 4 | 0 | 0 | 0 | 0 | 0 | 0 | 4 | 9 | 4 | 10 | 8 |
| ħ | 10 | 7 | 11 | 0 | 9 | 3 | 2 | 4 | 2 | 6 | 5 | 3 | 1 | 3 | 7 | 4 | 8 | 0 | 0 | 0 | 0 | 0 | 2 | 7 | 15 | 9 | 12 | 9 |
| ʕ | 11 | 10 | 10 | 1 | 9 | 1 | 2 | 2 | 1 | 4 | 3 | 3 | 0 | 4 | 3 | 3 | 9 | 0 | 0 | 0 | 0 | 2 | 1 | 14 | 16 | 6 | 14 | 12 |
| h | 6 | 5 | 10 | 3 | 2 | 2 | 1 | 0 | 0 | 5 | 1 | 0 | 0 | 5 | 2 | 5 | 1 | 0 | 0 | 0 | 7 | 0 | 5 | 8 | 10 | 4 | 5 | 3 |
| ʔ | 8 | 4 | 6 | 0 | 7 | 1 | 2 | 2 | 1 | 4 | 1 | 0 | 0 | 0 | 1 | 1 | 6 | 1 | 0 | 0 | 0 | 3 | 0 | 9 | 10 | 6 | 3 | 8 |
| l | 5 | 5 | 11 | 1 | 3 | 5 | 0 | 3 | 1 | 6 | 4 | 3 | 3 | 1 | 3 | 4 | 8 | 2 | 2 | 4 | 7 | 0 | 2 | 0 | 0 | 10 | 4 | 8 |
| r | 14 | 8 | 16 | 2 | 10 | 3 | 6 | 4 | 0 | 8 | 3 | 4 | 0 | 4 | 2 | 3 | 9 | 2 | 1 | 7 | 11 | 1 | 6 | 11 | 0 | 8 | 12 | 10 |
| n | 14 | 12 | 12 | 6 | 8 | 4 | 4 | 2 | 4 | 11 | 8 | 4 | 1 | 10 | 4 | 7 | 9 | 3 | 1 | 12 | 11 | 6 | 7 | 21 | 21 | 3 | 14 | 10 |
| w | 10 | 8 | 10 | 1 | 14 | 3 | 2 | 1 | 1 | 5 | 5 | 1 | 2 | 3 | 2 | 2 | 6 | 2 | 1 | 3 | 10 | 3 | 6 | 9 | 14 | 5 | 0 | 17 |
| y | 0 | 0 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 2 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 3 | 0 | 0 | 0 | 1 | 2 | 0 | 0 |

Notably, the uvular approximants /χ/ and /ʁ/ occur in both the class of velars and uvulars and the class of guttural approximants. In addition to these major classes, Greenberg demonstrates that the coronal obstruents form two subclasses, the coronal stops {t, d, T, D}, and the coronal fricatives {θ, ð, s, z, S, Z, ʃ}. These two coronal sub-classes have stronger cooccurrence restrictions within their classes than they do with each other.

Finally, McCarthy (1994) notes that the status of the glides {w, y} in Arabic is unclear. They may have a cooccurrence restriction with each other, but they are to some extent in complementary distribution as well. They do not have any clear cooccurrence restriction with any other consonants. The glides are not included in the analysis of OCP-Place effects by McCarthy or by FBP.

### *5.2 OCP Effects in Autosegmental Phonology*

McCarthy (1986, 1988, 1994) and others (Mester 1986, Padgett 1995a, Yip 1988) have attempted to formalize the cooccurrence restrictions in the Arabic verbal roots using the notation of autosegmental phonology (Goldsmith 1979, 1994). In autosegmental phonology, representations are split into separate levels, called tiers. McCarthy showed that the non-concatenative morphology of Arabic can be represented by separating the vowels and consonants of the word form onto separate tiers. Thus, a form like *kutib* 'to be written' is represented as in (38).

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| (38) | vocalic tier: | | u | | i | | |
| | | | \| | | \| | | |
| | skeletal tier: | C | V | C | V | C | |
| | | \| | | \| | | \| | |
| | consonantal tier: | k | | t | | b | |

In (38), there are three different tiers, each of which contributes to the meaning of the word form. The consonants *ktb* 'to write', the vowels *ui*, which indicates the passive, and the skeletal pattern CVCVC which indicates the infinitive. These three morphological components combine to make the complete form. If one pattern is changed, the resulting word changes. For example, we can change the consonants in (38) to *fʕl*, resulting in *fuʕil* 'to be done'; we can change the vowels to *aa*, resulting in *katab* 'to write'; or we can change the CV-skeleton to CVVCVC, resulting in *kuutib* 'to be corresponded with'.

The most interesting cases in Arabic morphology arise when there are mismatches between elements on different tiers. For example, *kuutib* 'to be corresponded with' has a CVVCVC skeleton, with a one-to-many association on the first vowel.

```
(39)       u        i
           | \      |
        C  VV  C  V  C
        |      |     |
        k      t     b
```

Restricting the use of one-to-many mappings is crucial to the traditional analysis of consonant cooccurrence restrictions in Arabic.

Recall that Arabic does not allow roots where the first two consonants are identical, like **ddm*, but does permit the second pair of consonants to be identical, as in *mdd* 'to stretch'. McCarthy claimed that the asymmetry follows from two separate restrictions. The first is the Obligatory Contour Principle or OCP (Goldsmith 1979, Leben 1973) stated in (40).

(40) Adjacent identical elements are prohibited.

The OCP was originally proposed as a naturalness condition on autosegmental representations. If elements are represented on separate levels with the possibility of one-to many-associations, there is no way, based on the surface phoneme string, to differentiate between a representation with repeated vowels (41a) and one with a shared vowel (41b).

```
(41)  a.   V  V        b.   VV
           |  |             | /
           a  a             a
```

If (40) is present as a universal, then the representation in (41b) will always result. The OCP has since been extended to apply in other domains. Itô & Mester (1986) applied the OCP to laryngeal features in Japanese, to explain patterns in consonant voicing. McCarthy (1986) applied the OCP to the consonantal tier, to account for the lack of adjacent identical consonants in the first and second position in the Arabic roots.

In order for the autosegmental OCP to allow *mdd* but not **ddm*, McCarthy (1986) also applied a restriction on one-to-many association: Multiple associations are not allowed word initially.[14] Together, the two constraints rule out **ddm* but permit *mdd* as follows. The underlying form of the surface consonant sequence *mdd*, with a repeated consonant, is *md* with only two consonants. When the consonants *md* are to be associated with a CVCVC tier, there are two consonants and three C slots, so a one-to-many mapping results. Since multiple association is not allowed word initially, the first consonant /m/ is associated with the first C slot, and the second consonant /d/ is associated with the second slot, and also the third, to avoid leaving an empty slot, as shown in (42).

[14] McCarthy (1986) formulates the rule derivationally, restricting multiple associations to apply in a left-to-right manner. I present a non-derivational alternative as I adopt this constraint in a non-derivational framework.

```
(42)   C   CC
       |   | /
       m   d
```

Vowels are inserted normally, giving surface forms where there appear to be two separate *d*s. Given these two constraints, it is impossible to derive a surface form like **dadam*. The idiosyncratic triconsonantal verbs with repeated second and third consonants are thus seen to be underlyingly biliteral roots. They do not violate the OCP.

Evidence for McCarthy's account of *mdd*-type roots comes from an Arabic language game (McCarthy 1994). In this game, the consonants of a verbal root are permuted. On a verb form like *kattab* there are five possible outcomes.

(43) battak
kabbat
takkab
bakkat
tabbak

Notice that in all of these outcomes, the middle consonant pair is always repeated. This is accounted for if *kattab* is represented as in (44), with a medial one-to-many association, and the consonants on the consonantal tier are permuted in the word game.[15]

```
(44)   skeletal tier:       C   V   CC   V   C
                            |       | /      |
       consonantal tier:    k       t        b
```

When a form like *maddad*, based on the root *md*, is involved in the same language game, there is only one outcome, *dammam* indicating that the final consonant must always be the same as the second. This is expected if *maddad* has only two underlying consonants.

McCarthy (1988, 1994) extends the restriction against adjacent identical consonants further, to attempt to account for the strong cooccurrence restriction between homorganic consonants. McCarthy proposes that there is a second OCP based constraint, OCP-Place, where the domain of the OCP is restricted to the place of articulation features. McCarthy separates this restriction from the total OCP. He states OCP-Place is not an absolute effect like the total OCP, but rather a strong tendency. He also claims that OCP-Place applies to consonants which are separated by intervening consonants. Note that, as pointed out in Pierrehumbert (1993), his treatment of the total OCP only applies to adjacent consonants; he fails to consider the possibility of a cooccurrence restriction between non-adjacent identical consonants, as discussed below.

---

[15] Note that the structure of this form is problematic for McCarthy's original derivational approach to association. If association proceeds left-to-right we would expect *katbab* for the word form. The non-derivational constraint, which avoids initial one-to-many associations, is satisfied by *kattab*.

Even though McCarthy admits that OCP-Place is not a categorical effect, it is formalized using the same absolute OCP constraint, with the domain restricted to the place of articulation tier.

First, let us examine effects of OCP-Place on adjacent consonants. A root like *ktb* ‘to write’ has the autosegmental representation as in (45), where irrelevant features and tiers are omitted for the sake of clarity.

```
(45)   skeletal tier:      C          C          C
                           |          |          |
       place tier:     [dorsal]  [coronal]  [labial]
```

An ill formed root, like **dtb*, would be represented as in (46).

```
(46)   *   C            C            C
           |            |            |
       [coronal]    [coronal]    [labial]
```

The adjacent identical coronal features on the place tier violate OCP-Place. In order for this analysis to account for the cooccurrence restriction, a constraint against one-to-many mappings of individual place features is required, so that a representation like (47) is disallowed.

```
(47)   *   C     C         C
            \   /          |
          [coronal]     [labial]
```

McCarthy (1994) proposes just such a constraint, claiming that one-to-many associations of place features are not allowed. The prohibition against shared place features is crucial to McCarthy’s account of OCP-Place effects. Without it, OCP-Place can be circumvented by multiple associations. Thus, OCP-Place effects depend on two independent constraints.

The general approach of autosegmental phonology is to use tier separation, in combination with the OCP, to account for cooccurrence restrictions. Tier separation makes objects on a tier ‘adjacent’ at the level of the tier, and moves irrelevant material to other tiers. This mechanism is extended to account for OCP-Place effects for non-adjacent consonants. Since OCP-Place effects are also found for non-adjacent consonants, McCarthy concludes that each of the different place features must be placed on its own tier. Intervening consonants with different places of articulation do not block application of the OCP-Place constraint because of tier separation. The root **ftb* is represented as:

```
(48)     [lab]        [lab]
           |            |
         * C     C      C
                 |
               [cor]
```

In (48), the place specifications for /f/ and /b/ are adjacent on the labial tier, violating OCP-Place. The [coronal] feature is transparent to the application of OCP-Place on the labial tier. In the non-adjacent case, the constraint against one-to-many mappings for place features is still required, to avoid associating a single labial feature with two separate segments.

One final modification to McCarthy's analysis is required to account for the split of the coronal consonants into two classes, sonorants and obstruents. Recall from section 5.1 that the coronal obstruents and coronal sonorants are in different cooccurrence classes, even though they share place of articulation. Padgett (1995a) has proposed that the domain of OCP-Place can be restricted by features in addition to place, which join with the place features into an 'articulator group'. He proposes that the domain of OCP-Place for a particular place feature, like [coronal], can be restricted by the other features in the articulator group, making the rule sensitive to manner features for coronals. FBP point out that this is a stipulation, which offers no explanation as to the cause of the coronal split.

Pierrehumbert (1993) claims that similarity is the basis for the Arabic cooccurrence restrictions. If cooccurrence restrictions are based on similarity, then we would expect all classes to show sub-regularities based on manner of articulation. The coronal class clearly has this pattern. Below, I review evidence from FBP that the dorso-guttural class is a single place class divided by manner features. In addition, voicing features as well as place and manner features are relevant to OCP effects (FBP). Thus OCP effects are based on all the features of a segment. The effect of additional features is to sub-classify and cross-classify the major classes.

### *5.3 Gradient Data and Patterns of Cross-Classification*

The FBP account of the OCP-Place constraint in Arabic is based on similarity computed over three place of articulation sub-lattices: [labial], [coronal], and [dorso-guttural]. Examples of such sub-lattices were given in chapter 2 for English. Similarity for the Arabic consonants is computed based on the features in (49).

(49) a. PLACE features:

| | b | f | m | t | d | T | D | θ | ð | s | z | S | Z | ʃ | k | g | q | χ | ʁ | ħ | ʕ | h | ʔ | l | r | n |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| labial | + | + | + | | | | | | | | | | | | | | | | | | | | | | | |
| coronal | | | | + | + | + | + | + | + | + | + | + | + | + | | | | | | | | | | + | + | + |
| dors-gut | | | | | | + | + | | | | | + | + | | + | + | + | + | + | + | + | + | + | | | |
| interdent | | | | | | | | + | + | | | | | | | | | | | | | | | | | |
| dental | | | | + | + | | | | | + | + | | | | | | | | | | | | | | | |
| alveolar | | | | | | + | + | | | | | + | + | | | | | | | | | | | | | |
| palatal | | | | | | | | | | | | | | + | | | | | | | | | | | | |
| dorsal | | | | | | | | | | | | | | | + | + | + | + | + | | | | | | | |
| guttural | | | | | | + | + | | | | | + | + | | | | | | | + | + | + | + | | | |
| velar | | | | | | | | | | | | | | | + | + | | | | | | | | | | |
| hi-uvular | | | | | | | | | | | | | | | | | + | + | + | | | | | | | |
| lo-uvular | | | | | | + | + | | | | | + | + | | | | | | | | | | | | | |
| pharyng | | | | | | | | | | | | | | | | | | | | + | + | | | | | |
| laryng | | | | | | | | | | | | | | | | | | | | | | + | + | | | |

b. MANNER features:

| | b | f | m | t | d | T | D | θ | ð | s | z | S | Z | ʃ | k | g | q | χ | ʁ | ħ | ʕ | h | ʔ | l | r | n |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| obstruent | + | + | | + | + | + | + | + | + | + | + | + | + | + | + | + | + | | | | | | | | | |
| sonorant | | | + | | | | | | | | | | | | | | | | | | | | | + | + | + |
| approx | | | | | | | | | | | | | | | | | | + | + | + | + | + | + | | | |
| stop | + | | + | + | + | + | + | | | | | | | | + | + | + | | | | | | | | | |
| fricative | | + | | | | | | + | + | + | + | + | + | + | | | | | | | | | | | | |
| lateral | | | | | | | | | | | | | | | | | | | | | | | | + | | |
| rhotic | | | | | | | | | | | | | | | | | | | | | | | | | + | |
| nasal | | | + | | | | | | | | | | | | | | | | | | | | | | | + |

c. LARYNGEAL features:

| | b | f | m | t | d | T | D | θ | ð | s | z | S | Z | ʃ | k | g | q | χ | ʁ | ħ | ʕ | h | ʔ | l | r | n |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| voice | + | | + | | + | | + | | + | | + | | + | | | + | | | + | | | | | + | + | + |
| voiceless | | + | | + | | + | | + | | + | | + | | + | + | | + | + | | + | | + | | | | |
| con glott | | | | | | | | | | | | | | | | | | | | | + | | + | | | |

Similarity for the Arabic consonants is given in Table 5.3. The similarity metric was computed over sub-lattices since OCP-Place only restricts the cooccurrence of homorganic consonant pairs. The emphatic coronals, which are found in both the coronal and dorso-guttural lattices, were given the maximal similarity value from the two sub-lattices. Consonant pairs which were not found on the same lattice, and thus do not share place of articulation, were given similarity zero.

To measure the strength of OCP-Place effects in Arabic, Pierrehumbert (1993) used the O/E measure, which was used in chapter 4 for speech errors. Expected cooccurrence is computed based on the overall frequency of consonants in each position in the root. Table 5.4 shows actual and expected counts along with O/E aggregated by similarity. This table reveals the gradience of the OCP-Place constraint.

| | b | f | m | t | d | T | D | θ | ð | s | z | S | Z | ʃ | k | g | q | χ | ʁ | ħ | ʕ | h | ʔ | l | r | n |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| b | 1 | | | | | | | | | | | | | | | | | | | | | | | | | |
| f | 0.4 | 1 | | | | | | | | | | | | | | | | | | | | | | | | |
| m | 0.4 | 0.2 | 1 | | | | | | | | | | | | | | | | | | | | | | | |
| t | 0 | 0 | 0 | 1 | | | | | | | | | | | | | | | | | | | | | | |
| d | 0 | 0 | 0 | 0.38 | 1 | | | | | | | | | | | | | | | | | | | | | |
| T | 0 | 0 | 0 | 0.43 | 0.22 | 1 | | | | | | | | | | | | | | | | | | | | |
| D | 0 | 0 | 0 | 0.22 | 0.5 | 0.38 | 1 | | | | | | | | | | | | | | | | | | | |
| θ | 0 | 0 | 0 | 0.21 | 0.12 | 0.21 | 0.12 | 1 | | | | | | | | | | | | | | | | | | |
| ð | 0 | 0 | 0 | 0.13 | 0.25 | 0.13 | 0.25 | 0.36 | 1 | | | | | | | | | | | | | | | | | |
| s | 0 | 0 | 0 | 0.36 | 0.17 | 0.19 | 0.11 | 0.45 | 0.21 | 1 | | | | | | | | | | | | | | | | |
| z | 0 | 0 | 0 | 0.18 | 0.38 | 0.11 | 0.22 | 0.21 | 0.5 | 0.36 | 1 | | | | | | | | | | | | | | | |
| S | 0 | 0 | 0 | 0.19 | 0.11 | 0.67 | 0.27 | 0.45 | 0.21 | 0.36 | 0.19 | 1 | | | | | | | | | | | | | | |
| Z | 0 | 0 | 0 | 0.11 | 0.22 | 0.27 | 0.64 | 0.21 | 0.45 | 0.19 | 0.38 | 0.36 | 1 | | | | | | | | | | | | | |
| ʃ | 0 | 0 | 0 | 0.23 | 0.13 | 0.23 | 0.13 | 0.63 | 0.27 | 0.5 | 0.23 | 0.5 | 0.23 | 1 | | | | | | | | | | | | |
| k | 0 | 0 | 0 | 0 | 0 | 0.33 | 0.15 | 0 | 0 | 0 | 0 | 0.25 | 0.12 | 0 | 1 | | | | | | | | | | | |
| g | 0 | 0 | 0 | 0 | 0 | 0.15 | 0.38 | 0 | 0 | 0 | 0 | 0.12 | 0.29 | 0 | 0.35 | 1 | | | | | | | | | | |
| q | 0 | 0 | 0 | 0 | 0 | 0.32 | 0.14 | 0 | 0 | 0 | 0 | 0.24 | 0.11 | 0 | 0.67 | 0.26 | 1 | | | | | | | | | |
| χ | 0 | 0 | 0 | 0 | 0 | 0.10 | 0.05 | 0 | 0 | 0 | 0 | 0.13 | 0.06 | 0 | 0.22 | 0.11 | 0.35 | 1 | | | | | | | | |
| ʁ | 0 | 0 | 0 | 0 | 0 | 0.05 | 0.12 | 0 | 0 | 0 | 0 | 0.07 | 0.15 | 0 | 0.11 | 0.27 | 0.17 | 0.38 | 1 | | | | | | | |
| ħ | 0 | 0 | 0 | 0 | 0 | 0.22 | 0.11 | 0 | 0 | 0 | 0 | 0.29 | 0.13 | 0 | 0.1 | 0.05 | 0.1 | 0.25 | 0.13 | 1 | | | | | | |
| ʕ | 0 | 0 | 0 | 0 | 0 | 0.12 | 0.13 | 0 | 0 | 0 | 0 | 0.15 | 0.17 | 0 | 0.06 | 0.06 | 0.05 | 0.13 | 0.15 | 0.42 | 1 | | | | | |
| h | 0 | 0 | 0 | 0 | 0 | 0.22 | 0.11 | 0 | 0 | 0 | 0 | 0.29 | 0.13 | 0 | 0.1 | 0.05 | 0.1 | 0.25 | 0.13 | 0.67 | 0.31 | 1 | | | | |
| ʔ | 0 | 0 | 0 | 0 | 0 | 0.12 | 0.13 | 0 | 0 | 0 | 0 | 0.15 | 0.15 | 0 | 0.06 | 0.06 | 0.05 | 0.13 | 0.15 | 0.31 | 0.56 | 0.42 | 1 | | | |
| l | 0 | 0 | 0 | 0.08 | 0.14 | 0.08 | 0.14 | 0.1 | 0.2 | 0.08 | 0.17 | 0.08 | 0.17 | 0.11 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | | |
| r | 0 | 0 | 0 | 0.08 | 0.14 | 0.08 | 0.14 | 0.1 | 0.2 | 0.08 | 0.17 | 0.08 | 0.17 | 0.11 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.6 | 1 | |
| n | 0 | 0 | 0 | 0.14 | 0.29 | 0.14 | 0.29 | 0.08 | 0.17 | 0.07 | 0.14 | 0.07 | 0.14 | 0.09 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.43 | 0.43 | 1 |

**Table 5.3**: Similarity of Arabic consonant pairs over place lattices.

**Table 5.4**: Adjacent and non-adjacent consonant pairs in the Arabic verbal roots.

| | Adjacent consonants | | | Non-adjacent consonants | | |
|---|---|---|---|---|---|---|
| Similarity | Actual | Expected | O/E | Actual | Expected | O/E |
| 0 | 2978 | 2349.3 | 1.27 | 1411 | 1248.4 | 1.13 |
| 0-0.1 | 481 | 365.2 | 1.23 | 219 | 203.2 | 1.08 |
| 0.1-0.2 | 492 | 550.6 | 1.18 | 308 | 288.7 | 1.09 |
| 0.2-0.3 | 151 | 260.2 | 0.58 | 96 | 124.2 | 0.77 |
| 0.3-0.4 | 29 | 131.2 | 0.22 | 50 | 67.0 | 0.75 |
| 0.4-0.5 | 14 | 180.2 | 0.08 | 75 | 103.2 | 0.73 |
| 0.5-0.6 | 3 | 165.0 | 0.07 | 8 | 25.0 | 0.32 |
| 0.6-0.7 | 0 | 90.2 | 0.00 | 13 | 40.4 | 0.32 |
| 1 | 1 | 199.6 | 0.01 | 16 | 103.9 | 0.15 |

The autosegmental model has a number of problems accounting for the gradience of the OCP-Place constraint (Pierrehumbert 1993, FBP). The autosegmental OCP-Place constraint cannot for the weakening of OCP-Place over distance, sub-classification within the place of articulation classes, and cross-classification between the major classes.

Comparing the adjacent to the non-adjacent O/E shows the weakening of the cooccurrence constraint for non-adjacent consonants. There is an effect of distance on OCP-Place. Pierrehumbert (1993) demonstrates the difficulties that the standard account has with the distance effects in Arabic. The autosegmental model predicts that these two cases are equivalent, since tier separation was used to make place features ‘adjacent’ for non-adjacent consonants. In addition, no account based on autosegmental adjacency can account for the strong cooccurrence restriction between identical non-adjacent consonants (Pierrehumbert 1993, 1994). Identical, non-adjacent consonants are not adjacent on every feature tier as there are always be some feature specifications present from the intervening consonant. Thus, in the autosegmental account, the restriction against non-adjacent identical consonants is predicted to be only as strong as the restriction between homorganic pairs. In fact, the restriction against non-adjacent identical pairs is stronger than the restriction against non-adjacent homorganic pairs, but weaker than the restriction against adjacent identical pairs.

Greenberg (1950) originally pointed out that the coronal obstruents actually break into two classes, the coronal stops and coronal fricatives. This sub-classification has not been accounted for in a non-arbitrary way in the autosegmental account. There are several other cases of sub-classification which also defy explanation in the autosegmental account. Pierrehumbert (1993) reports that the emphatic coronals /T/, /D/, /S/, and /Z/ have a stronger cooccurrence restriction with each other than they do with the other coronals. She also reports that the labial class does show some evidence of sub-classification by manner, parallel to the coronals. While there are no observed pairs of labial consonants in adjacent position, there are 17 such pairs in non-adjacent position. Of those, 16 involve /m/ with a labial obstruent (/b/ or /f/). There is only 1 pair with two obstruents. Finally, she reports that /l/ and /r/ form a subclass of the coronal sonorants, which have stronger cooccurrence restrictions with each other than they do with /n/.

FBP report additional regularities outside of the previously noted cooccurrence classes. Given the strong cooccurrence restrictions within the major classes, other consonant combinations outside of the major classes will necessarily be overrepresented (Pierrehumbert 1993). FBP show that there are clear patterns of cross-classification between the major classes. The cooccurrence of the velar and uvular stops {k, g, q} and the guttural approximants {ħ, ʕ, h, ʔ} is less than the expected amount of overrepresentation for classes which are unrestricted. While adjacent pairs of consonants in these two classes are not underrepresented (153 observed and 147 expected, O/E = 1.04), they do show far less overrepresentation than they should: the O/E for non-homorganic adjacent consonants is 1.27. FBP also demonstrate an effect of secondary place on cooccurrence. This is a result of secondary [dorso-guttural] place for the emphatic stops /T/ and /D/. The O/E for {k, g, q} with {T, D} is 0.53; while for {k, g, q} with {t, d}, O/E is 1.00. There is a similar effect for /S/ and /Z/, which also have [dorso-guttural] articulation. O/E for {k, g, q} with {S, Z} is 0.77; the O/E for {k, g, q} with the other coronal fricatives is 1.25. Thus there is underrepresentation between the coronal emphatics and the velar and uvular stops.

## *5.4 The Similarity Account and Contrastive Underspecification*

To summarize the preceding discussion, the effects of OCP-Place in the Arabic verbal roots are thoroughly gradient. Rather than attempting to fit a categorical statement of well-formedness to the data, Pierrehumbert (1993), Berkley (1994a, b), and FBP adopt a gradient constraint based on the continuous variable similarity. Pierrehumbert (1993) originally claimed that there is a cooccurrence restriction on homorganic consonants in proportion to their perceived similarity. Identical consonants are maximally similar, and thus have the strongest cooccurrence constraint. Consonants which differ in major class features, but are still homorganic, like the coronal obstruents and coronal sonorants, are subject to a weak cooccurrence restriction. Temporal distance affects perceived similarity by creating interference and allowing memory to decay, weakening the perception of similarity over distance.

To account for differences in OCP-Place effects within major classes, for example the strong cooccurrence restriction between labial obstruents and sonorants versus the weak restriction between coronal obstruents and sonorants, Pierrehumbert (1993) employed contrastive underspecification. I show below, following FBP, that the natural classes similarity model provides a better account of the differences within classes than the contrastive underspecification similarity model.

### *5.4.1 The similarity model.*

The OCP as a similarity effect is a universal, with languages differing on which features the effect is based upon (e.g. place features in Arabic, laryngeal features for OCP effects in tone languages). In addition, the degree to which similarity triggers an OCP effect can vary from language to language. Arabic is one of the strongest cases, where a small degree of similarity can trigger some OCP effects between homorganic consonants. English is a weaker case (Berkley 1994a, b), which is discussed in chapter 10. Languages which show OCP effects only on identical consonants have an OCP-Place effect which is subject to the highest similarity threshold before there is a cooccurrence restriction. The similarity model does not rely on adjacency or tier separation, and thus can apply to all features over any distance.

The similarity model accounts for the gradience of OCP effects, both between adjacent consonants and over distance. All of the classes originally noted in Greenberg (1950), which have strong cooccurrence restrictions, are highly similar. The sub-classification and cross-classification shown in section 2.3 are accounted for by the similarity model. Within the major classes there are sub-classes of more similar segments, like the coronal stops and coronal fricatives. Between the major classes there are segments which share some features, like the voiced coronal obstruents the coronal sonorant.

In addition to effects of sub and cross-classification, Pierrehumbert (1993) showed that the similarity model properly accounts for the effects of distance. It is well known that the ability to make comparisons is influenced by temporal distance and interference by intervening material. The effects of temporal spacing in the relatively short time scale of speech, on the order of one to two seconds, is found both in visual and auditory perception (Pisoni 1973, Eriksen & Shultze 1978). Massaro (1970) has shown that the ability to compare the first and third items in a three item sequence is impaired by the second. The exact character of the medial item determines the degree to which it masks the first. For non-adjacent consonants in first and third position, the consonant in second position acts as interference. Thus, the perception of similarity or dissimilarity between the first and third consonants is diminished by the intervening second consonant. As distance and intervening material increases, the evaluation of similarity will become increasingly unreliable. Consonants which are not perceived as similar or dissimilar tend to cooccur at random, resulting in an O/E of one.

### *5.4.2 The role of contrastive underspecification.*

Underspecification was employed by Pierrehumbert (1993) to account for the difference in the effects of manner on the coronals and dorso-gutturals, in comparison to the labials. Unlike the coronals and dorso-gutturals, the labials do not split into two major classes based on manner. Pierrehumbert used the contrastively underspecified featural specifications as shown in (50), where a smaller set of features are used to contrast /f/, /m/ than /s/, /n/. The additional features for /s/ are included due to the larger number of coronal fricatives which must be contrasted with /s/. These segments and features are also included in (50). In particular, contrasting values for [voice], [anterior], and [dental] differentiate /s/ from /z/, /ʃ/, and /θ/, respectively. The feature [nasal] is required to differentiate /n/ from the other coronal sonorants. There are no other labial fricatives or labial sonorants, so no additional features are needed to differentiate the labials.

(50)

| | f | m | s | n | z | ʃ | θ |
|---|---|---|---|---|---|---|---|
| lab | + | + | | | | | |
| cor | | | + | + | + | + | + |
| son | - | + | - | + | - | - | - |
| cont | + | | + | | + | + | + |
| nasal | | | | + | | | |
| voice | | | - | | + | - | - |
| ant | | | + | | + | | + |
| dental | | | - | | - | - | + |

Under the feature assignments based on contrastive underspecification, /f/-/m/ differ on the two features [sonorant] and [continuant] and share [labial]. The pair /s/-/n/ also differ on [sonorant] and [continuant], plus [nasal], [voice], [ant], and [dental]. Using the Pierrehumbert (1993) feature similarity model, presented in chapter 3, that computes similarity as the ratio of shared features to the shared and non-shared features, /s/-/n/ have similarity of 1/7, while /f/-/m/ has similarity 1/3. The greater similarity of /f/-/m/ under this approach accounts for the stronger OCP effect between /f/-/m/ than /s/-/n/.

The natural classes similarity model produces the same result without underspecification. Recall that since different natural classes are created by feature combinations only when those features are contrastive, features which are partially or totally redundant contribute to fewer distinct natural classes. For example, the similarity of /s/-/n/ is 0.07 in the natural classes model (from Table 5.3), while the similarity of /f/-/m/ is 0.2. Given that underspecification has formal difficulties (Broe 1993, discussed in chapter 2), FBP claim the natural classes model is superior.

FBP also found empirical evidence against underspecification. There is an effect of voicing on the strength of OCP-Place between the coronal obstruents and coronal sonorants. Sonorants are redundantly voiced, so the similarity model using contrastive underspecification predicts that voicing should be irrelevant to the similarity of sonorants to obstruents. However, the voiceless obstruents cooccur with the sonorants more than voiced obstruents do. The aggregate O/E is 1.15 for coronal sonorants and voiced coronal obstruents (239 actual and 207 expected). The aggregate O/E is 1.31 for coronal sonorants with voiceless obstruents (245 actual and 187 expected). The natural classes model does not eliminate redundant voicing, rather, redundant voicing has a lesser effect on similarity than non-redundant features.

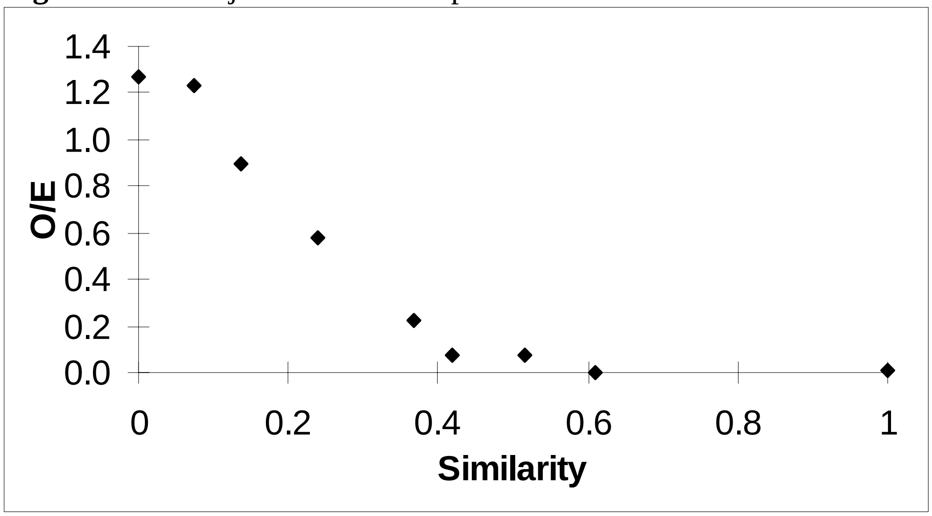
## *5.5 The Stochastic Constraint Model of OCP-Place*

FBP go beyond previous similarity based analyses of OCP-Place by presenting an explicit model of a gradient cooccurrence constraint. Figure 5.1a shows the aggregate O/E for adjacent consonant pairs in Arabic, and Figure 5.1b shows the comparable data for non-adjacent pairs, based on Table 5.4. Figure 5.1a shows an ogival pattern that is reminiscent of categorical perception (see Repp 1984 for a review). The data in Figure 5.1b are more scattered, but an ogival model does not appear at all unreasonable. The logistic function is often used to model ogival data (Tukey 1977). The equation of the logistic function is given in (51).

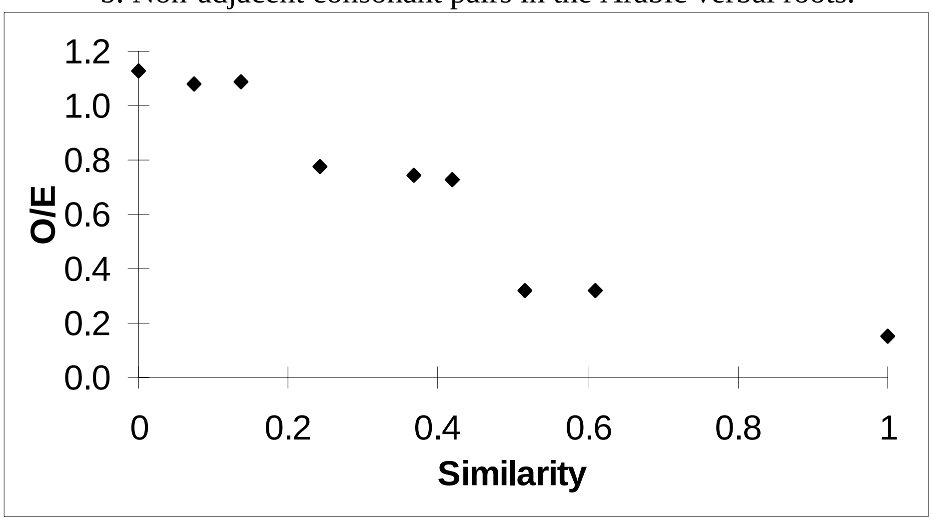
(51)

$$y = \frac{1}{1 + e^{K+Sx}}$$

**Figure 5.1**: a. Adjacent consonant pairs in the Arabic verbal roots.

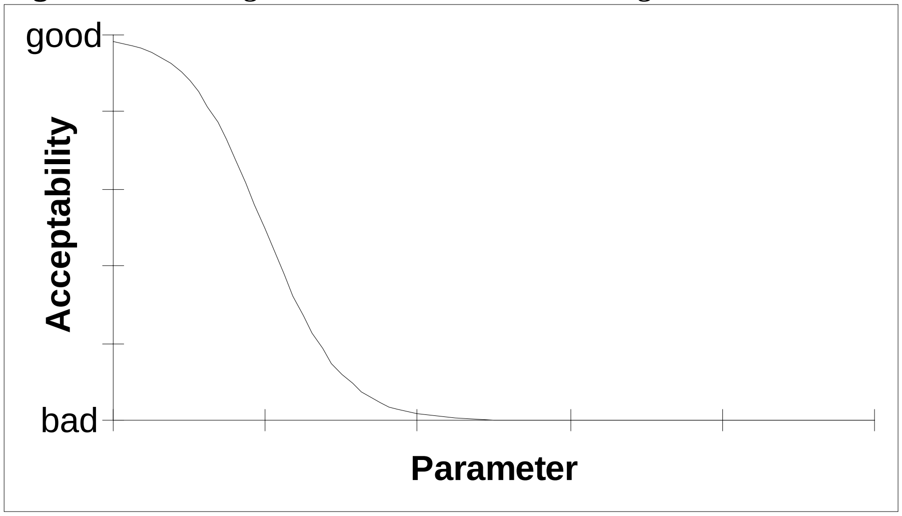



b. Non-adjacent consonant pairs in the Arabic verbal roots.




FBP offer the logistic function as a model of a gradient linguistic constraint based on a continuous variable like similarity. They call the model the STOCHASTIC CONSTRAINT model. The interpretation of the logistic function as a stochastic constraint model is presented in Figure 5.2. In the stochastic constraint model of a gradient linguistic constraint, the independent variable, $x$, is the relevant parameter upon which the constraint is based. In our case, this is similarity. The domain of the logistic is $(-\infty,\infty)$. The dependent variable, $y$, is the acceptability of a form. The range of the logistic is (0,1). The stochastic constraint model assumes that the statistical frequency of a form is a measure of the relative acceptability of a form with respect to other forms. The most acceptable forms are the most frequent, and unacceptable forms are so improbable that they are not expected to occur. Thus, in practice, the dependent variable is O/E.

**Figure 5.2**: The logistic function as a model of a gradient constraint.



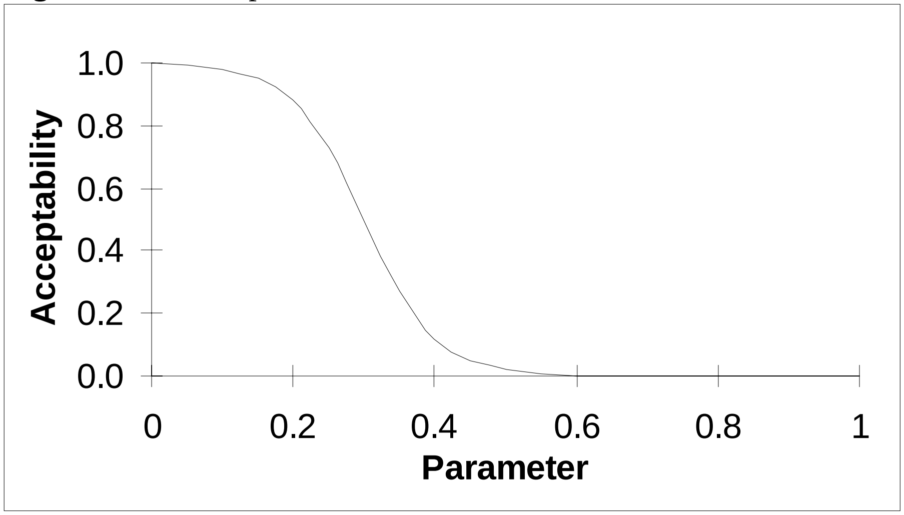

The stochastic constraint is used by FBP to model the OCP-Place constraint in Arabic. The basic model is shown in (52). The parameters of the stochastic constraint model; *A*, *K*, and *S*; are used to fit the constraint to actual data. The *A* parameter controls the upper asymptote of the logistic, and represents the assumption of a proportional relationship between acceptability and relative frequency. The logistic has range (0,1), but O/E can, in theory, range over [0,∞). The parameters *K* and *S* control the shape of the curve. *K* is called the INTERCEPT, and *K* controls the location of the midpoint (where $y = A/2$) of the curve, which is the informal 'constraint boundary'. *S* is called the SHARPNESS and it controls the steepness with which acceptability changes in response to changes in the parameter upon which the constraint is based. In general, stronger constraints have greater sharpness, and weaker constraints have lesser sharpness.

(52)

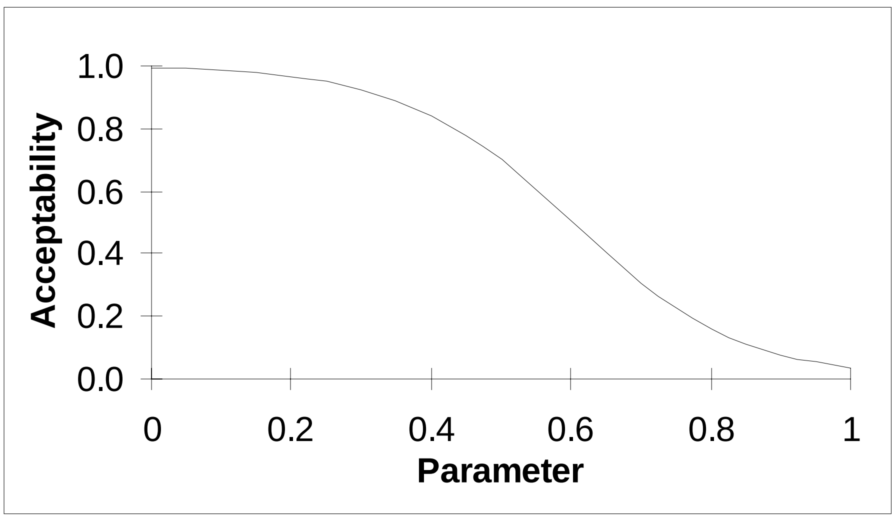
$$\mathrm{O/E} = \frac{A}{1 + e^{K+Sx}}$$

Figures 5.3a and 5.3b exemplify two possible parameterizations of the stochastic constraint model. Figure 5.3a shows a sharp constraint, which has a steep boundary between highly acceptable and highly unacceptable forms (modeled by the parameter $S = 20$), with constraint boundary x = 0.3 ($K = -6$). Figure 5.3b shows a weak constraint ($S = 8.3$) with boundary x = 0.6 ($K = -5$).

**Figure 5.3**: a. Sharp stochastic constraint.

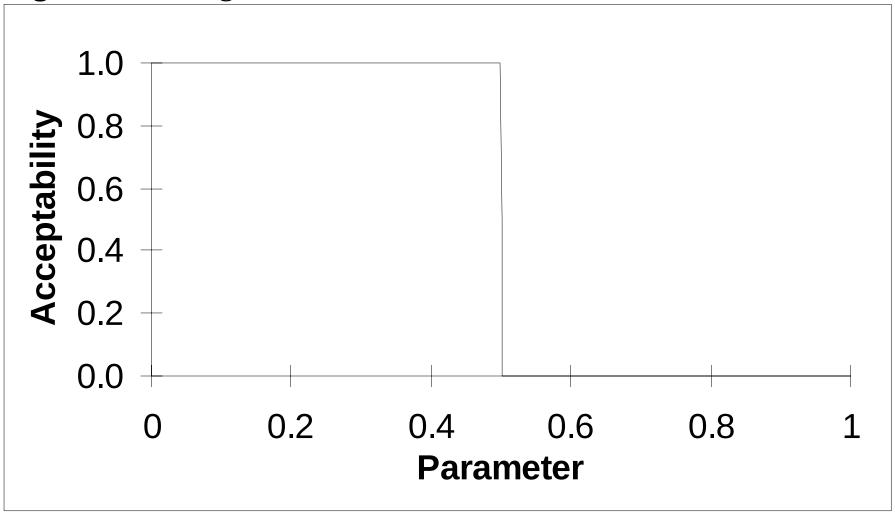



b. Weak stochastic constraint.




The stochastic constraint model has an advantage over traditional categorical constraints in that it can model gradient data. The stochastic constraint can also model a categorical constraint, as shown in Figure 5.4. The categorical constraint shown here has sharpness $S$ = 10,000 and boundary at x = 0.5 ($K$ = -5,000).

**Figure 5.4**: Categorical constraint modeled with the stochastic constraint.




Since the size of the lexicon is finite, any two forms which are distinct with respect to some relevant parameter, like similarity, must be separated by some minimal difference in the parameter (the only way there can be no lowest bound on the distance between distinct parameter values is if there are an infinite number of values). Since there is no upper bound on *S*, the stochastic constraint model can always be made sharp enough to categorically differentiate any two distinct parameter values. In other words, since the size of the lexicon is finite, it can always be made sharp enough to categorically differentiate between any two distinct forms.

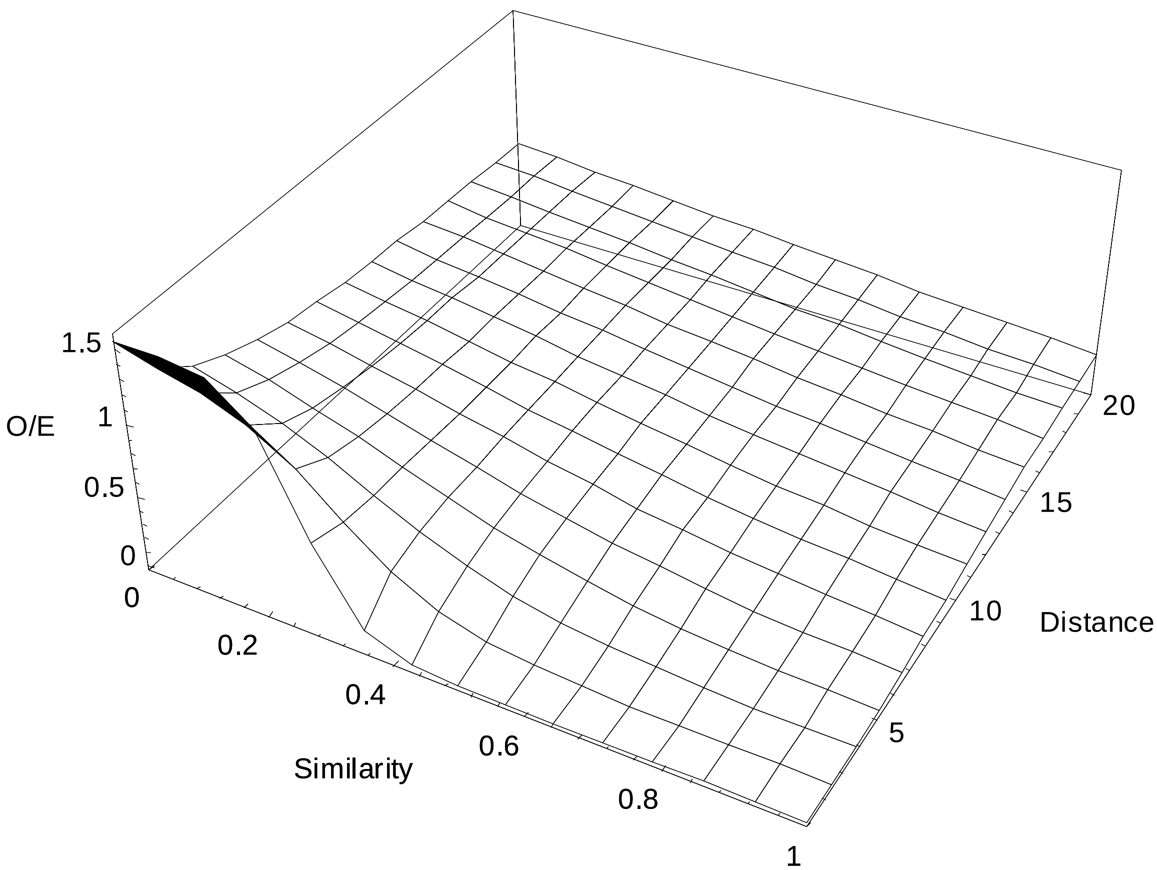
In order to account for the weakening of OCP-Place effects over distance, FBP modify the basic stochastic constraint model in (52). The addition of a distance variable allows them to simultaneously model adjacent and non-adjacent data. The modified model is shown in (53). FBP measure distance in terms of surface segments (distance = 2 for adjacent consonants, 4 for non-adjacent consonants).

(53)

$$\text{O/E} = \frac{1 + \text{A/distance}}{1 + e^{(\text{K} + \text{S}\cdot\text{similarity})/\text{distance}}}$$

The model in (53) is specifically designed for distance effects. As distance increases, OCP-Place effects weaken. At the theoretical limit, predicted O/E is one for any value of similarity (i.e. the mathematical limit as distance approaches infinity is one). Figure 5.5, shows the predicted O/E against similarity for distances from 1 to 20 phonemes.

**Figure 5.5**: Effects of distance on the stochastic constraint model of Arabic cooccurrence.




FBP quantitatively compare the logistic model in (53) to four other models of cooccurrence for the Arabic verbal roots. The first model is the frequency model, which assumes that there is no similarity effect in the Arabic roots. The predicted occurrences of a consonant pair are equal to the expected number of occurrences at random. Two other models are based on the traditional autosegmental approach. The categorical model is a strict interpretation of the formal account of McCarthy (1988, 1994) which prohibits homorganic consonants from cooccurring within a root. The soft model is the model intended by McCarthy (1994) which cannot be properly formalized using autosegmental theory. In the soft model, adjacent identical consonants are prohibited, and homorganic consonants are underrepresented at a constant rate. The final model, the feature similarity model, is the model of Pierrehumbert (1993), which computes similarity based on shared and non-shared features. I call the FBP model the natural classes model since in this model similarity is computed based on shared and non-shared natural classes. This model and the feature similarity model use the same features.

I compare models on $R^2$, as well as what I call the 'relative' $R^2$. The relative $R^2$ is based on the amount of reduction in the residual sum of squares of each model over a model which predicts occurrence by frequency, in other words the frequency model. The frequency model underlies all of the models presented here and accounts for a considerable portion of the variation in the Arabic data (see chapter 10 for more on the effects of frequency in phonotactics).

(54)
$$\text{Relative } R^2 = 1 - \frac{\text{Residual SS}}{\text{Frequency Model Residual SS}}$$

Finally, I also compare the relative $R^2$ over just the homorganic consonant pairs, which are the crucial data for OCP-Place effects.

Table 5.5 compares the different models of the Arabic data, and shows best fit model parameters. The similarity models provide a superior fit to the autosegmental models of McCarthy (1994). The Natural Classes model provides the best fit among the similarity models, providing additional evidence for the use of the natural classes similarity model to represent similarity among segments.

**Table 5.5**: Comparison of models of OCP-Place in Arabic.

| Model | $R^2$ | Relative $R^2$ | Homorganic Relative $R^2$ | Model Parameters |
|---|---|---|---|---|
| Frequency Model | 0.55 | - | - | O = E |
| Categorical Model | 0.74 | 0.42 | 0.56 | O/E = 0 for homorganic pairs<br>O/E = 1.20 otherwise |
| Soft Model | 0.77 | 0.50 | 0.69 | O/E = 0 for adjacent identical<br>O/E = 0.38 for homorganic<br>O/E = 1.20 otherwise |
| Feature Model | 0.78 | 0.52 | 0.73 | A = 0.48, S = 25.2, K = -10.1 |
| Natural Classes Model | 0.80 | 0.56 | 0.80 | A = 0.56, S = 27.5, K = -7.5 |

## *5.6 The Stochastic Constraint Model of a Linguistic Constraint*

The analysis of OCP-Place effects in FBP introduces the logistic function as a quantitative model of a probabilistic or stochastic constraint. FBP note that the gradient nature of the OCP-Place data is problematic for the standard model of a constraint in linguistic theory, which is categorical. I review their discussion here.

The gradient nature of OCP effects are not captured in any current formalism. In Declarative Phonology (Scobbie 1993), constraints are categorical statements of surface true well-formedness. The actual forms in a language simultaneously satisfy all constraints. An account of OCP effects based on current formulations of Declarative Phonology is quantitatively equivalent to the categorical autosegmental model presented above. Declarative Phonology could be extended to incorporate the stochastic constraint model. Instead of requiring absolute satisfaction of all constraints, constraints could be combined stochastically, with forms which violate many constraints becoming increasingly improbable, as suggested in Pierrehumbert & Nair (1995). This model of gradient constraint combination is implemented in chapter 7.

In Optimality Theory (Prince & Smolensky 1993), violable constraints are allowed, when violation of a constraint occurs in order to satisfy the needs of another constraint which has logical priority. The existence of violable constraints allows statistically variable data, since many forms which are in violation of a constraint might be found. However, Optimality Theory cannot account for the particular patterns of cooccurrence found in Arabic or English (Berkley 1994a, b). These cooccurrence constraints are based on statistical patterns over the lexicon. The architecture of Optimality Theory is based on finding the optimal pairing of input and output given the constraints. OCP-Place does not influence what the output is for any particular input, but rather it constrains the space of possible inputs and outputs in a probabilistic manner. Constraints like OCP-Place are not possible in the Optimality Theoretic architecture.

The stochastic constraint model shares many commonalities with so-called ‘phonetic implementation’ rules. Phonetic implementation is often divorced from phonology proper (e.g. Keating 1983, Pierrehumbert & Beckman 1988). This division is motivated by a desire to separate the categorical, symbolic phonological system from the probabilistic and gradient nature of real speech. Since phonotactics represent implicit linguistic knowledge about the possible words in a language, the results of FBP show that gradient phenomena must be incorporated within phonology proper, as argued for in Pierrehumbert (1994). Pierrehumbert (1994) found that incorporating probabilistic knowledge into the grammar of English reduces the set of admissible word medial clusters of three or more consonants from 8708 to 200 most likely combinations. There are 50 actually occurring medial triconsonantal clusters, and nearly all of them are in the most likely 200. I discuss Pierrehumbert (1994) in detail in chapter 10.

The rules of phonetic implementation are often language specific, which also undermines the existence of a dividing line between phonology and phonetics (see the review in Pierrehumbert 1990). Unification of phonological and phonetic knowledge in this way allows similar effects within the two domains to be accounted for with similar mechanisms. Evidence that native speakers are sensitive to gradient phonotactic patterns is presented in Treiman et al. (1996). They found that nonsense monosyllabic words with more probable VC sequences were judged as more wordlike than words with improbable, but possible VCs.

FBP note that no special linguistic faculty is needed to model OCP effects. The OCP based on perceived similarity is a cognitively grounded linguistic universal. The parameters which determine the sharpness of the effect of similarity and the effects of distance are fixed on a language particular basis. These changes in strength can be modeled directly with the stochastic constraint model, repeated in (55). Arabic is one of the strongest cases of consonant cooccurrence restrictions, the coefficient of sharpness *S* in FBP’s model is 27.5, creating a clear boundary between highly restricted and highly unrestricted consonants in the adjacent data.

(55)

$$\mathrm{O/E} = \frac{A}{1 + e^{K + Sx}}$$

In the stochastic constraint model, acceptability is an abstract measure of the goodness of a form. The best forms are maximally acceptable, but less acceptable forms may also be found. The constraint is probabilistic: less acceptable forms are found less frequently than more acceptable forms. A form is considered unacceptable if it is so improbable that less than one such form is expected to be found in the lexicon.

I showed above that different strengths of constraints can be modeled by altering the intercept and sharpness parameters, *K* and *S*, and a categorical effect can be created by using extremely large sharpness values, causing the logistic to become near vertical. The parameter *K* determines the point where the 'category boundary' is located, the midpoint of the logistic (where acceptability is 0.5) is found at $x = -K/S$. For a language with a cooccurrence restriction only on identical consonants, a large negative *K* places the category boundary toward similarity one.

**Figure 5.6**: Effects of changes in the intercept (K) on the stochastic constraint boundary.

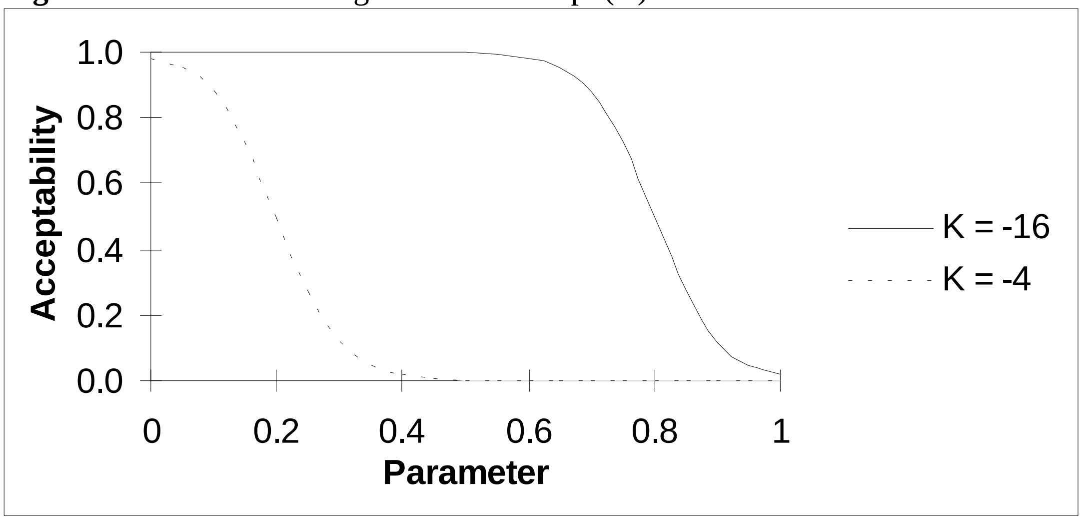
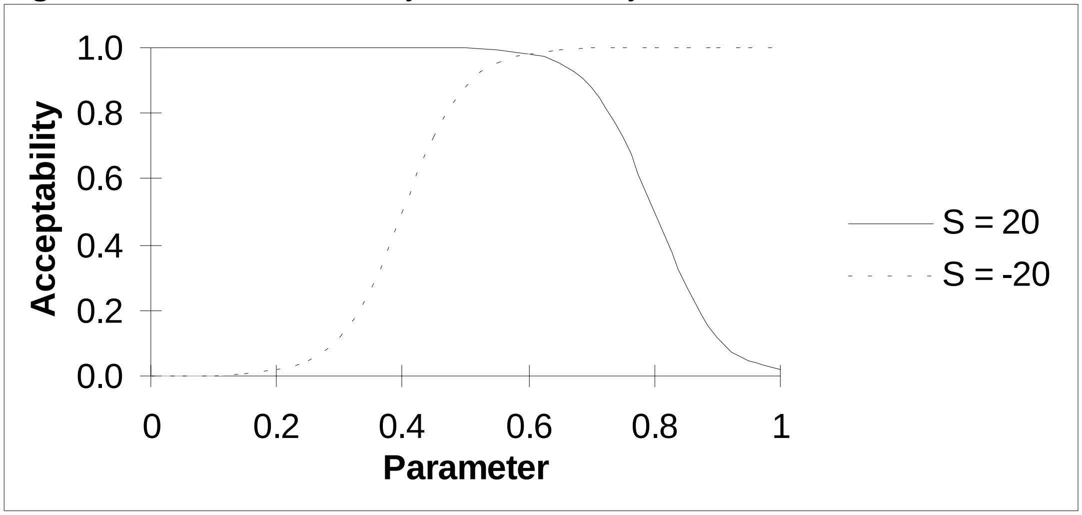



Figure 5.6 shows logistics with identical sharpness (S= 20), with two different intercepts (K = -4 and K = -16). The curve with *K* = -16 is right-shifted from the curve with *K* = -4. If these were two models of two different similarity effects, all but the most similar values are acceptable for the right-shifted curve. A model of this type is appropriate for languages which prohibit identical consonant pairs, but otherwise allow any consonant combination.

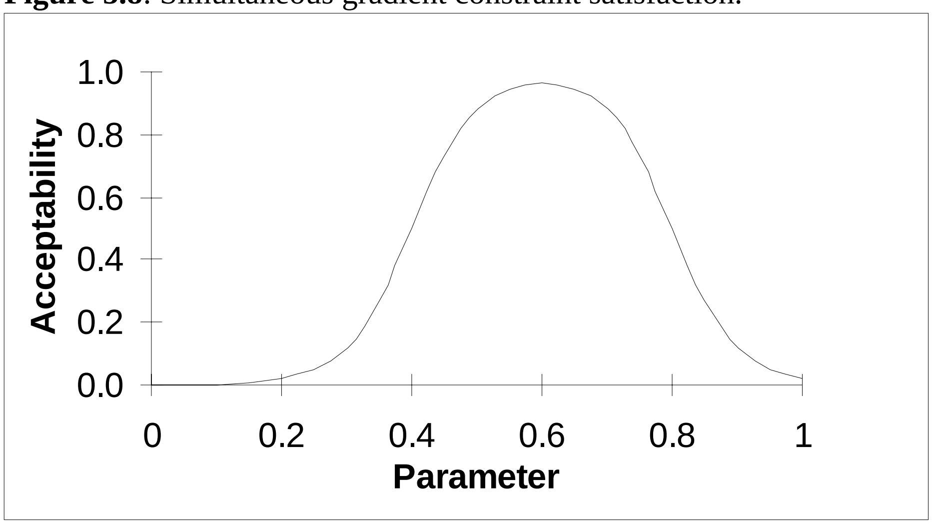
FBP point out that constraints enforcing similarity, for example vowel harmony, can be modeled via the stochastic constraint, by changing the signs on the parameters *K* and *S*. Figure 5.7 shows a stochastic dissimilarity constraint and a comparable stochastic similarity constraint. The curve on the right is a dissimilarity constraint. It has high acceptability for low values of similarity (*K* = -16 and *S* = 20). The curve on the left is a similarity constraint. It has high acceptability for high degrees of similarity (*K* = 8 and *S* = -20).

**Figure 5.7**: Stochastic similarity and dissimilarity constraints.




In the case of parasitic vowel harmony (e.g. Cole & Trigo 1989), discussed by FBP, there is an optimal degree of similarity which is desired, intermediate between maximal and minimal similarity. Such an effect can be modeled by invoking simultaneous stochastic constraints on similarity and dissimilarity. Multiplying the acceptability values of two stochastic constraints produces a simultaneous acceptability measure, shown in Figure 5.8. This has the desired result for parasitic vowel harmony: intermediate similarity provides optimal acceptability. Multiplying stochastic constraints is equivalent to the fuzzy set theory operation of set intersection. This model of constraint combination is implemented for Arabic OCP-Place in chapter 7.

**Figure 5.8**: Simultaneous gradient constraint satisfaction.




Note that, when two constraints are combined, as in Figure 5.8, no form is perfectly acceptable (acceptability = 1). Rather, some forms are relatively more acceptable than others. The stochastic constraint model thus shares the same conception of relative acceptability as

Optimality Theory (Prince & Smolensky 1993). Many occurring forms may be sub-optimal. In the stochastic constraint model, the most acceptable forms are the most frequently occurring forms. It is possible to have forms which violate a constraint to be of high frequency, if they are the most acceptable forms possible.

FBP argue that the consideration of similarity provides crucial understanding of the phenomena behind phonotactic constraints. Strong phonotactic constraints, which are often categorical, are found within adjacent consonants in consonants clusters. Perceived similarity is least diminished in these cases as the consonants are truly adjacent, and similarities and differences between them are most salient. However, the acceptability of similar pairs is language particular, with some languages allowing clusters of highly similar consonants, and other languages banning clusters entirely. Languages which allow no clusters have the strongest constraint against similarity between adjacent segments, such that adjacent phonemes must be vowels and consonants, the most dissimilar segments (cf. Ohala 1990, Lamontagne 1993). Gradient constraints like the OCP effects in Arabic are found over a greater distance, for example over intervening vowels, which reduces the strength of such effects considerably.

In the discussion above, the stochastic constraints modeled acceptability (on the *y*-axis) against similarity (on the *x*-axis). However, the logistic model as a stochastic constraint can be applied to model acceptability of a form to any parameter *x*. The stochastic constraint model can be applied in general in a phonological theory which utilizes gradient constraints. Incorporating these constraints, for example, into a current constraint based phonological framework like Declarative Phonology merely requires replacing the use of categorical constraints with stochastic constraints. However, doing so requires adopting the Optimality Theoretic perspective that occurring forms can violate constraints, and that the most acceptable forms may violate many constraints. Categorical effects can still be captured, using stochastic constraints with a near vertical category boundary, but smoothly gradient phenomena can also be accounted for. I return to the implications of the stochastic constraint model and gradient data for phonological theory in chapter 11.

# CHAPTER 6
# The Special Status of Word Onsets in Speech Errors

In this chapter, I review research on the effects of syllable and word position on production, by examining evidence from phonological segment errors. Errors between consonants in structurally parallel positions are as much as ten times more likely than errors between consonants in different structural positions (Dell 1984, Shattuck-Hufnagel 1992). I present the results of a new experiment, which shows that errors between similar consonants in word onset position are more likely than errors between similar consonants in other positions. For dissimilar consonants, no effect of word position is found. This pattern is predicted by a spreading activation or connectionist model of phonological encoding, as I discuss in section 6.3. This result is significant for three reasons. First, word onset errors are by far the most frequent in naturally occurring interaction errors. The higher error rate among similar word onsets than among similar consonants in other positions accounts for their preponderance in error analyses. Second, this demonstrates that there is differential behavior for consonants in different positions, and thus that future research should be aware of possible positional idiosyncracies from examining only word onset errors. Third, in the next chapter, I show that analogous effects of word onset position on similarity are found in OCP-Place. I claim that finding parallel effects in both speech production and phonotactics suggests that there is a deeper connection between processing in the lexicon and abstract linguistic constraints.

## *6.1 Phonological Structure in Phonological Speech Errors*

Recall that similar segments (e.g. p/f) are typically mis-selected for one another in both naturally occurring errors and those elicited in experiments (e.g. Fromkin 1971, Levitt & Healy 1985, Nooteboom 1969, MacKay 1970, Motley & Baars 1975, Shattuck-Hufnagel & Klatt 1979). While similarity is an important factor in speech errors, there are additional constraints on which consonants in an utterance are likely to interact. Consonantal speech errors usually involve the interaction of an onset with an onset or a coda with a coda (MacKay 1969). Examples are from Fromkin (1971):

(56) Onset errors:
- a. [alʃo] share (also share)
- b. reek long race (week long race)

Coda errors:
- c. sub... such observation
- d. great wrist (great risk)

There is additional evidence that other aspects of phonological structure are relevant for speech errors. Shattuck-Hufnagel (1987) examined the frequency of errors in word initial position and in stressed syllables in the MIT-Arizona corpus of naturally occurring errors. She found that most errors occurred in word onsets, even though word onset positions are less frequent than other syllable onset positions in the corpus. In addition, stressed syllables were disproportionately involved in errors.

Since many errors in natural error corpora occur in the onset of stressed monosyllabic words, differences in the effects of word position and stress in error rates is not clear. Shattuck-Hufnagel (1992) performed a series of experiments to determine whether there are word onset effects or stress effects or both which influence error rates. She found both that shared word onset position between two consonants increased their propensity to interact in an error, and that shared stress between two syllables also increased error rates between the onsets of those syllables. I review Shattuck-Hufnagel's experiments in some detail, in part because I use her paradigm in my own experiment on interaction between the effect of word position and similarity on error rate.

Shattuck-Hufnagel recorded subjects reading and reciting tongue twisters containing similar consonants which shared prosodic position to different degrees. All of the stimuli were based on two bisyllabic words and two monosyllabic words. The stimuli were created in sets of four, with the four stimuli in each set based around the same onset consonants. In the example stimuli given below, the onset consonants are /p/, /f/, and /r/. Since /p/ and /f/ are similar, these phonemes are the target pair for the four stimuli in the example set. The other onset /r/ was chosen to be dissimilar from the target pair.

The first type of stimulus placed the target phonemes in stressed word onset position. Shattuck-Hufnagel refers to these as the 'Both Same' stimuli (B), since the target phonemes have both the same word position and the same stress position.

(57) Type 1: B(oth same)
peril fad foot parrot

Each stimulus in the set of 4 used the same monosyllabic words. The bisyllabic words were varied, to vary the amount of prosodic parallelism between the onset of the monosyllable (which is a stressed word onset) and a similar onset in the bisyllable. The second type of stimulus, the 'Word Same' stimuli (W) have target phonemes which are both word onsets, but with different stresses.

(58) Type 2: W(ord same)
parade fad foot parole

The third type of stimulus, called the 'Stress Same' stimuli (S) have target phonemes which are both stressed, but which differ in word position.

(59) Type 3: S(tress same)
repeat fad foot repaid

The final type of stimulus, the 'Neither Same' stimuli (N) have target phonemes which are both onsets, but which differ on word position and stress.

(60) Type 4: N(either same)
ripple fad foot rapid

There were 24 sets of consonants, with 4 stimuli per set, for a total of 96 twisters. A total of 20 subjects participated in the experiment. Each subject read half of the twisters for 20 of the consonant sets, and all of the twisters for 2 of the sets. The twisters were presented individually on index cards. The subjects read each twister out loud three times, and then recited the twister from memory three times. Thus, each subject produced 1,152 words, for a total of 23,040 words produced in the experiment.

Shattuck-Hufnagel (1992) tabulated errors between the target segments (e.g. p/f). Her results, collapsed across speakers, are presented in Table 6.1. Phonemes in parallel prosodic position are more likely to interact in speech errors. In the first row (B), p/f are both in stressed word onset position, and many errors occurred. In the second row (W), /p/ is in unstressed word onset position, while /f/ is a stressed word onset. In the third row (S), /p/ is stressed, but is not a word onset. In the last row, the phonemes differ in both word and stress position. Clearly, structural parallelism increases error rate.

**Table 6.1**: Results of Shattuck-Hufnagel (1992) Experiment 1.

| Stimulus type (e.g. p/f) | Prosodic positions shared | Errors |
|---|---|---|
| parrot fad foot peril | word onset, stressed (B) | 182 |
| parade fad foot parole | word onset (W) | 130 |
| repeat fad foot repaid | stressed (S) | 55 |
| ripple fad foot rapid | none (N) | 14 |

Shattuck-Hufnagel (1992) replicated this result in two other experimental conditions. In Shattuck-Hufnagel (1992) Experiment 1a, the word lists were embedded in nonsense phrases. The phrase condition was intended to test the results of Experiment 1 in a more natural speech task. The examples above are shown here in their corresponding phrases.

(61) Type 1: B(oth same)
The parrot is a fad and the foot is in peril.

(62) Type 2: W(ord same)
The parade is a fad and the foot is on parole.

(63) Type 3: S(tress same)
To repeat is a fad and the foot is repaid.

(64) Type 4: N(either same)
The ripple is a fad and the foot is rapid.

In Shattuck-Hufnagel (1992) Experiment 2, a partially different set of bisyllabic words was used. These words were given in lists of the form ‘monosyllable bisyllable bisyllable monosyllable’. Additionally, different monosyllables were used for each type of stimulus for a particular phoneme pair. Sample stimuli are:

(65) Type 1: B(oth same)
pack fussy fossil pig

(66) Type 2: W(ord same)
pad forsake foresee pot

(67) Type 3: S(tress same)
pin suffuse suffice pet

(68) Type 4: N(either same)
pod sofa suffer peg

Experiments 1a and 2 from Shattuck-Hufnagel (1992) support the results of Experiment 1. In all three experiments, consonants with parallel phonological structure had more interaction errors than consonants which had some structural mismatches. Results for all three experiments are shown together in Table 6.2. Note that in all of the errors in this experiment, there was interaction between a stressed word onset (e.g. /f/, the onset of the monosyllable) and another onset, so all errors involved at least one word onset.

**Table 6.2**: Results of Shattuck-Hufnagel (1992) Experiments 1, 1a, and 2.

| | | #Errors | | |
|---|---|---|---|---|
| Stimulus type (e.g. p/f) | Prosodic positions shared | Exp. 1 | Exp. 1a | Exp. 2 |
| parrot fad foot peril | word onset, stressed (B) | 182 | 202 | 253 |
| parade fad foot parole | word onset (W) | 130 | 134 | 132 |
| repeat fad foot repaid | stressed (S) | 55 | 59 | 75 |
| ripple fad foot rapid | none (N) | 14 | 15 | 26 |

In addition to tabulating errors between the similar target segments, Shattuck-Hufnagel (1992) also tabulated errors between the dissimilar segments (e.g. /r/ with /f/ or /p/ in the sample stimuli). The dissimilar segments are in the same set of structural relations as the similar ones. In *parrot fad foot peril* (where p/f are in the B configuration), /r/ and /f/ share no prosodic structure other than syllable onset. Thus, an error between /r/ and /f/ would be of type N. Table 6.3 shows the complete set of corresponding error types.

**Table 6.3**: Error types tabulated in Shattuck-Hufnagel (1992).

| Stimulus type | Similar error type (p/f) | Dissimilar error type (r/f) |
|---|---|---|
| parrot fad foot peril | word onset, stressed (B) | none (N) |
| parade fad foot parole | word onset (W) | stressed (S) |
| repeat fad foot repaid | stressed (S) | word onset (W) |
| ripple fad foot rapid | none (N) | word onset, stressed (B) |

Table 6.4 shows the number of errors for dissimilar phonemes in the three experiments. The basic trend found above is preserved: stimuli where more prosodic structure is shared have higher error rates. Note, however, that the dissimilar B and W cases are not as different from one another as the similar B and W cases are. To summarize, Shattuck-Hufnagel (1992) demonstrated that shared prosodic structure has a strong effect on error rate.

**Table 6.4**: Errors between dissimilar consonants in Shattuck-Hufnagel (1992) Expt. 1, 1a, and 2.

| Stimulus type (e.g. r/f) | Prosodic positions shared | Expt.1 | Expt.1a | Expt.2 |
|---|---|---|---|---|
| ripple fad foot rapid | word onset, stressed (B) | 24 | 45 | 58 |
| repeat fad foot repaid | word onset (W) | 25 | 44 | 21 |
| parade fad foot parole | stressed (S) | 11 | 14 | 4 |
| parrot fad foot peril | none (N) | 0 | 0 | 1 |

Other aspects of structural parallelism have been shown to increase error rate as well. Spontaneous error corpora show a tendency for phoneme exchanges (e.g. *like box* for *bike locks*) to occur when the phoneme adjacent to the exchanging ones are identical (Nooteboom 1969, MacKay 1970). In a series of experiments Dell (1984) explored what he terms the REPEATED PHONEME EFFECT: the effect of repetition of the nuclei or the codas of two syllables on error rates between the onsets of those syllables. Dell conducted experiments using the SLIPS paradigm (Motley & Baars 1975).

In SLIPS, word pairs are presented to subjects at a rate of about one pair per second. Immediately after the presentation of some pairs, a cue is given that the subject is to say the word pair as quickly as possible. Certain cued pairs, called the critical cued pairs, are preceded by pairs which bias the subject to produce an exchange error. These interference pairs prime the subject say the word onset consonants of the critical pair in reverse order. For example, the subject could see a sequence like the following:

(69) sad sack (primes correct nucleus and coda)
bid meek (interference pair)
bud meek (interference pair)
big men (interference pair)
mad back (critical pair)

In addition to the interference and critical pairs, there are filler pairs between the stimulus blocks, some of which are cued, that are used to prevent the subject from noticing the priming pattern.

Dell constructed SLIPS stimuli which controlled the vowels and coda consonants to determine if shared vowels or shared codas increased error rate. He examined the effect of shared vowels in Dell (1984) Experiment 1. He contrasted stimuli with repeated vowels in the two words (e.g. mad back, made bake) and stimuli which did not repeat vowels (e.g. made back, mad bake). There were twice as many errors between word onset consonants when vowels were repeated. The error rate was 11.0% for onsets before repeated vowels, and 5.4% when the vowels following the onsets were not repeated.

Dell (1984) Experiment 2 examined the effect of shared coda consonants. In this experiment, cued pairs with repeated coda consonants (e.g. boot coat, boom comb) were contrasted with cued pairs which did not have repeated coda consonants (e.g. boom coat, boot comb). Error rates between word onsets were again higher when other phonemes were repeated. The error rate between word onsets was 6.3% when final consonants were repeated and only 2.4% when they were not.

Together Dell (1984) and Shattuck-Hufnagel (1992) show that error rates are highly dependent on structural parallelism. The presence of shared prosodic context or shared segmental context influences error rates.

## *6.2 Experiment 1: Word Onsets in Phonological Speech Errors*

Word onsets are disproportionately involved in speech error analyses as they are the most frequent errors found in naturally occurring corpora (Dell & Reich 1981, Garrett 1980, Shattuck-Hufnagel & Klatt 1979, Stemberger 1983), and word onset errors are easy to induce experimentally (Levitt & Healy 1985, Motley & Baars 1975, Shattuck-Hufnagel 1987). We might suspect that word onsets are inherently error prone. In the previous section, I reviewed Shattuck-Hufnagel (1992), which showed that parallel prosodic position increases error rates. No study, to my knowledge, has shown that stressed word onsets are inherently more likely to err than other syllable onsets.

I have undertaken a follow-up to Shattuck-Hufnagel's experiments in which all consonant pairs are in fully parallel prosodic positions, but the prosodic position is varied across stimuli. Thus, all stimuli shared prosodic structure equally, but vary as to the exact position under examination. I have found that word onsets are inherently more susceptible to error when they are similar, but not when they are dissimilar.

### *6.2.1 Materials.*

Each stimulus in the experiment is based on a pair of similar consonants (e.g. b/p) and a pair of dissimilar consonants (e.g. k/s) appearing in syllable onset position. There were twelve quadruplets of consonants. Each quadruplet was used in a set of four tongue twisters, where the pairs of similar and dissimilar consonants were placed in structural correspondence. Each twister consisted of four bisyllables. The first type of tongue twister had the similar consonants in stressed word onset position, and the dissimilar consonants in unstressed, second syllable onset position.

(70) Type 1: beacon possum piercing bookie
(b/p are stressed word onsets, k/s are unstressed second syllable onsets)

This stimulus roughly corresponds to the B stimulus from Shattuck-Hufnagel's experiments. The second type of tongue twister had the similar consonants in the unstressed word onset position, and the dissimilar consonants in the stressed second syllable onset position.

(71) Type 2: became pursue percent bouquet
(b/p are unstressed word onsets, k/s are stressed second syllable onsets)

The third type of tongue twister had the similar consonants in the stressed second syllable onset and the dissimilar consonants in the unstressed word onset.

(72) Type 3: caboose support suppose kebab
(b/p are stressed second syllable onsets, k/s are unstressed word onsets)

The fourth type of tongue twister had the similar consonants in the unstressed second syllable onset and the dissimilar consonants in the stressed word onset.

(73) Type 4: cabin soapy supper cobble
(b/p are unstressed second syllable onsets, k/s are stressed word onsets)

I constructed 36 such twisters, one of each type for each quadruplet of consonants. These 36 twisters were reordered (e.g. beacon possum piercing bookie → possum beacon bookie piercing) to create a second set of 36 twisters. Each tongue twister was printed on a three-by-five index card in 18 point type with a laser printer. The complete set of stimuli is shown in section 6.4, at the end of the chapter.

### *6.2.2 Method.*

The stimuli were given to subjects in two blocks of 36, each block contained one of the two orderings for a particular set of words. Each block was randomized so that no adjacent stimuli had the same similar phoneme pair or the same structural position for the similar phonemes. All subjects read the stimuli in the same order.

Following the procedure of Shattuck-Hufnagel (1992), the subjects were instructed to read each stimulus three times out loud, and then to repeat the stimulus from memory three times. Shattuck-Hufnagel (1992) found the error rate to be higher during the repetition portion, but that the overall pattern of errors was the same in both the reading portion and the repeating portion. I do not differentiate errors while reading from errors while repeating in the analysis below. To avoid outright memory errors, the subjects were allowed to look at the stimuli between repetitions if they became uncertain about what the words were.

Subjects were recorded one at a time in a sound controlled booth on a Marantz cassette recorder using a Sony electret lavaliere microphone. The equipment made a very high quality recording, and the electret microphone was hopefully less obtrusive than a full-sized microphone, resulting in more natural speech behavior. The subjects were instructed that the experiment was concerned with the mistakes that they made, and thus that it was important for them to speak at a normal to slightly fast rate. The subjects were given a five minute break and distracted with ordinary conversation between stimulus groups. The entire recording procedure generally took less than 45 minutes.

#### *6.2.3 Tabulation.*

The cassettes were later analyzed using Sony professional headphones. Errors that were unambiguously on the target word onset consonants were tabulated. Errors on larger units, or that were interrupted before a clear determination of the error unit could be made, were not counted. I tabulated consonant exchange errors, e.g. *peacon bossum* for *beacon possum*, separately. Dell & Reich (1981) present evidence that the first error in the exchange is the primary error, and the second error is not an independent error. The exchange errors are thus presented below based on the status of the first error in the pair. Multiple errors on the same word in different repetitions of the stimulus were counted as separate errors long as a correct repetition of that word occurred between the errors. Thus, it was possible for as many as three errors to be counted for a single word on a single stimulus over the six repetitions.[16]

#### *6.2.4 Subjects.*

The subjects were 33 Northwestern undergraduates, primarily from introductory classes in linguistics and cognitive science. The data from one subject were unusable due to technical difficulties, and the data from a second subject were removed from the study as the subject had a conspicuously non-native accent which was not detected during the pre-experimental screening. The remaining 31 subjects were monolingual speakers of American English. Subjects were paid for their participation in the experiment.

#### *6.2.5 Results and discussion.*

Table 6.5 presents the results of the experiment. The second row of data in the table is shaded as it involves an extra complication which I discuss below. Setting the unstressed word onsets aside, the error rate between similar stressed word onsets is higher than the error rate between similar second syllable onsets is significant on an ANOVA with stimulus type and subject as factors ($F(2,30) = 3.05$, $p = 0.054$).

---

[16] This method was chosen as subjects occasionally adopted an error as the correct word in the stimulus and repeated the error for the remaining repetitions of the stimulus. The repeated errors which I did include do not significantly affect the results presented below. All results are the same if they are left out of the analysis.

**Table 6.5**: Results of Experiment 1.

| Stimulus type (e.g. b/p) | Prosody shared | #Similar | #Dissimilar | #S/#D |
|---|---|---|---|---|
| beacon possum piercing bookie | word onset, stress | 124 | 30 | 4.1 |
| became pursue percent bouquet | word onset, no stress | 226 | 163 | 1.4 |
| caboose support suppose kebab | $\sigma_2$ onset, stress | 89 | 31 | 2.9 |
| cabin soapy supper cobble | $\sigma_2$ onset, no stress | 79 | 27 | 2.9 |

The difference in error rate between similar segments in different positions is not significant with stimulus type and consonant pair as factor ($F(2,8) = 0.96$, ns). The lack of significance with consonant pair as factor can be blamed on the small difference between the two second syllable onset types. Additionally, there appear to be particular consonant pairs which do not reflect the overall trend. Note first that nearly half of the errors between stressed second syllable onsets involved the similar consonant pair l/r (in two different quadruplets of stimuli). If this pair is set aside, the difference between stressed word onsets and stressed second syllable onsets across stimuli is significant on a paired *t*-test ($t = 2.2$, $p = 0.033$). In the case of unstressed second syllable onsets, the similar pair p/f do not behave like the rest of the group. This pair is distinct from the rest as it is the only manner of articulation contrast among the similar consonant pairs. If this consonant group is set aside, the difference between the remaining stressed word onsets and unstressed second syllable onsets is significant on a paired *t*-test ($t = 1.8$, $p = 0.056$). Whether the exceptional patterns just discussed are accidental or principled, and whether if principled they can be given a satisfactory account, is left as an open question for future work.

The unstressed word onsets, which have a very high error rate in Experiment 1, involve the added complication of the repeated phoneme effect (Dell 1984), discussed in section 6.1 above. In these stimuli, the consonants are word onsets of iambic words. As a result, all of these syllables had reduced, /ə/ vowels. Dell found that sharing a following vowels increased error rates between word onsets by 150% to 200%. The repeated phoneme effect accounts for the difference between similar stressed and unstressed word onset error rates. The 226 similar unstressed word onset errors is between the predicted 186 to 248 errors based on the similar stressed word onsets (150-200% of the stressed word onset error count). The extremely high error rate for dissimilar unstressed word onsets cannot be accounted for so straightforwardly. Another possible source of errors in these stimuli is whole syllable errors. Given the reduced vowels in the unstressed word initial syllables, a syllable error is indistinguishable from an onset consonant error. Thus, the error tabulation for this type of stimuli may include more than one type of error.

The overall effect of word position on similarity in Experiment 1 is also found in the MIT-Arizona corpus of naturally occurring no-source errors, which was presented in chapter 4. The 521 single segment consonant errors in this corpus consist of 209 word initial errors and 312 word medial or word final errors. Using the similarity values for English consonant pairs presented in chapter 3, the mean similarity for word initial errors in this corpus is 0.36 (with variance 0.018), while the mean similarity for word medial and word final errors in this corpus is 0.33 (with variance 0.024). The mean similarity is significantly less in the non-initial case by a *t*-

test ($t$ = 2.01, $p$ = 0.023). In addition, the difference in variance is significant ($F$-test, $F$ = 1.36, $p$ = 0.009). The lower mean similarity and greater variance for the non-initial case indicates that the non-initial no source errors cover a broader range of similarity values that is more evenly distributed than the initial errors. The initial errors show greater mean similarity and less variation in similarity.

In this section, I have shown that, while structural parallelism increases the error rate between consonants, the word onset position has special status compared to the onset of the second syllable in a bisyllable. Similar consonants in the word onset are inherently more likely to interact than similar consonants elsewhere in the word. This result may account, in part, for the large number of word onset errors in naturally occurring error corpora. Since most words have onsets and interactions between word onsets are especially likely to happen, the large number of word onset errors in natural corpora are accounted for.

### *6.3 Word Onsets in an Activation Model of Phonological Encoding*

Current linguistic models of speech production are based on evidence from speech errors. Any correct model of speech production must be able to account for the many regular patterns of errors which have been established in the literature. In this section, I argue that the extra sensitivity to similarity of word onset consonants in speech errors is the result of the sequential access of segmental information within the word. Thus, I claim that the procedure for phonological encoding for speech production accesses the segmental contents of words sequentially from left-to-right. In this model, words which are phonologically similar to the target word are also activated.

In the Dell (1986) model of phonological encoding presented in chapter 6, the phonemes of a syllable are encoded all at once, based on which onset, nucleus, and coda are most highly activated. Sevald & Dell (1994) found experimental evidence against the atemporal aspect of the model. They propose a sequentially cued model, where segmental activation proceeds from left-to-right within the word. I first review their evidence for sequential segmental encoding, then present an account of the greater strength of similarity effects for word onsets based on the Sevald & Dell (1994) model.

Sevald & Dell (1994) found evidence for sequential cueing by testing the effect of repetition on production time. They counted the number of repetitions of sequences of four CVCs subjects could make in eight seconds. Using $C_iVC_f$ sequences, they tested the effects of repeating $C_i$, V, $C_f$, $C_iV$, $VC_f$, and the entire $C_iVC_f$ sequence on production time. Stimuli that were four words each were constructed with four different patterns of repetition. ‘Immediate repetition’ involved repeating a unit in all four CVCs (AAAA pattern). ‘Near repetition’ followed an ABBA pattern. ‘Far repetition’ used an ABAB pattern. ‘Nonrepetition’ (ABCD pattern) was only used for levels of structure above the segment to constrain the size of the stimulus space. Examples of each type of repetition on a $C_iV$ sequence, in conjunction with a far repetition of $C_f$, are:

(74) Immediate repetition of $C_iV$ (AAAA):
pick pin pick pin

(75) Near repetition of $C_iV$ (ABBA):
pick tin tick pin

(76) Far repetition of $C_iV$ (ABAB):
pick ton pick ton

(77) Nonrepetition of $C_iV$ (ABCD):
pick ton tick pun

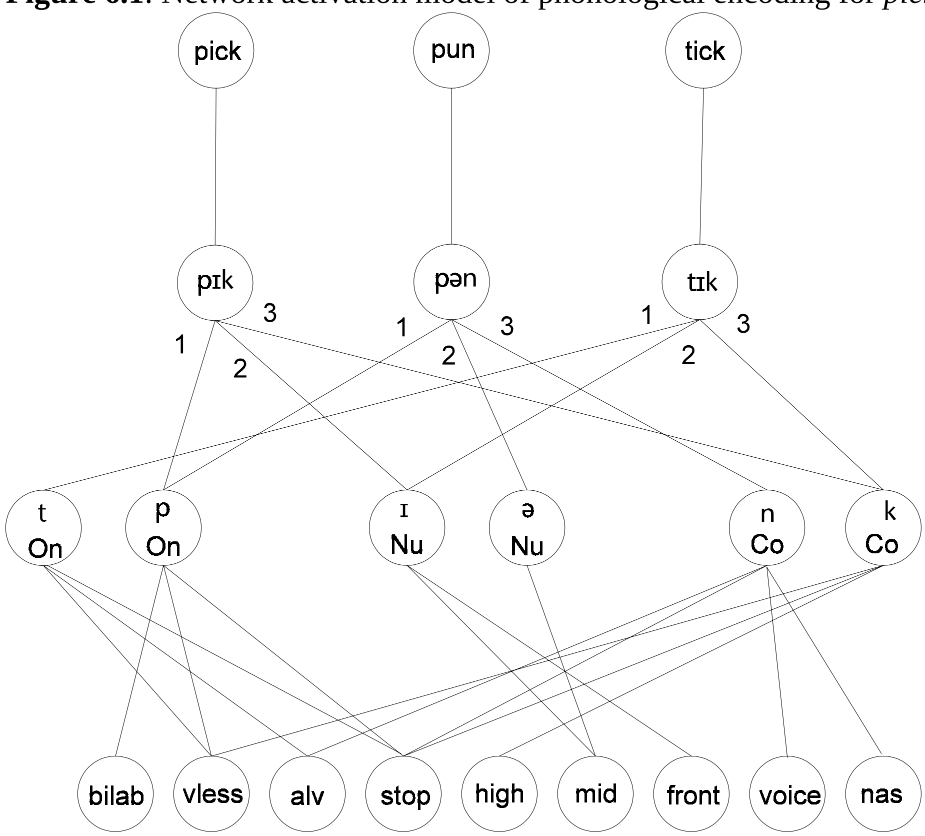
The stimuli contained all possible combinations of repetition types over the segments ($C_i$, V, and $C_f$) for two completely different sets of segments. When taken together, individual repetition patterns for segments create the patterns of $C_iV$, $VC_f$, and $C_iVC_f$ repetition. The first example involves immediate repetition of $C_i$ and V, giving immediate repetition of $C_iV$. The second example has near repetition of $C_i$, and immediate repetition of V, resulting in near repetition of $C_iV$. The third example has far repetition of $C_i$ and V, resulting in far repetition of $C_iV$. The fourth example has near repetition of $C_i$ and far repetition of V, resulting in no repetition of $C_iV$.

Sevald & Dell (1994) found that repetition of the entire $C_iVC_f$ was the most significant factor in reducing production time. The only other factor which significantly reduced production time was $C_f$ repetition, in other words repeating the coda. By contrast, they found that repeated $C_i$ and $C_iV$ significantly *increased* production time. In general, the inhibitory or facilitative effects of repetition were enhanced by the nearness of the repetition. They concluded that a production model which employs simultaneous activation of onset, nucleus, and coda segments cannot account for the fact that repeated $C_f$ facilitates production, while repeated $C_i$ inhibits production. Sevald & Dell (1994) proposed instead that $C_i$ is activated first, and $C_f$ is activated later. The sequential activation of $C_i$ and $C_f$ means that there are different context effects for the selection of $C_i$ and the selection of $C_f$.

Recall that the Dell (1986) model involved spreading activation across nodes representing morphemes, syllables, rimes, and cluster or singleton onsets, nuclei, and codas. Activation spreads from the initially activated morpheme node, and the onset, nucleus, and coda of a syllable is selected after an interval of time has passed. Since the time from activation of the morpheme node to the selection of the segments is a fixed number of time units, this model cannot be used to model differences in production times in its current form. Instead, Sevald & Dell (1994) assume that competition between segments delays encoding and production. Situations with greater competition between the intended segment and other segments are slower to be produced than situations without competition. With sequentially cued segments, repetition of $C_i$ creates competition between possible completions of the word. Other words which are soon to be produced, and also begin with $C_i$, become activated and compete with the target word. The repetition of $C_f$ creates no competition, as $C_f$ is the last segment encoded in the word. Thus, repetition of $C_i$ slows production through inhibition, while repetition of $C_f$ facilitates production because it receives additional activation that does not create competition.

To understand the effects of sequential activation on repeated forms, compare the production of *pick pun* (with repeated $C_i$) and *pick tick* (with repeated $VC_f$). Figure 6.1 shows a representation of the network including *pick*, *pun*, and *tick*. When producing *pun* in *pick pun*, the nodes for /p/, /ə/, and /n/ are activated in that order. Activation spreads from the node for *pun* to the /p/ node. Activation from the /p/ spreads back to *pun* but also to all other words that start with /p/. Since the node for *pick* was recently activated, it has some residual activation which is enhanced by the new activation from the /p/ node. The reactivation of *pick* creates competition between the proper encoding of the rest of *pun* and repeating *pick*. This competition delays proper encoding, and slows production. By contrast, when producing *tick* in *pick tick*, the activation spreads from the node for *tuck* to the /t/ node. Since the /t/ node does not spread activation to *pick*, *pick* is not reactivated, and no extraordinary competition results. Finally, consider the production of *pick pick*. In this case, activating the second *pick* spreads activation to nodes which are already activated from the preceding production of *pick*. In this case, the result is pure benefit, leading to the fastest encoding of all.

**Figure 6.1**: Network activation model of phonological encoding for *pick*, *pun*, and *tick*.




Sevald & Dell (1994) tested the prediction of sequential cueing in a second experiment. They examined subject's speed of production for pairs of CVC words with repeated initial or final segment sequences. In this experiment, there were four types of stimuli:

(78) repeated $C_i$:
cat cub

(79) repeated $C_iV$:
cat cab

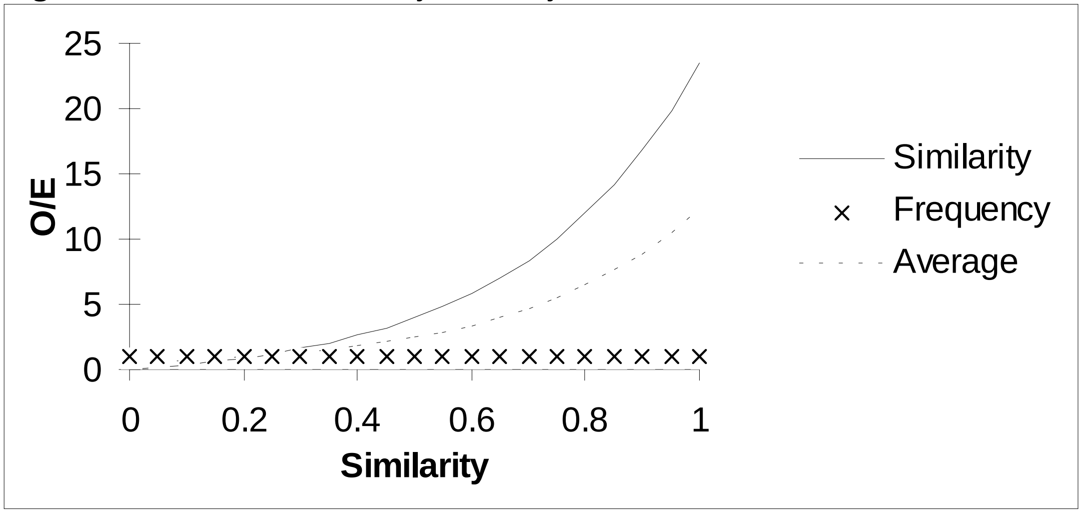
(80) repeated $VC_f$:
cat bat

(81) repeated $C_f$:
cat but

Sequential cueing predicts that repeated final segments should facilitate production while repeated initial segments should delay production. Also, repeating two segments should affect production more than repeating one segment, since two repeated segments spread more activation than one. The subject's mean production times were consistent with these predictions. To summarize, Sevald & Dell (1994) showed that segment repetition had different effects depending on whether the segment was initial or final, and accounted for the difference through sequential cueing.

An analogous difference was found for similarity influences on speech errors in Experiment 1. Initial segment errors were more heavily influenced by similarity than later segment errors. I claim that this is also a result of sequential cueing. The production of non-initial consonants occurs after the activation of the previous segments. This leftward context spreads activation to the entire cohort of words which begin with that leftward context. Thus, all of the possible completions of the words compete with the intended completion. In this case, the similarity of the competing segments to the intended one plays a lesser role, as all possible completions are activated by the leftward context regardless of their similarity to the target. In the case of word onsets, however, there are no such leftward context effects. In this case, the only plausible candidates for competition are ones which are similar to the intended segment, which would be activated by feature nodes which are shared between the target and the competing segments. The effect of context is represented graphically in Figure 6.2. The solid line in Figure 6.2 shows the relative error rate (O/E) by similarity, based on the regression model presented in chapter 4. This line can roughly represent the effect of the activation only by similarity, since the majority of errors in naturally occurring corpora are word initial. The *X*s represent the predicted relative error rate by context. Since the likelihood of a consonant being one of the activated completions is on average the consonant's frequency, the predicted O/E is one for the context effects. Assuming an intrusion is equally likely to occur because of activation due to frequency or due to context the predicted non-initial error rate is the average of these two curves, shown as a dashed line in Figure 6.2.

**Figure 6.2**: Relative error rate by similarity, context, and their combination.




Consider, for example, the production of the word *peril*. In the Sevald & Dell (1994) version of the Dell (1986) model, activation of the word node for *peril* activates the first syllable node, which in turn activates the /p/ node. Activation in the /p/ node spreads downward to all of the feature nodes connected to the /p/ node, and upward to all other syllables which begin with /p/. Spreading activation back down from the other syllables does not cause competition for /p/, since all of the syllables begin with /p/. Spreading activation back up from the feature nodes activates all phoneme nodes which share features with /p/, and the nodes which share the most features with /p/ receive the most activation. Thus, any potential competition with /p/ within the first few cycles of activation is similar to /p/. Now suppose /p/ and /ɛ/ have been encoded, and /r/ is next. The activation of the /p/ and /ɛ/ nodes spreads activation to all words that begin with /pɛ/, like *peril*, *pelican*, *petrify*, *pebble*, etc. Thus, the encoding and production of /r/ has to compete with many segments from other possible completions of the word. Similar segments should still be more competitive since they receive the extra activation of the shared feature nodes, but the similarity effect should be weaker. This is exactly the pattern found in Experiment 1.

In conclusion, I have shown in this section that similarity effects on word onset errors can be attributed to phonological competition in a model where segments are encoded sequentially for production as in Sevald & Dell (1994). Word onsets have special status as they have no left hand context within the word, and so are subject to intrusion only by segments which share a number of features (i.e. which are similar). Other segments have some left hand context which helps to promote dissimilar errors by activating other words in the target word's cohort. The organization and encoding of the lexicon for production is thus similar to the cohort model used for perception (Marslen-Wilson 1984). In the next section, I show that sequential access of segmental information has influences beyond production and perception.

*6.4 Stimuli Used in Experiment 1*

| Target pair: b,p | | Non-target pair: k,s | | |
|---|---|---|---|---|
| Type 1: | beacon | possum | piercing | bookie |
| Type 2: | became | pursue | percent | bouquet |
| Type 3: | caboose | support | suppose | kebab |
| Type 4: | cabin | soapy | supper | cobble |

| Target pair: b,d | | Non-target pair: k,s | | |
|---|---|---|---|---|
| Type 1: | docile | backer | booking | decent |
| Type 2: | descend | because | becalm | deceit |
| Type 3: | sedate | combine | combust | sedan |
| Type 4: | sided | cabbage | carbon | sadden |

| Target pair: f,p | | Non-target pair: s,l | | |
|---|---|---|---|---|
| Type 1: | fossil | pillow | pilot | facile |
| Type 2: | facade | polite | police | forsake |
| Type 3: | suffuse | lapel | lampoon | suffice |
| Type 4: | siphon | leper | lipid | sofa |

| Target pair: p,b | | Non-target pair: s,d | | |
|---|---|---|---|---|
| Type 1: | passive | bundle | bedding | person |
| Type 2: | percent | bedeck | bedew | pursue |
| Type 3: | surpass | debase | debate | superb |
| Type 4: | super | debit | double | sappy |

| Target pair: l,r | | Non-target pair: k,s | | |
|---|---|---|---|---|
| Type 1: | Lisa | wrinkle | racket | lasso |
| Type 2: | LaSalle | recant | raccoon | Lucille |
| Type 3: | Celeste | career | corrupt | saloon |
| Type 4: | Sally | carrot | courage | salad |

| Target pair: s,f | | Non-target pair: r,b | | |
|---|---|---|---|---|
| Type 1: | sorry | phobic | feeble | serum |
| Type 2: | serene | forbid | forbode | surround |
| Type 3: | recite | befoul | before | receipt |
| Type 4: | racy | barfing | beefy | wrestle |

| Target pair: p,k | | Non-target pair: r,d | | |
|---|---|---|---|---|
| Type 1: | parent | coddle | kidding | parish |
| Type 2: | parole | cadet | condone | Peru |
| Type 3: | repeal | discard | decay | riposte |
| Type 4: | ripen | decade | decoy | romper |

| | | | | |
|---|---|---|---|---|
| Target pair: l,r | | Non-target pair: k,m | | |
| Type 1: | rumor | lucky | liquor | roaming |
| Type 2: | remain | locale | LaCoste | remote |
| Type 3: | maroon | collapse | collide | morass |
| Type 4: | mirror | colon | cooling | marrow |

| | | | | |
|---|---|---|---|---|
| Target pair: m,r | | Non-target pair: s,p | | |
| Type 1: | massive | ripple | rapid | messy |
| Type 2: | myself | repeat | repulse | masseuse |
| Type 3: | cement | parade | peruse | surmount |
| Type 4: | summer | porous | purring | salmon |

# CHAPTER 7
# Word Onsets in OCP-Place Effects

In standard generative phonology, there is no role for the mechanism of lexical access as a part of linguistic competence. Lexical access is assumed to be relevant only to aspects of linguistic performance such as speech production and perception. In fact, much of what is known about lexical access comes from the study of errors in production and perception. Errors are traditionally considered aspects of linguistic performance, as linguistic competence is assumed to be error free. In this chapter, I present data from OCP-Place effects in Arabic which suggests that the mechanism of lexical access does play a role in determining the phonotactic acceptability of a form. In particular, I show that OCP-Place is more strictly enforced at the beginning of the word. In the previous chapter, I demonstrated an analogous pattern in language production, in a speech error elicitation experiment. As shown in chapters 4 and 5, both OCP-Place and phonological speech errors share an underlying dependence on the *similarity* of the consonants involved. I propose that the special status of word initial consonants in OCP-Place follows from evaluating the similarity of consonants on a sequentially encoded representation of the word. In other words, that access of the segmental content of the lexical item takes place in its natural temporal sequence, rather than all-at-once, in *both phonological processing and phonotactic constraint evaluation*.

This result calls into question the traditional competence/performance distinction. Since lexical access is considered an aspect of performance, and access has an effect on the phonotactics, there is a direct influence of performance on competence. This influence can impact both the synchronic grammar, through the productivity of morphemes and the assimilation of borrowed words, and the diachronic grammar, as the cumulative effects of the influence of performance become grammaticized.

Recall from chapter 5 that OCP-Place is enforced to the degree to which homorganic consonants are similar. In this chapter, I demonstrate that the OCP-Place constraint is more sensitive to similarity for word onsets consonants. In Arabic, OCP-Place effects are stronger word initially than later in the word. In modeling the Arabic data, I present a model of the Arabic verbal root lexicon that employs stochastic constraint combination. The Arabic model I present is a significant innovation in modeling using the stochastic constraint, as it is the first model that predicts the occurrence of complete root forms instead of individual consonant pairs. The chapter concludes with a discussion of the implications of sequential encoding of phonological material for linguistic theory.

## *7.1 Word Onsets in Arabic*

The Frisch, Broe, & Pierrehumbert (1995) analysis of OCP-Place in Arabic was presented in chapter 5 for both adjacent ($C_1C_2$ and $C_2C_3$) and non-adjacent ($C_1C_3$) consonant pairs in Arabic $C_1C_2C_3$ roots. In this section, I consider differences between the adjacent $C_1C_2$ and $C_2C_3$ consonant pairs, and the effects that simultaneous violations of OCP-Place by both $C_1C_2$ and $C_2C_3$ have on the acceptability of an entire root.

*7.1.1 Data.*

Tables 7.1 and 7.2 show observed and expected counts of verbal roots of Arabic aggregated by similarity. Table 7.3 gives the relative rate of cooccurrence as measured by O/E. Shading in these tables indicates cells off the main diagonal with sufficiently high expected cooccurrence to be used in the statistical test for asymmetry discussed below. Data in these tables are aggregated separately by the similarity of adjacent consonant pairs. Expected counts are computed in the same way as those in chapter 5, by assuming that consonants should cooccur at random, and are based on the consonant frequencies in first, second, and third position in the root. The observed and expected counts clearly covary, showing that frequency predicts occurrence to a large degree, a point to which I return below.

**Table 7.1**: Actual Arabic verbal roots with consonant pairs aggregated by similarity.

| Similarity | $C_2C_3$ | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| $C_1C_2$ | 0 | 0-0.1 | 0.1-0.2 | 0.2-0.3 | 0.3-0.4 | 0.4-0.5 | 0.5-0.6 | 0.6-0.7 | 1 | Total |
| 0 | 1694 | 146 | 196 | 45 | 9 | 4 | 0 | 0 | 0 | 2094 |
| 0-0.1 | 170 | 19 | 23 | 7 | 0 | 1 | 1 | 0 | 0 | 220 |
| 0.1-0.2 | 177 | 27 | 31 | 12 | 1 | 2 | 1 | 0 | 0 | 251 |
| 0.2-0.3 | 55 | 12 | 12 | 3 | 1 | 0 | 1 | 0 | 0 | 84 |
| 0.3-0.4 | 12 | 4 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 18 |
| 0.4-0.5 | 7 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 7 |
| 0.5-0.6 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0.6-0.7 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |
| Total | 2116 | 208 | 264 | 67 | 11 | 7 | 3 | 0 | 0 | 2676 |

**Table 7.2**: Expected counts of Arabic roots aggregated by similarity.

| Similarity | $C_2C_3$ | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| $C_1C_2$ | 0 | 0-0.1 | 0.1-0.2 | 0.2-0.3 | 0.3-0.4 | 0.4-0.5 | 0.5-0.6 | 0.6-0.7 | 1 | Total |
| 0 | 1232 | 78.0 | 130.2 | 59.7 | 28.6 | 60.2 | 8.3 | 26.0 | 87.2 | 1711 |
| 0-0.1 | 97.5 | 21.1 | 22.0 | 7.6 | 4.9 | 6.2 | 1.7 | 5.3 | 8.1 | 174.4 |
| 0.1-0.2 | 166.2 | 24.7 | 57.9 | 19.8 | 11.5 | 7.0 | 3.5 | 7.3 | 11.6 | 309.3 |
| 0.2-0.3 | 80.6 | 10.2 | 23.7 | 14.1 | 5.6 | 3.1 | 1.5 | 2.0 | 4.2 | 145.0 |
| 0.3-0.4 | 41.3 | 6.9 | 11.4 | 4.5 | 4.0 | 0.9 | 0.7 | 0.8 | 2.0 | 72.4 |
| 0.4-0.5 | 57.0 | 6.0 | 7.6 | 2.5 | 0.8 | 6.4 | 0.2 | 3.2 | 5.7 | 90.3 |
| 0.5-0.6 | 12.4 | 2.3 | 4.4 | 1.7 | 0.7 | 0.2 | 0.8 | 0.2 | 0.7 | 23.4 |
| 0.6-0.7 | 23.7 | 3.6 | 5.3 | 1.5 | 1.0 | 1.4 | 0.1 | 2.3 | 2.2 | 41.0 |
| 1 | 73.4 | 6.9 | 10.0 | 3.9 | 1.9 | 4.7 | 0.7 | 2.3 | 5.7 | 109.4 |
| Total | 1788 | 160.0 | 272.5 | 115.2 | 58.9 | 90.0 | 17.5 | 49.3 | 127.2 | 2676 |

**Table 7.3**: Relative cooccurrence of Arabic roots aggregated by similarity.

| Similarity | $C_2C_3$ | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| $C_1C_2$ | 0 | 0-0.1 | 0.1-0.2 | 0.2-0.3 | 0.3-0.4 | 0.4-0.5 | 0.5-0.6 | 0.6-0.7 | 1 | Total |
| 0 | 1.37 | 1.87 | 1.51 | 0.75 | 0.31 | 0.07 | 0 | 0 | 0 | 1.22 |
| 0-0.1 | 1.73 | 0.90 | 1.04 | 0.92 | 0 | 0.16 | 0.60 | 0 | 0 | 1.26 |
| 0.1-0.2 | 1.06 | 1.09 | 0.54 | 0.61 | 0.09 | 0.29 | 0.29 | 0 | 0 | 0.81 |
| 0.2-0.3 | 0.68 | 1.18 | 0.51 | 0.21 | 0.19 | 0 | 0.66 | 0 | 0 | 0.58 |
| 0.3-0.4 | 0.29 | 0.58 | 0.17 | 0 | 0 | 0 | 0 | 0 | 0 | 0.25 |
| 0.4-0.5 | 0.12 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.08 |
| 0.5-0.6 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0.6-0.7 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 1 | 0.01 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.01 |
| Total | 1.18 | 1.30 | 0.97 | 0.58 | 0.19 | 0.08 | 0.17 | 0 | 0 | 1 |

Table 7.3 reveals two factors which were not apparent in the presentation of the Arabic data in Frisch, Broe, & Pierrehumbert (1995). While there is a clear influence of the similarity of the consonant pair on cooccurrence, the effect seems to be stronger for $C_1C_2$ than $C_2C_3$. The difference becomes apparent when shaded cells across the main diagonal are compared. For example, there are relatively fewer pairs with $C_1C_2$ similarity of 0.1-0.2 and $C_2C_3$ similarity of 0

(O/E = 1.06) than pairs with $C_2C_3$ similarity of 0.1-0.2 and $C_1C_2$ similarity of 0 (O/E = 1.51). This asymmetry is significant, based on a chi-square test.

If the OCP-Place effects were symmetric, the observed number of pairs in cells across the diagonal are expected to be evenly distributed based on the positional frequency of the consonant pairs involved. For example, there are a total of 373 pairs in the two cells with $C_1C_2$ similarity of 0.1-0.2 and $C_2C_3$ similarity of 0 and with $C_2C_3$ similarity of 0.1-0.2 and $C_1C_2$ similarity of 0. There are 296.4 expected pairs of this type, 166.2 with $C_1C_2$ similarity of 0.1-0.2 and $C_2C_3$ similarity of 0 and 130.2 with $C_2C_3$ similarity of 0.1-0.2 and $C_1C_2$ similarity of 0. If the actually occurring pairs were divided between these two cells only by frequency, there should be 209.2 (= 373×166.2/296.4) pairs with $C_1C_2$ similarity of 0.1-0.2 and $C_2C_3$ similarity of 0 and 163.8 (= 373×130.2/296.4) pairs with $C_2C_3$ similarity of 0.1-0.2 and $C_1C_2$ similarity of 0. Shaded cells in the tables indicate where there are a sufficient number of occurrences for the chi-square test to apply. There are a total of 8 such pairs ($\chi^2(7) = 13.4$, $p = 0.06$).

Table 7.3 also reveals a cumulative interaction of the similarity of the first and second consonant with the similarity of the second and third consonant. This is apparently an effect of multiple OCP-Place constraint violations in the Arabic verbal roots. Examining cells on the diagonal we see, for example, that roots where the similarity of $C_1C_2$ and $C_2C_3$ of 0.1-0.2 (O/E = 0.54) are far more underrepresented than roots where the similarity of either $C_1C_2$ or $C_2C_3$ are 0.1-0.2 and the other pair is non-homorganic, in other words the similarity of the other pair is 0 (O/E = 1.06 and O/E = 1.51 respectively).

### *7.1.2 Model of gradient constraint combination.*

Frisch, Broe, and Pierrehumbert (1995) modeled the cooccurrence restrictions of Arabic by fitting a logistic model of the constraint to all of the adjacent consonant pairs aggregated together. In other words, they fit a single stochastic constraint model to the first and second pairs and the second and third pairs together, ignoring the fact that these pairs combine to form a complete triconsonantal root. I present here an improved model of the Arabic data which considers each root in its entirety. In so doing, I implement the model of stochastic constraint combination suggested by Pierrehumbert & Nair (1995) and Frisch, Broe, & Pierrehumbert (1995).

A complete model of the Arabic cooccurrence constraints simultaneously considers the acceptability of a root by the similarity of the first and second consonants and the similarity of the second and third consonants, as well as the similarity of the first and third consonants, to which I return below. Combinations are allowed for each consonant pair as a function of their similarity. A model of combined stochastic constraints for adjacent consonant pairs is given in (82):

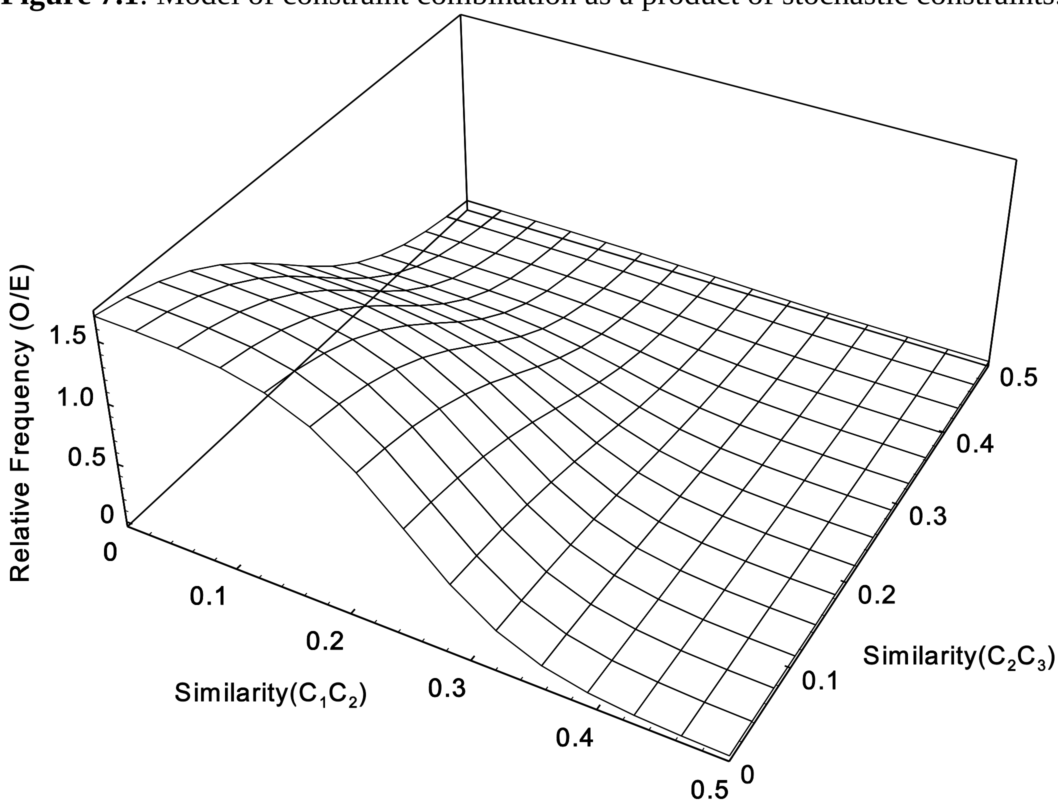
(82)
$$\text{Observed} = \frac{\text{Expected} \cdot \text{A}}{(1 + e^{(\text{K}_1 + \text{S}_1 \cdot \text{similarity}_{12})}) \cdot (1 + e^{(\text{K}_2 + \text{S}_2 \cdot \text{similarity}_{23})})}$$

Where $\text{similarity}_{12}$ is the similarity of $C_1C_2$,
and $\text{similarity}_{23}$ is the similarity of $C_2C_3$.

Each term in the denominator of (82) is a stochastic constraint model; one for each pair of adjacent consonants in the triconsonantal root. Stochastic constraints are combined by multiplication, which corresponds to fuzzy set intersection (Zadeh 1965, Kosko 1990). This model of stochastic constraint combination is a gradient generalization of the categorical model of constraint combination as simultaneous constraint satisfaction.

The model of constraint combination as multiplication of stochastic constraints is shown graphically in Figure 7.1. The *x* and *y* axes are the similarity of $C_1C_2$ and $C_2C_3$ respectively, and the *z* axis is the relative acceptability of the combination. Figure 7.1 shows the effect of cumulative interaction, as acceptability near either axis (where the similarity of one pair is near zero) is much higher than acceptability near the diagonal, where both pairs have non-zero similarity.

**Figure 7.1**: Model of constraint combination as a product of stochastic constraints.




Note that there are roots in Arabic in which both $C_1C_2$ and $C_2C_3$ are homorganic. In a root of this form, $C_1$ and $C_3$ are also, of course, homorganic. In chapter 5, I presented the Frisch, Broe, & Pierrehumbert (1995) model of OCP-Place for non-adjacent consonants. However, Frisch, Broe, & Pierrehumbert (1995) did not consider the effects of simultaneous violations of OCP-Place for $C_1C_2$ and $C_2C_3$. Since they looked only at initial and final consonants, without considering the intervening consonant in the root, it is likely that the $C_1C_3$ pattern they examined is affected by simultaneous $C_1C_2$ and $C_2C_3$ violations. Table 7.4 presents actual and expected $C_1C_3$ pairs, aggregated by similarity, for roots which have non-homorganic $C_1C_2$ and $C_2C_3$ pairs.

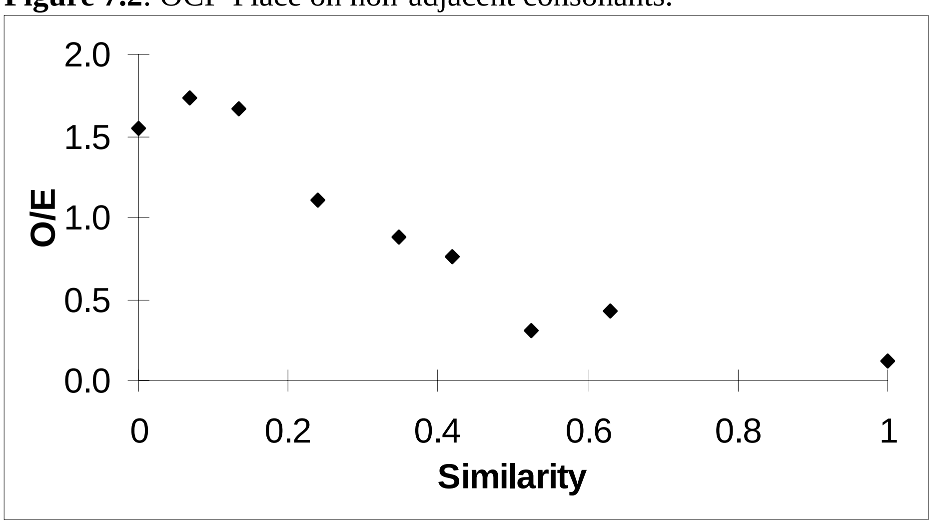
These are the data presented in the upper leftmost cell of Tables 7.1-7.3.

**Table 7.4**: Cooccurrence between homorganic non-adjacent consonants.

| Similarity | Actual | Expected | O/E |
|---|---|---|---|
| 0 | 1108 | 713.6 | 1.55 |
| 0-0.1 | 169 | 97.7 | 1.73 |
| 0.1-0.2 | 244 | 146.8 | 1.66 |
| 0.2-0.3 | 69 | 62.5 | 1.10 |
| 0.3-0.4 | 31 | 35.2 | 0.88 |
| 0.4-0.5 | 50 | 66.1 | 0.76 |
| 0.5-0.6 | 4 | 13.1 | 0.31 |
| 0.6-0.7 | 10 | 23.2 | 0.43 |
| 1 | 9 | 74.6 | 0.12 |
| Total | 1694 | 1232.9 | |

Table 7.4 clearly shows that there is an OCP-Place effect between non-adjacent homorganic consonants independent of the simultaneous occurrence of $C_1C_2$ and $C_2C_3$ violations. The non-adjacent data (with no adjacent pair OCP-Place violations) that are aggregated in Table 7.4 can be fit with a stochastic constraint model. The data are shown in Figure 7.2. The best fit stochastic constraint model has parameters $A$ = 1.58, $K$ = -4.06, and $S$ = 8.89.

**Figure 7.2**: OCP-Place on non-adjacent consonants.




Given that there is an independent, non-adjacent $C_1C_3$ constraint, the complete model of the Arabic data has to combine three stochastic constraints, one for each pair of adjacent consonants, and one for the non-adjacent pair. The revised stochastic constraint model, based on the original analysis of Frisch, Broe, & Pierrehumbert (1995) is shown in (83).

(83)
$$\text{Observed} = \frac{\text{Expected} \cdot \text{A}}{(1 + e^{(K_1 + S_1 \cdot \text{similarity}_{12})}) \cdot (1 + e^{(K_1 + S_1 \cdot \text{similarity}_{23})}) \cdot (1 + e^{(K_2 + S_2 \cdot \text{similarity}_{13})})}$$

Following the analysis of Frisch, Broe, & Pierrehumbert (1995), $K_1$ and $S_1$ are parameters of the stochastic constraint against adjacent consonant pairs, which applies twice: once to $C_1C_2$ and once to $C_2C_3$. $K_2$ and $S_2$ are parameters of the stochastic constraint against non-adjacent pairs, which applies to $C_1C_3$.

In Frisch, Broe, & Pierrehumbert (1995), the stochasic constraint model was fit to individual consonant pairs. This represents a form of aggregation, as consonant pairs occur in different words in different positions. When modeling entire roots, it is not possible to achieve a reasonable model fit to the actual consonant combinations (which is equivalent to modeling the actual lexicon) as this amounts to predicting the pattern of accidental as well as systematic gaps. For example, the non-occurrence of *bleeg,* as opposed to *bleat*, *bleed*, and *bleak* in English is not a principled gap. To avoid this problem, the models in this chapter are fit to roots which are aggregated by similarity. Data were aggregated by the 53 unique similarity values that were computed, so there was less aggregation than shown in the table above. The majority of the aggregation involved the non-homorganic consonant pairs, which all have similarity zero.

The revised stochastic constraint model with gradient constraint combination, which I refer to as the 'stochastic model', provides a better fit of the Arabic data than the categorical models of OCP-Place in Arabic which are not similarity based. Recall that the 'categorical model' is a strict interpretation of OCP-Place where adjacent and non-adjacent homorganic consonants are disallowed. The 'soft model' has a soft constraint against homorganic consonant pairs. The soft model prohibits roots with adjacent identical consonants, and has roots with homorganic consonants (whether adjacent or non-adjacent) underrepresented at a constant rate. In implementing these two models in this chapter, I determined their model fit to the same similarity aggregated data as the stochastic model. Consonants with similarity greater than zero were considered homorganic and consonants with similarity one are identical. Note that these models do not make special consideration for the obstruent/sonorant split in the coronals. In addition, they treat the emphatics {T, D, S, Z} as homorganic to the coronals and dorso-gutterals. Thus, these models do not employ the ad hoc stipulations added to the OCP-Place constraint in McCarthy (1994).

These models are compared on the basis of the $R^2$ as well as the relative $R^2$, as in chapter 5. The relative $R^2$ is based on the amount of reduction in the residual sum of squares of each model over the 'frequency model'.

(84)
$$\text{Relative } R^2 = 1 - \frac{\text{Residual SS}}{\text{Frequency Model Residual SS}}$$

To highlight differences in the models, relative $R^2$ values are computed twice, once for all root forms, and once for just the roots which contain at least one homorganic consonant pair, in other words roots which are expected to have some OCP-Place effect. Best fit model parameters and relative $R^2$ values are presented in Table 7.5.

**Table 7.5**: Model parameters and fits of the Arabic triconsonantal roots.

| Model | $R^2$ | Relative $R^2$ | Homorganic Relative $R^2$ | Model Parameters |
|---|---|---|---|---|
| Frequency model | 0.85 | - | - | O = E |
| Categorical model | 0.97 | 0.79 | - | O/E = 0 for homorganic<br>O/E = 1.55 otherwise |
| Soft model | 0.99 | 0.91 | 0.45 | O/E = 0 for identical<br>O/E = 0.89 for homorganic<br>O/E = 1.58 otherwise |
| Stochastic model | 0.995 | 0.97 | 0.81 | A = 1.63<br>Adjacent: $S_1 = 19.0$, $K_1 = -4.6$<br>Non-adjacent: $S_2 = 7.6$, $K_2 = -3.6$ |

The categorical model does capture a great deal of variation in the data, when all roots are considered, primarily by allowing overrepresentation of roots with non-homorganic consonants. This model fairs much worse when just the homorganic roots are examined. In fact, over this subset of the data, the categorical model fits worse than the frequency model, so no relative $R^2$ is given. The categorical model is significantly improved upon by both the soft model and the stochastic model. Both of these models predict gradient acceptability of forms among homorganic consonant pairs. The stochastic model is far superior to the soft model, especially when variation only within roots that contain homorganic consonants is considered. The soft model predicts that there is no variation within this class, apart from expected frequency. The stochastic model predicts a great deal of variation, and that variation is cumulative across all three consonant pairs. Notice that the parameters for the stochastic model reveal the effects of distance on the dissimilarity constraint. The constraint for non-adjacent consonants ($S_2 = 7.6$) is much weaker than the constraint for adjacent consonants ($S_1 = 19.0$), a fact not captured by the categorical model or the soft model. The stochastic constraint for non-adjacent consonants in the stochastic model, which uses constraint combination, is quite close in form to the one found above when only non-adjacent violations were examined independently (K = -4.06, S = 8.89).

### *7.1.3 Modeling asymmetries in word position.*

The stochastic model which I just presented is based on the original Frisch, Broe, & Pierrehumbert (1995) model of Arabic cooccurrence. This model does not capture the significant asymmetry in the OCP-Place effects found above. A minor modification to the symmetric stochastic model is needed to model this asymmetry. The ‘asymmetric stochastic model’ allows the two adjacent consonant constraints (for the first and second consonant pair and for the second and third consonant pair) to be parameterized separately. With independent parameterizations for the word initial adjacent pair and the non-initial adjacent pair, the asymmetric stochastic model reflects the relative strength of OCP-Place in each position.

(85)
$$\text{Observed} = \frac{\text{Expected} \cdot A}{(1 + e^{K_1 + S_1 \cdot \text{similarity}_{12}}) \cdot (1 + e^{K_2 + S_2 \cdot \text{similarity}_{23}}) \cdot (1 + e^{K_3 + S_3 \cdot \text{similarity}_{13}})}$$

In the asymmetric model, $K_1$ and $S_1$ are stochastic constraint parameters for $C_1C_2$. $K_2$ and $S_2$ are stochastic constraint parameters for $C_2C_3$. $K_3$ and $S_3$ are stochastic constraint parameters for $C_1C_3$. Table 7.6 compares the two stochastic constraint models of Arabic. The asymmetric stochastic model has a relative $R^2 = 0.87$ over roots with homorganic consonant pairs. This is a 10% improvement in residual sum of squares over the symmetric stochastic model over the data set for which they make different predictions. The improvement in model fit is expected, since the asymmetric model has two extra degrees of freedom due to the two extra stochastic constraint parameters, but I feel it is a significant improvement that warrants the additional model complexity.

**Table 7.6**: Symmetric and asymmetric stochastic models of the Arabic roots.

| Model | $R^2$ | Relative $R^2$ | Homorganic Relative $R^2$ | Model Parameters |
|---|---|---|---|---|
| Stochastic model | 0.995 | 0.97 | 0.80 | $A = 1.63$<br>Adjacent: $S_1 = 19.0$, $K_1 = -4.6$<br>Non-adjacent: $S_2 = 7.6$, $K_2 = -3.6$ |
| Asymmetric stochastic model | 0.997 | 0.98 | 0.87 | $A = 1.80$<br>$C_1C_2$: $S_1 = 21.9$, $K_1 = -5.5$<br>$C_2C_3$: $S_2 = 13.0$, $K_2 = -2.9$<br>$C_1C_3$: $S_3 = 5.2$, $K_3 = -2.5$ |

In the asymmetric stochastic model, the stochastic constraint for the first and second consonants is sharper ($S_1 = 21.9$, 95% confidence interval for $S_1$ is $18.3 \leq S_1 \leq 25.4$) than the stochastic constraint for the second and third consonants ($S_2 = 13.0$, 95% confidence interval for $S_2$ is $11.3 \leq S_2 \leq 14.6$). The difference in constraint parameters captures the asymmetry between the initial consonant pair and the final consonant pair. The sharpness of the stochastic constraint parameters shows that the dissimilarity constraint is stronger word initially than later in the word. This is the same effect that was found in the speech error data in Experiment 1.

### *7.1.4 Summary.*

In this section, I have shown that OCP-Place effects are stronger word initially than elsewhere in the word. The results presented above were effectively modeled using a stochastic constraint, which models gradient data using two parameters that determine the overall shape of the constraint. The variation in constraint sharpness by word position was modeled with different parameterizations for different word positions. The stronger word initial constraint had greater constraint sharpness than the other constraints. These results presents a challenge to traditional models of grammar which employ categorical constraints. Categorical constraints cannot model the gradient nature of the data, nor the variation in gradient effects within and between languages demonstrated in this section.

In addition, I presented a model of gradient constraint combination which models cooccurrence in the Arabic verbal root lexicon. This model considered the triconsonantal root as a whole and thus modeled patterns in the entire lexicon rather than particular patterns of individual consonant pairs.

### *7.2 Accounting for Word Onset OCP-Place Effects Using Sequential Encoding*

In chapter 6, I presented the model of phonological encoding in Sevald & Dell (1994) which proposes sequential activation of segments in phonological encoding. Above, I presented evidence that a phonotactic constraint involving similarity has stronger effects when word onsets are involved than otherwise. The results for OCP-Place parallel the results for speech errors. The account of word onset effects for speech errors can be extended to the effects of word position on OCP-Place.

Recall that word onset consonants are accessed with a minimum of competing activation, while later consonants compete with consonants from other possible words involving the initial segmental context. The competition which is created by the access the initial segments of a word serves as interference in the comparison of segments for purposes of determining similarity. Thus, in the case of the first consonants in the word, there is very little interference and the similarity comparison can sharply differentiate consonants. As more consonants are accessed, the access of those consonants as well as the activation of consonants in similar words creates proactive interference for the judgement of similarity. Proactive interference has been shown to degrade performance in memory tasks (e.g. Watkins 1975). In the noisy context, extreme similarity and extreme dissimilarity are more difficult to detect, as there is a background of essentially random activation.

A similar account is given by Pierrehumbert (1993) and Berkley (1994a) for distance effects on the OCP-Place constraint. The weakening of OCP-Place over distance was seen above in the stochastic constraint model: the stochastic constraint model had a weaker constraint for non-adjacent consonants (sharpness S = 7.6) than for adjacent consonants (sharpness S = 19.0). Pierrehumbert (1993) analyzed the weakening of the effects of OCP-Place over distance (between $C_1C_3$) as an effect of interference of the intervening $C_2$. This interference is predicted by the sequential activation model, but not by a model of lexical access where all segments are accessed simultaneously. Thus, sequential access of segmental material accounts for two different weakening effects on OCP-Place. In fact, it is difficult to imagine any account for the effects of distance if a completely atemporal representation of the word is assumed. If the entire representation of the word is available for access all at once, there is no reason why intervening consonants should have an effect on the comparison of distant ones, since the two distant consonants could presumably be accessed just like two adjacent consonants.

The common finding of a word onset effect between speech production and phonotactics suggests that the same mechanisms are used for phonological encoding both in evaluating phonotactic constraints and in production. Thus, I claim that linguistic constraints like OCP-Place operate on an inherently temporal representation of phonological material. This is a radical departure from the traditional conception, where an abstract representation of the word's segmental context is assumed to be accessible in an all-at-once fashion.

The constraints imposed by sequential encoding on the OCP are reminiscent of the running window model of phonetic implementation of phrasal phonology (Pierrehumbert & Beckman 1988, Pierrehumbert 1995). In this model, segments and tones are implemented in temporal order, mimicking the production of speech in time. The temporal model provides a more natural framework than an atemporal hierarchical model (e.g. Coleman 1992) for modeling allophonic glottalization in English (Pierrehumbert & Frisch 1996). Thus, the architecture for implementing word level and phrase level phonology both appear to be sensitive to real time processing.

There is evidence for sequential access of segmental material and competition between outcomes in speech perception as well. The cohort model of lexical access for recognition (Marslen-Wilson & Welsch 1978; Marslen-Wilson 1984, 1987) involves a narrowing cohort of possible words, based on the segmental information which has thus far been perceived. The cohort is brought to a unique choice when enough of the segments have been identified. Other models of perception share the assumption of competition between potential percepts (e.g. logogen theory (Morton 1982), the neighborhood activation model (Luce 1986), TRACE (McClelland & Elman 1986), see Goldinger, Pisoni, & Luce 1996 for a recent review). There is thus evidence for sequential access of segmental information and competition in production, perception, and more abstract phonological processes like linguistic constraint application. This uniformity suggests a single mechanism for the storage and retrieval of segmental information. The Sevald & Dell (1994) model of phonological encoding presented above is a valid model of this mechanism for all three purposes.

# CHAPTER 8
# Underspecification in Speech Errors I: Similarity[17]

In this chapter, I demonstrate that the use of UNDERSPECIFICATION (Kiparsky 1982; Archangeli 1984, 1988) in phonological feature matrices is not supported by the speech error data examined in this thesis. Stemberger (1991b) proposed that underspecified features have less influence on the similarity of consonants, and hence speech error rates, since underspecified features are blanks during early portions of the derivation. However, he did not consider the effect that underspecification would have on similarity outside of the small number of minimal contrasts which he examined. I demonstrate that computing similarity over lattices provides a more accurate prediction of error rate than measures of similarity based on underspecified feature matrices, when the entire corpus of phonological segment errors are examined.

## *8.1 Introduction to Underspecification*

Underspecification refers to the practice of leaving blanks in feature matrices which are filled in during the course of the phonological derivation. There are two types of underspecification which are generally practiced: CONTRASTIVE UNDERSPECIFICATION and RADICAL UNDERSPECIFICATION. In contrastive underspecification (Steriade 1987), non-contrastive features are left out of the feature matrix. As mentioned in chapter 2, the [+voice] feature of sonorant consonants in English is predictable, since all sonorants are voiced. Thus, in contrastive underspecification, stops and fricatives are marked as [±voice] but sonorants would be underspecified. A redundancy rule [+son] → [+voice] applies during the phonological derivation. Some phonological processes, such as devoicing coda consonants in German, are proposed to apply before the redundant voicing feature is filled in, with the result that coda sonorants do not devoice in German.

Radical underspecification (Archangeli 1988) proposes that one value for each feature is considered the default and is always left blank in underlying representation. For example, [-voice] is considered to be the unmarked value of voicing in obstruents, so voiceless obstruents are underspecified for voicing. In radical underspecification, the [-voice] specification is inserted by a default rule which marks any consonant without a voicing feature as [-voice]. As with contrastive underspecification, predictable features are also left blank, so [+voice] in sonorants is also underspecified and filled in by a redundancy rule.

Underspecification is motivated by phonological processes which treat some segments as though they were transparent to the process and as a means of achieving minimal underlying representations (see Steriade 1995 for a review). Underspecification is used in autosegmental

[17] The preparation of this chapter benefitted from reading a review of Stemberger (1991b) written by Will Thompson for the introductory graduate phonology course at Northwestern in the fall quarter of 1995-6. I refer to this uncirculated manuscript in the text and references as Thompson (1995).

phonology as a means of allowing feature spreading to apply across intervening segments. I do not discuss the use of underspecification in phonological theory to account for assimilation or cross-linguistic tendencies in epenthetic segments (see e.g. Archangeli 1988, Mohanan 1991, and Paradis & Prunet 1991), but instead refer the interested reader to Broe (1993) and Steriade (1995) for detailed discussions of underspecification and some alternative analyses. In this thesis, I examine evidence against underspecification in the similarity based speech errors and OCP-Place effects introduced in chapters 4 and 5.

Broe (1993) proposes that the transparency of phonological rules attributed to underspecification can instead be captured by sensitizing rules or constraints to the redundant or marked status of features, as encoded by the redundancy hierarchy and the markedness hierarchy. Steriade (1995) concludes that the evidence for the eventual specification of underspecified features is not compelling, and instead proposes that all cases of underspecification are INHERENT UNDERSPECIFICATION or TRIVIAL UNDERSPECIFICATION in which the underspecified features are *never* filled in. I showed in chapter 2, following Broe (1995), that inherent underspecification leads to a loss of distinctness of underlying representations. For example, Lombardi (1991) proposes that voicing is a privative feature and thus that [-voice] is never specified. Without [-voice], though, /t/ and /d/, /s/ and /z/, and so on, cannot be individuated. Thus, I adopt structured specification and reject all forms of underspecification.

Underspecification contrasts with full specification, shown in the feature matrix in (86). Full specification of features underlies the redundancy hierarchy. In the fully specified feature matrices in this chapter, I use the mixture of bivalent and monovalent features which is traditionally used (e.g. Kenstowicz 1994). In this matrix, feature values are only left blank if they are irrelevant or do not apply to a particular phoneme. For example, the feature [anterior] only applies to coronals, and so is left blank for labials and velars (Sagey 1986, Yip 1989). Place features, spelled with capital letters in this matrix, are monovalent, and are either present (%) or absent. I have also used a privative feature [Affricate] for the affricates /tʃ/ and /dʒ/, since the particular featural representation of these complex segments does not bear on the analysis of underspecification.

| (86) | p | b | f | v | m | t | d | θ | ð | s | z | ʃ | ʒ | tʃ | dʒ | k | g | ŋ | l | r | n | w | y | h |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| sonorant | - | - | - | - | + | - | - | - | - | - | - | - | - | - | - | - | - | + | + | + | + | + | + | - |
| continuant | - | - | + | + | - | - | - | + | + | + | + | + | + | - | - | - | - | - | - | + | - | + | + | + |
| voice | - | + | - | + | + | - | + | - | + | - | + | - | + | - | + | - | + | + | + | + | + | + | + | - |
| Labial | % | % | % | % | % | | | | | | | | | | | | | | | % | | % | | |
| Velar | | | | | | | | | | | | | | | | % | % | % | | | | | | |
| Coronal | | | | | | % | % | % | % | % | % | % | % | % | % | | | | % | % | % | | % | |
| anterior | | | | | | + | + | - | - | + | + | - | - | - | - | | | | + | - | + | | - | |
| dental | | | | | | - | - | + | + | - | - | - | - | - | - | | | | - | - | - | | - | |
| Affricate | | | | | | | | | | | | | | % | % | | | | | | | | | |
| glide | | | | | - | | | | | | | | | | | | | - | - | - | - | + | + | |
| nasal | | | | | + | | | | | | | | | | | | | + | - | - | + | - | - | |
| lateral | | | | | - | | | | | | | | | | | | | - | + | - | - | - | - | |

In (87), I show a feature matrix for the English consonants over the same features assuming contrastive underspecification, as practiced for example in Pierrehumbert (1993).

Major class features (cf. the primary features of Stevens & Keyser (1989)) are assigned first, and features for smaller classes are included only if they are relevant and contrastive within the major classes. For example, voicing is not contrastive for sonorants, and is not included. The feature [anterior] is inapplicable for labials and so is left blank. Note that approaches using underspecification confuse non-contrastive feature specifications with inapplicable ones (Broe 1993).

| (87) | p | b | f | v | m | t | d | θ | ð | s | z | ʃ | ʒ | tʃ | dʒ | k | g | ŋ | l | r | n | w | y | h |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| sonorant | - | - | - | - | + | - | - | - | - | - | - | - | - | - | - | - | - | + | + | + | + | + | + | - |
| continuant | - | - | + | + | | - | - | + | + | + | + | + | + | | | - | - | | | | | | | + |
| voice | - | + | - | + | | - | + | - | + | - | + | - | + | - | + | - | + | | | | | | | - |
| Labial | % | % | % | % | % | | | | | | | | | | | | | | | % | | % | | |
| Velar | | | | | | | | | | | | | | | | % | % | % | | | | | | |
| Coronal | | | | | | % | % | % | % | % | % | % | % | % | % | | | | % | % | % | | % | |
| anterior | | | | | | | | | | + | + | - | - | - | - | | | | + | - | + | | - | |
| dental | | | | | | | | + | + | - | - | | | | | | | | | | | | | |
| Affricate | | | | | | | | | | | | | | % | % | | | | | | | | | |
| glide | | | | | | | | | | | | | | | | | | | | | | + | + | |
| nasal | | | | | + | | | | | | | | | | | | | + | | | + | | | |
| lateral | | | | | | | | | | | | | | | | | | | + | - | | | | |

In (88), I show a feature matrix for the same consonants assuming radical under-specification, adapted from Stemberger (1991b). For all bivalent features except [anterior], the minus value is underspecified. Place features are privative, and in radical underspecification, coronal place is underspecified. Other defaults include [-continuant], [-sonorant], [-voice], and [+anterior].

| (88) | p | b | f | v | m | t | d | θ | ð | s | z | ʃ | ʒ | tʃ | dʒ | k | g | ŋ | l | r | n | w | y | h |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| sonorant | | | | | | | | | | | | | | | | | | | + | + | | + | + | |
| continuant | | | + | + | | | | + | + | + | + | + | + | | | | | | | | | | | + |
| voice | | + | | + | | | + | | + | | + | | + | | + | | + | | | | | | | |
| Labial | % | % | % | % | % | | | | | | | | | | | | | | | % | | % | | |
| Velar | | | | | | | | | | | | | | | | % | % | % | | | | | | |
| Coronal | | | | | | | | % | % | | | % | % | % | % | | | | | % | % | | % | |
| anterior | | | | | | | | | | | | - | - | - | - | | | | | - | + | | - | |
| dental | | | | | | | | + | + | | | | | | | | | | | | | | | |
| Affricate | | | | | | | | | | | | | | % | % | | | | | | | | | |
| glide | | | | | | | | | | | | | | | | | | | | | | | | |
| nasal | | | | | + | | | | | | | | | | | | | + | | | + | | | |
| lateral | | | | | | | | | | | | | | | | | | | | | | | | |

Stemberger (1991b) assumes, as is often practiced (see Paradis & Prunet 1991), that [Coronal] cannot be underspecified when [anterior] is specified. This dependency is based on the theory of feature geometry, discussed in chapter 2. Recall that Sagey (1986) proposed that [anterior] is a hierarchical dependent of [Coronal] in the feature geometry. Proponents of underspecification, in an attempt to incorporate Sagey's proposal, have assumed that the requirements of underspecification can be overridden in this case, and that [Coronal] is specified for the [-anterior] segments due to their hierarchical relationship.

As mentioned in chapter 2, structured specification solves a number of formal problems with underspecification. Underspecification is incompatible with structured specification (Broe 1993), as underspecification leads to a loss of individuation. I adopt structured specification in this thesis, and use it the basis of a similarity metric. Thus, Stemberger's evidence for underspecification in similarity phenomena must be reanalyzed. Stemberger's analysis is discussed in detail in section 8.2. In section 8.3, I compare similarity computed using natural classes with similarity computed over the three feature matrices presented above, and show that the natural classes model gives a superior model of speech error rates in Stemberger's (1991a) corpus, originally presented in chapter 4.

## *8.2 Review of Stemberger (1991b)*

Stemberger (1991b) considers the effect which the assumption of radical underspecification might have on similarity and hence on speech error rates. He hypothesizes that, if speech errors can occur at a point in the phonological derivation where segments are still underspecified for some features, then underspecified features should have no effect on the similarity of consonants at that time. Thus, underspecified features should play no role in determining error rates early in the derivation. If speech errors can occur at more than one point in the derivation, underspecified features may only be relevant for some error opportunities and thus have a reduced effect on error rate.

Consider, for example, the difference between the following fully specified and radically underspecified representations of {p, f, t, s}.

(89) Fully specified feature matrix

| | p | f | t | s |
|---|---|---|---|---|
| sonorant | - | - | - | - |
| continuant | - | + | - | + |
| voice | - | - | - | - |
| Labial | % | % | | |
| Coronal | | | % | % |
| anterior | | | + | + |

(90) Radically underspecified feature matrix

| | p | f | t | s |
|---|---|---|---|---|
| sonorant | | | | |
| continuant | | + | | + |
| voice | | | | |
| Labial | % | % | | |
| Coronal | | | | |
| anterior | | | | |

In the radically underspecified feature matrix, [Labial] is specified, but [Coronal] is left blank, and filled in by a default rule in the course of the phonological derivation. Thus, Stemberger observes, in the radically underspecified representation, /p/ and /f/ share a feature for which the corresponding pair /t/ and /s/ are underspecified. Since shared features increase similarity (Tversky 1977), the radically underspecified feature matrix predicts that /p/ and /f/ should be more similar to one another than /t/ and /s/. Thus, /p/ and /f/ should have higher error rates with one another than /t/ and /s/.

In support of his claims, Stemberger (1991b) considers quantitative evidence from his corpus of naturally occurring speech errors, presented in the confusion matrix in chapter 4. To quantify error rate, Stemberger presents three different methods for estimating random chance and settles on a measure similar to O/E, used in chapter 4, though he does not find large differences between the different measures of chance.

Stemberger's estimate of chance is based on the total number of errors for a particular contrast being examined. The chance error rate for each contrast is the proportion of the total number of errors involving each group. Thus, his estimate of expected error rate is slightly different from the estimate used in chapter 4 as he adjusts the estimates of chance based on the total number of errors within the set of consonants being compared, rather than over the entire inventory of consonants. As a result, the expected numbers of errors for any particular class varies depending on what other class it is being compared with. For example, there are 158 errors involving /p/ and 240 errors involving /t/ in the error corpus. There are a total of 30 errors between /s/ and either /p/ or /t/. Stemberger's estimate of chance would predict 12 errors ($30 \times 158 \div 398$) between /s/ and /p/ and 18 ($30 \times 240 \div 398$) errors between /s/ and /t/. Consider a second example. There are 34 errors in the corpus involving /z/, and a total of 25 errors between /s/ and either /z/ or /t/. In this case, Stemberger's estimate of chance would predict 3 errors ($25 \times 34 \div 274$) between /s/ and /z/ and 22 ($25 \times 240 \div 274$) errors between /s/ and /t/. The number of errors predicted between /s/ and /t/ varies depending on what other pair it is being compared to.

In my method of estimating chance (adopted from Pierrehumbert 1993) the expected number of errors does not vary for any particular consonant pair. I believe my estimate of chance better reflects the effect of similarity on error rate, as it reveals the overall error rate rather than a relative comparison of error rate within a particular contrast. Stemberger's method of predicting chance has the advantage of allowing significance to be computed for each contrast with a chi-square test to find differences in error rate. This was Stemberger's motivation for using this variable measure of chance.

Stemberger (1991b) first examines coronal place of articulation. Within the coronals he only considers alveolars /t d s z n/, since they are the only coronals which are underspecified for place. His results are presented in Table 8.1. Each row of Table 8.1 is one comparison, with predicted errors computed as outlined above. Significance values are also presented for a chi-square test with one degree of freedom. For example, the first row contrasts errors between labials and labials with errors between labials and alveolars. Within this contrast, labial-labial errors are more frequent than labial-alveolar errors, as expected on any similarity based account of speech errors. Note though that comparing the actual error counts to the predicted counts (using the O/E measure as in chapter 4) is not valid given Stemberger's method of predicting chance. The predicted values are relativized to the total number of errors. It is possible that both groups in a particular contrast have O/E greater than or less than one.

**Table 8.1**: Interaction errors across place of articulation from Stemberger (1991b).

| contrast | | actual | | predicted | | $\chi^2(1)$ | p |
|---|---|---|---|---|---|---|---|
| lab-lab | lab-alv | 96 | 175 | 59.22 | 211.78 | 28.45 | 0.001 |
| alv-alv | lab-alv | 88 | 175 | 89.54 | 173.46 | 0.02 | ns |
| alv-alv | vel-alv | 88 | 118 | 105.45 | 100.52 | 5.60 | 0.02 |
| vel-vel | vel-alv | 15 | 118 | 8.97 | 124.21 | 3.97 | 0.05 |

Notice also that the ‘lab-alv’, ‘alv-alv’, and the ‘vel-alv’ groups appear more than once in the table. Each time they appear, the predicted number of errors is different, because they are being contrasted with a different group, resulting in a different total number of errors in the comparison, as explained above.

Examining the first and fourth rows of the table, Stemberger (1991b) concludes that errors within the labial and velar classes occur more frequently when place of articulation is shared than when it is not. By contrast, the second and third rows do not show that sharing alveolar place of articulation increases error rate relative to when place of articulation is not shared.

Stemberger notes, though, that not all contrasts between labials and alveolars are equivalent to contrasts within the alveolars. For example, /p/-/t/ differ only by place of articulation, and have a very high error rate. A more reasonable comparison would contrast /p/-/s/ and /t/-/f/ errors with /t/-/s/ errors, since /p/-/s/ and /t/-/f/ have the exact same contrast as /t/-/s/ with the addition of a difference in place of articulation. If all errors for consonants which differ only on place are eliminated, an effect of coronal is found. This effect is shown in Table 8.2.

**Table 8.2**: Equivalent interaction errors from Stemberger (1991b).

| contrast | | actual | | predicted | | $\chi^2(1)$ | p |
|---|---|---|---|---|---|---|---|
| alv-alv | lab-alv | 88 | 27 | 46.07 | 68.93 | 62.16 | 0.001 |
| alv-alv | vel-alv | 88 | 22 | 64.84 | 45.16 | 19.29 | 0.001 |

Stemberger concludes that shared alveolar does have an effect, though the results of Table 8.2 are equally compatible with the conclusion that it is the difference between the specified labials and velars and the underspecified alveolars which reduces the error rate between alveolars and labials or velars, rather than the shared (underspecified) alveolar which increases the error rate between alveolars. In general, Stemberger’s data do not reveal the source of the deviation from the predicted error rate, since his adjusted estimate of chance does not reveal the true error rate of the consonants involved.

Stemberger next shows that the effect of underspecified coronal place on error rate is weaker than shared labial or shared velar, based on direct comparisons between the error rates within the different place classes, presented in Table 8.3.

**Table 8.3**: Interaction errors within each place of articulation from Stemberger (1991b)

| contrast | | actual | | predicted | | $\chi^2(1)$ | p |
|---|---|---|---|---|---|---|---|
| lab-lab | alv-alv | 96 | 88 | 64.65 | 119.35 | 22.70 | 0.001 |
| lab-lab | vel-vel | 96 | 15 | 98.71 | 12.29 | 0.45 | ns |
| alv-alv | vel-vel | 88 | 15 | 96.49 | 6.51 | 10.47 | 0.005 |

The first and third rows of table 8.3 show that the relative error rate within the labials and velars is higher than the error rate within the alveolars. The second row shows that the relative error rates within the labials and velars are comparable. Thus, while shared alveolar place does presumably have an effect, the effect is weaker than shared labial or velar. Stemberger (1991b) attributes this weaker effect to underspecification of [Coronal] for the alveolar consonants.

Note that the difference in error rates cannot be explained by frequency. All other things being equal, we might expect that since alveolars are high frequency consonants they might be processed more accurately and thus be involved less frequently in errors (Stemberger & MacWhinney 1986, Dell 1990). However, the estimate of chance is based on the frequency of consonants in the error corpus. Any advantage for the high frequency alveolars is already factored out by this intrinsic measure of chance (see Stemberger 1991b).

Stemberger (1991b) demonstrates analogous differences in error rate based on manner of articulation contrasts and voicing contrasts. He shows that, while shared [-continuant] and [-nasal] for stops does have some effect on error rate, the effect is weaker than shared [+continuant] for fricatives or shared [+nasal] for nasal stops. Similarly, he shows shared [-voice] does affect error rates among obstruents, but that shared [+voice] affects error rates more.

Stemberger also demonstrates that the presence of shared redundant features increases error rate. These are features which are left blank in underlying form in both radical underspecification and contrastive underspecification. For example, he shows that the feature [+voice] has an effect in the interaction of sonorants with obstruents, even though [+voice] is predictable for sonorants. Stemberger takes the effects of predictable features on error rate as additional evidence that at least some of the errors must occur late in the derivation, after underspecified features are filled in.

Stemberger takes these differences in error rate as support for underspecified underlying representations. He proposes that errors occur at more than one place in the phonological derivation, and thus that specified features have a greater effect on error rate than underspecified features, as they are present throughout the derivation. Specifically, Stemberger (1991b) proposes a two stage model of speech errors in which errors can occur either before or after underspecified features are filled in. In this model, the similarity of consonants at each stage is different because at the early stage consonants are underspecified and at the later stage they are fully specified. Error rates in the early stage are only influenced by specified features, while errors rates at the second stage are influenced by all features. However, Stemberger (1991b) does not provide an explicit similarity metric (Thompson 1995) or test his assumptions about underspecification and similarity over any other sets of consonants.

### *8.3 Underspecification and Similarity*

Stemberger (1991b) bases his results on the claim that underspecified features have no effect on similarity (and therefore speech error rates) early in the derivation. Recall that underspecification raises a number of formal problems due to the use of feature blanks for undefined features, contextually determined features, as well as default features. In addition, underspecification is dependent upon derivational phonological theory and is thus incompatible with current constraint based approaches. In this section, I show that Stemberger's conclusions cannot be supported when the speech error data are considered in greater detail.

First, recall that the natural classes similarity metric, introduced in chapter 3, differentiates redundant features from contrastive ones for purposes of similarity. Thus, part of Stemberger's result can be captured using similarity computed over lattices, without the formally problematic use of underspecification. Second, I demonstrate below that the degree to which 'underspecification' reduces the error rate between segments varies, and it is not always the case that shared specified features affect similarity more than shared underspecified features. Third, I show that the similarity of English consonants computed using the natural classes model provides a far better fit to the speech error data in Stemberger's consonant confusion matrix than Stemberger's two stage model where errors can apply either before features are specified or after. This suggests that the method of feature assignment in structured specification introduced in chapter 2, using monovalent features, specifying redundant features, and assuring individuation of phonemes is a better representation for determining similarity than feature representations based on underspecification.

#### *8.3.1 Redundancy and underspecification.*

Recall from chapter 4 that the natural classes model of similarity predicts that partially and totally redundant features influence similarity less than fully contrastive features. In particular, I showed that speech error rates among fricatives were sensitive to the reciprocal dependency between [voiceless] and [obstruent]. Since [+voiceless] ⇒ [+obstruent], the similarity between voiceless obstruents is less than the similarity between voiced obstruents. As predicted, the error rate between voiceless fricatives is less than the error rate between voiced fricatives. While radical underspecification and contrastive underspecifiation assume that redundant features are left blank in the representation, and could thus potentially capture the same result, there is no algorithm for reducing the representation to a minimal set of specifications which produces a unique non-arbitrary result (Halle 1959, Stanley 1967, Broe 1993). As demonstrated in chapter 2, structured specification provides a unique redundancy hierarchy for any feature matrix, and is therefore preferable.

#### *8.3.2 The influence of specified and underspecified features.*

Stemberger (1991b) presents his results from the perspective of the effect that underspecification has on the similarity of the consonants involved. He claims that the effects of specified features are stronger than the effects of underspecified features. Stemberger examined the major featural contrasts which are assumed in radical underspecification as well as a case of redundant features. I have already shown how the natural classes similarity model accounts for differential behavior in the case of redundant features. In the case of differences between radically underspecified features, left blank as defaults, and specified features, the natural classes similarity model as given makes no predictions as defaults are not encoded in the redundancy hierarchy.

While Stemberger (1991b) examined minimal contrasts involving underspecification within the major classes and found evidence that underspecified features affected similarity less than specified ones, he did not examine any cases of [coronal] place of articulation which was *not* underspecified. The interdentals, {θ, ð}, and the palatals {ʃ, ʒ, tʃ, dʒ} are assumed to be specified for [Coronal] given the structural requirements of the feature geometry tree.

I examined the effect of specified [Coronal] on speech error rates, a case which Stemberger (1991b) did not discuss. When contrasts involving specified [Coronal] are considered, the results are difficult to interpret based on Stemberger's claim that underspecified features affect similarity less than specified features. I examined errors between all English fricatives, as the [-anterior] coronals are all either fricatives or affricates, and affricates are not found at any other place of articulation. The groups I examined are given in (91).

(91) lab-alv: errors between {f, v} and {s, z}
lab-cor: errors between {f, v} and {θ, ð, ʃ, ʒ}
alv-cor: errors between {s, z} and {θ, ð, ʃ, ʒ}

Table 8.4 compares each group using the method for the relative estimate of chance as in Stemberger (1991b). In all cases, there is a significant difference in the error rate between the contrasted groups.

**Table 8.4**: Error rates for labial, underspecified coronal, and specified coronal fricatives.

| contrast | | actual | | predicted | | $\chi^2(1)$ | p |
|---|---|---|---|---|---|---|---|
| lab-alv | lab-cor | 46 | 12 | 34.34 | 23.66 | 9.71 | 0.002 |
| lab-alv | alv-cor | 46 | 139 | 83.18 | 101.82 | 30.19 | 0.001 |
| lab-cor | alv-cor | 12 | 139 | 54.38 | 96.62 | 51.62 | 0.001 |

The first row shows that there is a significantly stronger interaction between labials and alveolars than between labials and (specified) coronals. Both have an equal degree of mismatch in terms of the place of articulation features: labial and underspecified or specified coronal mismatch equally. Stemberger does not discuss cases of this kind. In addition, he presents no explicit model of similarity (Thompson 1995, see below), so there is no way to determine if these data support his hypothesis or not.

The second row of Table 8.4 shows that there is a relatively greater error rate between alveolars and coronals than between labials and coronals. This contrast is expected in Stemberger's two stage model, as after the alveolars are specified for [coronal] they interact with the underlyingly specified coronals more than the labials. Stemberger predicts that the effect should be a weak one, since it involves errors late in the derivation only. Finally, the third row shows that there is a relatively greater error rate between alveolars and coronals than between labials and coronals. This asymmetry might be predicted in two different ways by the two stage model. First, labials and coronals mismatch underlyingly, and thus should have very low error rate in the early stages of the derivation. Second, coronals and alveolars match after defaults are specified, so they should have a higher error rate, but once again the effect should be weak. In the case of the third row, it is impossible to tell, using Stemberger's method of comparing relative error rates, which is the stronger effect. We can see which is stronger if we consider the true error rate, as reflected in the O/E measure, shown in Table 8.5.

**Table 8.5**: Error rates for labial, underspecified coronal, and specified coronal fricatives.

| group | actual | expected | O/E |
|---|---|---|---|
| lab-alv | 46 | 22.2 | 2.07 |
| alv-cor | 139 | 27.2 | 5.10 |
| lab-cor | 12 | 15.3 | 0.78 |

Table 8.5 shows that the interaction between alveolars and specified coronals is very strong, and the inhibition between labials and specified coronals caused by a mismatch in underlying place of articulation has a weaker effect on the relative error rate. Thus, this is a case where the underspecified feature, which is supposed to be filled in late in the derivation, has a very strong effect on error rate. This is not predicted by Stemberger's two stage model which uses underspecification as a primary predictor of error rate. Further, the discrepancy between error rates of labial consonants with the alveolar consonants and with the specified coronals is not predicted straightforwardly. A null prediction is made by Stemberger (1991b) as he does not give an explicit similarity model. For example, he does not indicate whether contrasts between [Labial] and underspecified place reduces similarity less than contrast between [Labial] and specified [Coronal].

Let us now turn to the case of redundant voicing in sonorants. Stemberger demonstrates there is an effect of the underspecified [-voice] feature on error rate. Table 8.6 gives O/E for interactions between sonorants {m, n, ŋ, l, r} and voiced obstruents {b, d, ð, z, ʒ, dʒ, g} or voiceless obstruents {p, t, θ, s, ʃ, tʃ, k}.

**Table 8.6**: Error rate between sonorants and voiced or voiceless obstruents.

| group | actual | expected | O/E |
|---|---|---|---|
| son-vcd | 77 | 104.1 | 0.74 |
| son-vless | 44 | 276.5 | 0.16 |

Comparing the interactions of sonorants with the voiced obstruents with interactions of sonorants with the voiceless obstruents, we see the overall effect of default voicing increases the error rate by a factor of five. That is much stronger than the effect of shared [Labial] in section 8.2, for example, which only doubles the error rate. Since in Stemberger's model specified [Labial] is present throughout the entire derivation, it should affect error rate more than shared redundant [+voice].

The magnitude of the difference in error rate cannot be observed using simple comparisons of contrast of the type presented in Stemberger (1991b). Differences in error rate are apparently not as simple as differences between the status of features as specified and underspecified.

### *8.3.3 Similarity over the entire inventory.*

As mentioned above, Stemberger (1991b) only considers minimal contrasts involving underspecification and does not consider the effect that underspecification has on similarity, and therefore error rate, across the entire inventory. Stemberger does not present an explicit similarity metric, and opts instead to point out in general how the contrasts he examines could be predicted by similarity and underspecification on a case by case basis. I next model the effects of under-specification on similarity across the entire phoneme inventory, based on Stemberger's discussion.

In his analysis, Stemberger makes the following assumptions about similarity (Thompson 1996):

(92)
1. Shared specified, but not underspecified, features increase similarity.
2. Different specified, but not underspecified, features decrease similarity.[18]
3. For two consonants where one feature is specified and one is underspecified, similarity is less than between two consonants which are both underspecified.

The similarity model which is closest to the natural classes model that satisfies these assumptions is the metric of similarity in Pierrehumbert (1993), repeated in (93). I take this to be the implicit model behind Stemberger's discussion, since this model satisfies all of the assumptions in (92).

(93)

$$\text{similarity} = \frac{\text{shared features}}{\text{shared features} + \text{non}-\text{shared features}}$$

In attempting to implement an underspecified similarity model using this metric, a number of problems were encountered. First, some discussion is needed of what is considered a shared feature and what is considered a non-shared feature when feature matrices use a mix of bivalent and monovalent features. Following Stemberger (1991b), I consider a feature to be shared if it is specified and has the same value for both consonants, for example [+continuant]

[18] Note that this situation does not arise for monovalent features. It can only occur with bivalent or multivalent features.

for /s/ and /f/. I consider a feature to be non-shared under any other circumstance. Some decision must be made as to whether a difference in specification for the same feature counts as one difference or two (Frisch, Broe, & Pierrehumbert 1995). This is a decision as to whether there is a difference between (92[2]) and (92[3]) above. For example, should the difference between [+continuant] and [-continuant] count as just one non-shared feature or two? A similar decision must be made for the difference in monovalent place specifications between, e.g. /p/ and /k/. I count two differences in both cases, eliminating the difference between different specifications for a binary feature and the use of two privative features. This is the only a priori coherent position that can be taken on the issue, as implementing a quantitative difference between non-matching bivalent feature specifications and monovalent feature specifications imparts a computational distinction to a purely notational difference between feature specifications (see chapter 2). Recall that any bivalent contrast can be represented by two monovalent features. Bivalency is merely the special case of complementary features.

A second problem arises in comparing the underspecification model to the similarity model for English speech errors in chapter 4. It would be desirable for the features used to be the same, with the underspecification of some features the only difference. However, underspecification is inherently incompatible with the principles of contrast and individuation used the feature matrix in chapter 4. When those feature assignments are examined for redundancies, it immediately becomes apparent that the usual sorts of feature matrices used when assuming underspecification will not result by eliminating some of the features as redundant. For example, Stemberger (1991b) presents the radically underspecified feature assignments in (94) for the sonorants.

| (94) | m | n | ŋ | l | r | w | y |
|---|---|---|---|---|---|---|---|
| sonorant | | | | + | + | + | + |
| continuant | | | | | | | |
| voice | | | | | | | |
| nasal | + | + | + | | | | |
| Labial | % | | | | % | % | |
| Dorsal | | | % | | | | |
| Coronal | | | | | % | | % |
| anterior | | | | | - | | - |

Compare (94) with my choice of features, based on the principles of individuation and contrast using only monovalent features, given in (95). There are a number of redundancies in (95) which could be eliminated if a minimally contrastive feature matrix was desired. Clearly voicing is predictable for all sonorants. The feature [sonorant] itself is predictable for each consonant, and thus could be underspecified. The traditional specifications used in current feature matrices are the result of gradual refinements in the *SPE* feature system which accomodate the requirements of feature geometry and specific analyses of phonological processes which employed underspecification.

| (95) | m | n | ŋ | l | r | w | y |
|---|---|---|---|---|---|---|---|
| labial | + | | | | + | + | |
| dorsal | | | + | | | | |
| coronal | | | | + | | + | |
| sonorant | + | + | + | + | + | + | + |
| voice | + | + | + | + | + | + | + |
| stop | + | + | + | | | | |
| nasal | + | + | + | | | | |
| alveolar | | + | | + | + | | |
| palatal | | | | | | | + |
| consonantal | + | + | + | + | + | | |
| glide | | | | | | + | + |
| lateral | | | | + | | | |
| rhotic | | | | | + | | |

Stemberger's feature matrix represents a more traditional position (see e.g. Archangeli & Pulleyblank 1989, Kiparsky 1982, Lombardi 1991, Paradis & Prunet 1991), where [sonorant] is considered predictable from [nasal] but is specified for the other sonorants. Stemberger leaves as underspecified the manner features which differentiate the liquids and glides from the nasals ([glide], [lateral], and [rhotic] in my matrix), as they are predictable from the combination of sonorant and the place features. There is no coherent algorithm for eliminating redundancy which can produce this result from (95). For example, the feature sonorant is equally predictable from [lateral] as it is from [nasal].

In addition, in the standard theory, the status of place nodes in the feature geometry tree affects the underspecification of [Coronal], as mentioned above. The presence of a specified coronal dependent like [-anterior] presumably forces the presence of [Coronal] for structural reasons. Since structured specification abandons notions of universal articulator dependencies in feature geometry in favor of describing dependencies between natural classes based on language particular contrasts, feature dependencies of this type play no role in feature assignment. There is thus no algorithm which can produce the correct pattern of underspecification for place features out of (95), either.

Stemberger's feature assignments are based on a survey of contemporary literature involving underspecification as well as feature geometry. I simply adopt his feature assignment as representative of the effect that the assumption of underspecification has on similarity. Similarity was computed using Pierrehumbert's metric for the radically underspecified, contrastively underspecified, and fully specified feature matrices given in section 8.1. Similarity for each consonant pair using each feature matrix is given in Tables 8.7, 8.8, and 8.9.[19]

[19] Note that the similarity of /t/ to itself, which is not relevant for speech errors, is undefined based on the radically underspecified feature matrix. In radical underspecification, all of the features of /t/ are either default or predictable. It thus has no shared or non-shared features with itself, and division by zero occurs in the Pierrehumbert (1993) similarity metric.

| | p | b | f | v | m | t | d | θ | ð | s | z | ʃ | ʒ | tʃ | dʒ | k | g | ŋ | l | r | n | w | y | h |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| p | 1 | | | | | | | | | | | | | | | | | | | | | | | |
| b | 0.5 | 1 | | | | | | | | | | | | | | | | | | | | | | |
| f | 0.5 | 0.33 | 1 | | | | | | | | | | | | | | | | | | | | | |
| v | 0.33 | 0.67 | 0.67 | 1 | | | | | | | | | | | | | | | | | | | | |
| m | 0.5 | 0.33 | 0.33 | 0.25 | 1 | | | | | | | | | | | | | | | | | | | |
| t | 0 | 0 | 0 | 0 | 0 | - | | | | | | | | | | | | | | | | | | |
| d | 0 | 0.5 | 0 | 0.33 | 0 | 0 | 1 | | | | | | | | | | | | | | | | | |
| θ | 0 | 0 | 0.25 | 0.2 | 0 | 0 | 0 | 1 | | | | | | | | | | | | | | | | |
| ð | 0 | 0.2 | 0.2 | 0.4 | 0 | 0 | 0.25 | 0.75 | 1 | | | | | | | | | | | | | | | |
| s | 0 | 0 | 0.5 | 0.33 | 0 | 0 | 0 | 0.33 | 0.25 | 1 | | | | | | | | | | | | | | |
| z | 0 | 0.33 | 0.33 | 0.67 | 0 | 0 | 0.5 | 0.25 | 0.5 | 0.5 | 1 | | | | | | | | | | | | | |
| ʃ | 0 | 0 | 0.25 | 0.2 | 0 | 0 | 0 | 0.5 | 0.4 | 0.33 | 0.25 | 1 | | | | | | | | | | | | |
| ʒ | 0 | 0.2 | 0.2 | 0.4 | 0 | 0 | 0.25 | 0.4 | 0.6 | 0.25 | 0.5 | 0.75 | 1 | | | | | | | | | | | |
| tʃ | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.2 | 0.17 | 0 | 0 | 0.5 | 0.4 | 1 | | | | | | | | | | |
| dʒ | 0 | 0.2 | 0 | 0.17 | 0 | 0 | 0.25 | 0.17 | 0.33 | 0 | 0.2 | 0.4 | 0.6 | 0.75 | 1 | | | | | | | | | |
| k | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | | | | | | | | |
| g | 0 | 0.33 | 0 | 0.25 | 0 | 0 | 0.5 | 0 | 0.2 | 0 | 0.33 | 0 | 0.2 | 0 | 0.2 | 0.5 | 1 | | | | | | | |
| ŋ | 0 | 0 | 0 | 0 | 0.33 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.5 | 0.33 | 1 | | | | | | |
| l | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | | | | | |
| r | 0.25 | 0.2 | 0.2 | 0.17 | 0.2 | 0 | 0 | 0.17 | 0.14 | 0 | 0 | 0.4 | 0.33 | 0.4 | 0.33 | 0 | 0 | 0 | 0.25 | 1 | | | | |
| n | 0 | 0 | 0 | 0 | 0.5 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.5 | 0 | 0 | 1 | | | |
| w | 0.5 | 0.33 | 0.33 | 0.25 | 0.33 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.5 | 0.5 | 0 | 1 | | |
| y | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.2 | 0.17 | 0 | 0 | 0.5 | 0.4 | 0.5 | 0.4 | 0 | 0 | 0 | 0.33 | 0.75 | 0 | 0.25 | 1 | |
| h | 0 | 0 | 0.33 | 0.25 | 0 | 0 | 0 | 0.25 | 0.2 | 0.5 | 0.33 | 0.25 | 0.2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |

**Table 8.7**: Similarity based on radical underspecification.

**Table 8.8**: Similarity based on contrastive underspecification.

| | p | b | f | v | m | t | d | θ | ð | s | z | ʃ | ʒ | tʃ | dʒ | k | g | ŋ | l | r | n | w | y | h |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| p | 1 | | | | | | | | | | | | | | | | | | | | | | | |
| b | 0.6 | 1 | | | | | | | | | | | | | | | | | | | | | | |
| f | 0.6 | 0.33 | 1 | | | | | | | | | | | | | | | | | | | | | |
| v | 0.33 | 0.6 | 0.6 | 1 | | | | | | | | | | | | | | | | | | | | |
| m | 0.17 | 0.17 | 0.17 | 0.17 | 1 | | | | | | | | | | | | | | | | | | | |
| t | 0.6 | 0.33 | 0.33 | 0.14 | 0 | 1 | | | | | | | | | | | | | | | | | | |
| d | 0.33 | 0.6 | 0.14 | 0.33 | 0 | 0.6 | 1 | | | | | | | | | | | | | | | | | |
| θ | 0.29 | 0.13 | 0.5 | 0.29 | 0 | 0.5 | 0.29 | 1 | | | | | | | | | | | | | | | | |
| ð | 0.13 | 0.29 | 0.29 | 0.5 | 0 | 0.29 | 0.5 | 0.67 | 1 | | | | | | | | | | | | | | | |
| s | 0.25 | 0.11 | 0.43 | 0.25 | 0 | 0.43 | 0.25 | 0.57 | 0.38 | 1 | | | | | | | | | | | | | | |
| z | 0.11 | 0.25 | 0.25 | 0.43 | 0 | 0.25 | 0.43 | 0.38 | 0.57 | 0.71 | 1 | | | | | | | | | | | | | |
| ʃ | 0.29 | 0.13 | 0.5 | 0.29 | 0 | 0.5 | 0.29 | 0.67 | 0.43 | 0.57 | 0.38 | 1 | | | | | | | | | | | | |
| ʒ | 0.13 | 0.29 | 0.29 | 0.5 | 0 | 0.29 | 0.5 | 0.43 | 0.67 | 0.38 | 0.57 | 0.67 | 1 | | | | | | | | | | | |
| tʃ | 0.43 | 0.25 | 0.25 | 0.11 | 0 | 0.67 | 0.43 | 0.38 | 0.22 | 0.33 | 0.2 | 0.57 | 0.38 | 1 | | | | | | | | | | |
| dʒ | 0.25 | 0.43 | 0.11 | 0.25 | 0 | 0.43 | 0.67 | 0.22 | 0.38 | 0.2 | 0.33 | 0.38 | 0.57 | 0.71 | 1 | | | | | | | | | |
| k | 0.6 | 0.33 | 0.33 | 0.14 | 0 | 0.6 | 0.33 | 0.29 | 0.13 | 0.25 | 0.11 | 0.29 | 0.13 | 0.43 | 0.25 | 1 | | | | | | | | |
| g | 0.33 | 0.6 | 0.14 | 0.33 | 0 | 0.33 | 0.6 | 0.13 | 0.29 | 0.11 | 0.25 | 0.13 | 0.29 | 0.25 | 0.43 | 0.6 | 1 | | | | | | | |
| ŋ | 0 | 0 | 0 | 0 | 0.5 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.17 | 0.17 | 1 | | | | | | |
| l | 0 | 0 | 0 | 0 | 0.17 | 0.14 | 0.14 | 0.13 | 0.13 | 0.25 | 0.25 | 0.13 | 0.13 | 0.11 | 0.11 | 0 | 0 | 0.17 | 1 | | | | | |
| r | 0.13 | 0.13 | 0.13 | 0.13 | 0.33 | 0.13 | 0.13 | 0.11 | 0.11 | 0.1 | 0.1 | 0.25 | 0.25 | 0.22 | 0.22 | 0 | 0 | 0.14 | 0.29 | 1 | | | | |
| n | 0 | 0 | 0 | 0 | 0.4 | 0.14 | 0.14 | 0.13 | 0.13 | 0.25 | 0.25 | 0.13 | 0.13 | 0.11 | 0.11 | 0 | 0 | 0.4 | 0.6 | 0.29 | 1 | | | |
| w | 0.17 | 0.17 | 0.17 | 0.17 | 0.5 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.2 | 0.17 | 0.33 | 0.17 | 1 | | |
| y | 0 | 0 | 0 | 0 | 0.17 | 0.14 | 0.14 | 0.13 | 0.13 | 0.11 | 0.11 | 0.29 | 0.29 | 0.25 | 0.25 | 0 | 0 | 0.17 | 0.33 | 0.5 | 0.33 | 0.4 | 1 | |
| h | 0.2 | 0.2 | 0.2 | 0.2 | 0 | 0.2 | 0.2 | 0.17 | 0.17 | 0.14 | 0.14 | 0.17 | 0.17 | 0.14 | 0.14 | 0.2 | 0.2 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |

**Table 8.9**: Similarity based on full specification.

| | p | b | f | v | m | t | d | θ | ð | s | z | ʃ | ʒ | tʃ | dʒ | k | g | ŋ | l | r | n | w | y | h |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| p | 1 | | | | | | | | | | | | | | | | | | | | | | | |
| b | 0.6 | 1 | | | | | | | | | | | | | | | | | | | | | | |
| f | 0.6 | 0.33 | 1 | | | | | | | | | | | | | | | | | | | | | |
| v | 0.33 | 0.6 | 0.6 | 1 | | | | | | | | | | | | | | | | | | | | |
| m | 0.25 | 0.43 | 0.11 | 0.25 | 1 | | | | | | | | | | | | | | | | | | | |
| t | 0.5 | 0.29 | 0.29 | 0.13 | 0.1 | 1 | | | | | | | | | | | | | | | | | | |
| d | 0.29 | 0.5 | 0.13 | 0.29 | 0.22 | 0.67 | 1 | | | | | | | | | | | | | | | | | |
| θ | 0.25 | 0.11 | 0.43 | 0.25 | 0 | 0.38 | 0.22 | 1 | | | | | | | | | | | | | | | | |
| ð | 0.11 | 0.25 | 0.25 | 0.429 | 0.09 | 0.22 | 0.38 | 0.71 | 1 | | | | | | | | | | | | | | | |
| s | 0.25 | 0.11 | 0.43 | 0.25 | 0 | 0.57 | 0.38 | 0.5 | 0.33 | 1 | | | | | | | | | | | | | | |
| z | 0.11 | 0.25 | 0.25 | 0.43 | 0.09 | 0.38 | 0.57 | 0.33 | 0.5 | 0.71 | 1 | | | | | | | | | | | | | |
| ʃ | 0.25 | 0.11 | 0.43 | 0.25 | 0 | 0.57 | 0.38 | 0.71 | 0.5 | 0.71 | 0.5 | 1 | | | | | | | | | | | | |
| ʒ | 0.11 | 0.25 | 0.25 | 0.43 | 0.09 | 0.38 | 0.57 | 0.5 | 0.71 | 0.5 | 0.71 | 0.71 | 1 | | | | | | | | | | | |
| tʃ | 0.38 | 0.22 | 0.22 | 0.1 | 0.08 | 0.71 | 0.5 | 0.44 | 0.3 | 0.44 | 0.3 | 0.63 | 0.44 | 1 | | | | | | | | | | |
| dʒ | 0.22 | 0.38 | 0.1 | 0.22 | 0.18 | 0.5 | 0.71 | 0.3 | 0.44 | 0.3 | 0.44 | 0.44 | 0.63 | 0.75 | 1 | | | | | | | | | |
| k | 0.6 | 0.33 | 0.33 | 0.14 | 0.11 | 0.5 | 0.29 | 0.25 | 0.11 | 0.25 | 0.11 | 0.25 | 0.11 | 0.38 | 0.22 | 1 | | | | | | | | |
| g | 0.33 | 0.6 | 0.14 | 0.33 | 0.25 | 0.29 | 0.5 | 0.11 | 0.25 | 0.11 | 0.25 | 0.11 | 0.25 | 0.22 | 0.38 | 0.6 | 1 | | | | | | | |
| ŋ | 0.11 | 0.25 | 0 | 0.11 | 0.71 | 0.1 | 0.22 | 0 | 0.09 | 0 | 0.091 | 0 | 0.09 | 0.08 | 0.18 | 0.25 | 0.43 | 1 | | | | | | |
| l | 0.08 | 0.18 | 0 | 0.08 | 0.36 | 0.27 | 0.4 | 0.07 | 0.15 | 0.25 | 0.36 | 0.15 | 0.25 | 0.23 | 0.33 | 0.08 | 0.18 | 0.36 | 1 | | | | | |
| r | 0.08 | 0.17 | 0.17 | 0.27 | 0.33 | 0.15 | 0.25 | 0.23 | 0.33 | 0.23 | 0.33 | 0.33 | 0.45 | 0.21 | 0.31 | 0 | 0.08 | 0.23 | 0.46 | 1 | | | | |
| n | 0.09 | 0.2 | 0 | 0.09 | 0.56 | 0.3 | 0.44 | 0.08 | 0.17 | 0.27 | 0.4 | 0.17 | 0.27 | 0.25 | 0.36 | 0.09 | 0.2 | 0.56 | 0.7 | 0.38 | 1 | | | |
| w | 0.11 | 0.25 | 0.25 | 0.43 | 0.33 | 0 | 0.1 | 0.09 | 0.2 | 0.09 | 0.2 | 0.09 | 0.2 | 0 | 0.08 | 0 | 0.11 | 0.2 | 0.25 | 0.45 | 0.17 | 1 | | |
| y | 0 | 0.09 | 0.09 | 0.2 | 0.17 | 0.18 | 0.3 | 0.27 | 0.4 | 0.27 | 0.4 | 0.4 | 0.56 | 0.25 | 0.36 | 0 | 0.09 | 0.17 | 0.42 | 0.64 | 0.33 | 0.56 | 1 | |
| h | 0.33 | 0.14 | 0.6 | 0.33 | 0 | 0.29 | 0.13 | 0.43 | 0.25 | 0.43 | 0.25 | 0.43 | 0.25 | 0.22 | 0.1 | 0.33 | 0.14 | 0 | 0 | 0.08 | 0 | 0.11 | 0.09 | 1 |

Stemberger (1991b) claims that specified features have the greatest effect on similarity, and that underspecified features have a smaller influence. Stemberger's model thus involves errors at two stages, before and after specification. Other plausible models of similarity involve a single stage using specified features under either radical underspecification or contrastive underspecification, and a model based only on the fully specified similarity computed with analagous features. Finally, all of these models can be compared to the natural classes similarity model.

The five models are compared by fitting them to Stemberger's (1991a) consonant confusion matrix. Each model is fit with a non-linear regression to predict the actual number of errors for each pair, with the expected number of errors and similarity as predictor variables, as in chapter 4. The regression equation is shown in (96).

(96) Observed = Expected × (A + B × Similarity)

In the case of Stemberger's two stage model, the equation is:

(97) Observed = Expected × (A + B × Underspecified Sim + C × Specified Sim)

In Stemberger's two stage model, both the specified and underspecified similarity can each contribute to the error rate.

Table 8.10 shows the fit for each similarity model, along with the parameters for the corresponding non-linear regression. Among the models based on the standard linguistic feature matrix, the fully specified model has the best fit, followed by the contrastive underspecification model, and then the radical underspecification model. Of course, the two stage model provides a better fit than either the radical underspecification model or the specified feature model, since it is equivalent to either of these models with an additional parameter. However, the natural classes similarity model is far superior to the two stage model, even though it has one less free parameter. The model based only on radical underspecification fared the poorest of all of the models of the speech error data, indicating that Stemberger's emphasis on underspecification, and the primary role of specified features in determining error rate, does not account for the majority of the speech error data. The fully specified feature model fit better, where no special status was given to redundant or default features.

**Table 8.10**: Five models of Stemberger's corpus of consonant speech errors.

| Model | $R^2$ | Parameters |
|---|---|---|
| Radical Underspecification | 0.35 | A (constant) = 0.65, B (similarity) = 5.4 |
| Contrastive Underspecification | 0.44 | A = -0.19, B = 5.1 |
| Full Specification | 0.49 | A = 0.65, B = 5.8 |
| Two-stage model | 0.57 | A = -0.57, B (underspec) = 3.8, C (spec) = 4.9 |
| Natural classes model | 0.72 | A = -0.69, B = 9.88 |

# CHAPTER 9
# Underspecification in Speech Errors II: Anti-Frequency Effects

In another paper, Stemberger (1991a) analyzed asymmetries in speech error rates. Stemberger proposes that there is a general ADDITION BIAS in language production, in which errors tend to add segments or features rather than deleting segments or features. In combination with underspecification, the addition bias predicts that errors are more likely to produce specified segments than underspecified segments. Stemberger (1991a) demonstrates that this asymmetry is found using a SLIPS experiment. I performed a replication of Stemberger's experiment using the tongue twister task as in Experiment 1 (chapter 6). I found that the asymmetries in error rates which Stemberger (1991a) attributed to radical underspecification are influenced by the status of the error outcome as either a word or non-word outcome. Thus, these asymmetries are not due to a property of the erring consonant alone, such as underspecification, and are instead process dependent.

## *9.1 Review of Stemberger (1991a)*

In this section, I review more of Stemberger's arguments for the influence of radical underspecification in adult speech errors. Stemberger (1991a) observes that, in general, high frequency items are processed more accurately than low frequency items in both perception (e.g. Treisman 1978) and production (e.g. Stemberger & MacWhinney 1986, Dell 1990). However, he presents a number of cases of an apparent 'anti-frequency' effect in speech production, where less frequent consonants replace more frequent ones, rather than the reverse. For example, Shattuck-Hufnagel & Klatt (1979) originally noted that, in naturally occurring errors, the low frequency palatal consonants are more often intrusions on target high frequency alveolars than the reverse. Shattuck-Hufnagel & Klatt (1979) concluded that there is something special about palatal consonants which leads to this error asymmetry, calling it the PALATAL BIAS. Stemberger proposes that this asymmetry results from the combination of two more general principles. First, following Stemberger & Treiman (1986), he proposes that there is an ADDITION BIAS in speech production. Stemberger & Treiman (1986) found that errors which formed a cluster by adding a consonant to a singleton were more common than errors which formed a singleton from a cluster by deleting a consonant. Second, Stemberger proposes once again that speech errors can occur at a point in the derivation where segments are underspecified. In the case of palatals versus alveolars, alveolar consonants are underspecified for the feature [+anterior], while palatals are specified [-anterior]. Thus, an error which replaces /s/ with /ʃ/ would involve the addition of the [-anterior] feature to /s/, while an error which replaces /ʃ/ with /s/ would involve the deletion of the [-anterior] feature (with [+anterior] to filled in later by a default rule as discussed in chapter 8). The combination of underspecification and the addition bias accounts for the palatal bias.

Stemberger (1991a) replicated the palatal bias using the SLIPS error induction technique (Motley & Baars 1975) to support his claim. Recall from chapter 6 that, in SLIPS, subjects are primed to make consonant exchange errors on the word onsets of monosyllabic words. Subjects look at word pairs displayed on a screen one at a time. After some word pairs, they are shown a cue which indicates they are to produce the immediately preceding pair. The word pairs of interest are preceded by word pairs that prime the subjects to produce consonants in a certain order. In the target pair, that order is reversed. A sample stimulus sequence for Stemberger's SLIPS experiment on palatals and alveolars is shown in (98). In this example, the target phoneme pair is /s/-/ʃ/.

| | | |
|---|---|---|
| (98) | Priming pairs | muck rift |
| | | sub shin |
| | | sum ship |
| | Target pair | shuck sift |

In 103 errors produced by 24 subjects on 27 target pairs, Stemberger found 65 errors where a palatal {ʃ, tʃ, dʒ} replaced an alveolar {s, t, d}, and only 38 errors where an alveolar replaced a palatal.

To demonstrate the addition bias and verify the results of Stemberger & Treiman (1986), Stemberger (1991a) also used SLIPS to replicate the asymmetry in error rate between singleton consonants and consonant clusters. Stemberger & Treiman (1986) found that consonants are added to singletons to make clusters more frequently than consonants were deleted from clusters to create singletons. Two sample stimuli, one priming for a deletion error, and a comparable one priming for an addition error, are shown in (99). All of the addition and deletion errors were primed to occur on a liquid or glide following an obstruent.

| | | | |
|---|---|---|---|
| (99) | Priming pairs | spite bowl | bowl spite |
| | | tile squirrel | trout squirrel |
| | | type host | trove host |
| | Target pair | trite toll | toll trite |

In 95 errors produced by 30 subjects on 44 stimuli, Stemberger found 72 errors where a cluster was created by the addition of a liquid or glide, and 23 errors where a singleton onset was created by the deletion of a liquid or glide.

The addition bias of Stemberger & Treiman (1986) was originally proposed to deal with the preference for addition of an entire segment to create a cluster. Stemberger (1991a) generalizes the addition bias to also include addition of a feature to an underspecified segment. He tests this generalization by looking for asymmetries between consonants which are specified and underspecified for place of articulation, manner of articulation, and voicing. The feature contrasts Stemberger (1991a) examined using SLIPS experiments are:

| (100) | Contrast | Specified | Underspecified |
|---|---|---|---|
| a. | Place of articulation I | [Labial], e.g. /p/ | [Coronal], e.g. /t/ |
| b | Place of articulation II | [Dorsal], e.g. /k/ | [Coronal], e.g. /t/ |
| c. | Manner I | [+cont], e.g. /s/ | [-cont], e.g. /t/ |
| d | Manner II | [+nas], e.g. /n/ | [-nas], e.g. /t/ |
| e. | Voicing | [+voice], e.g. /d/ | [-voice], e.g. /t/ |

For each contrast between specified and underspecified features, Stemberger created SLIPS stimuli over a variety of consonant pairs which minimally contrasted on the target features. The results from Stemberger's experiments are summarized in Table 9.1. For each category, the specified and underspecified features are given, along with the consonant pairs Stemberger examined. Total errors aggregated across subjects and consonant pairs are shown for errors where the specified consonant replaced the underspecified one (an 'addition' of the specified feature) and where the underspecified consonant replaced the specified one (a 'deletion' of the specified feature).

**Table 9.1**: Results of SLIPS experiments in Stemberger (1991a).

| Category | Specified/ Underspecified | Consonant pairs (Specified/Underspecified) | 'addition' errors | 'deletion' errors |
|---|---|---|---|---|
| Place (I) | [Labial]/[Coronal] | p/t, b/d, f/s, v/z, m/n | 59 | 32 |
| Place (II) | [Velar]/[Coronal] | k/t, g/d | 4 | 2 |
| Manner (I) | [+cont]/[-cont] | f/p, v/b, s/t, θ/t, z/d | 43 | 19 |
| Manner (II) | [+nas]/[-nas] | m/p, n/t, m/b, n/d | 29 | 14 |
| Voicing | [+voice]/[-voice] | b/p, d/t, g/k, v/f, z/s, dʒ/tʃ | 21 | 9 |

Every case is consistent with Stemberger's hypothesis: there were more errors where a feature specification was added to an underspecified consonant than errors where a specified feature was lost, resulting in default specification.

Stemberger also examined his corpus of naturally occurring speech errors for evidence of asymmetries. He found significant asymmetries only between the consonant pairs /p/-/f/ and /d/-/g/. These asymmetries are consistent with his account of the experimental data, though we might consider it surprising that no other asymmetries were clearly present. An account of Stemberger's asymmetries which can also explain why they were difficult to find in the naturally occurring corpus has an additional advantage over the underspecification account. Such an account, which does not employ underspecification, is presented in section 9.3.

### *9.2 Experiment 2: Asymmetries in Speech Errors*

Stemberger's results are compelling because some asymmetry was found in every experiment, and only consistent asymmetries were found in the natural error corpus. In addition, Stemberger's account is theoretically appealing since he gives a unified account to the palatal bias and asymmetries involving underspecification using only the addition bias, which is independently needed to account for the pattern of errors between consonant clusters and singletons. In the previous section, we saw that similarity based on underspecification is a poor predictor of error rate. Therefore, before Stemberger's results can be accepted, they must be shown to be robust. In this section, I present the results of an experiment of my own which was designed to replicate Stemberger's findings. I found evidence for asymmetries only in some of the cases above, and I also found crucial evidence that the asymmetries cannot be attributed to underspecification, as I discuss below.

### *9.2.1 Materials.*

Closely following the design of Stemberger (1991a), I attempted to replicate the finding of an addition bias between 'specified' and 'underspecified' consonant pairs. The materials were designed to examine the same contrasts as Stemberger (1991a). These contrasts are presented along with all of the consonant pairs I examined in (101). In most cases, these are the same pairs as studied by Stemberger. The specified consonant is given first in each pair, and the underspecified consonant is given second.

| (101) | | Contrast | Consonant pairs |
|---|---|---|---|
| | a. | Palatal vs. alveolar | tʃ/t, dʒ/d, ʃ/s |
| | b | Labial/dorsal vs. coronal | p/t, b/d, f/s, m/n, k/t, g/d |
| | c. | Fricative vs. stop | f/p, v/b, s/t, z/d |
| | d | Nasal vs. stop | m/p, m/b, n/t, n/d |
| | e. | Voiced vs. voiceless | b/p, v/f, d/t, z/s, g/k |

To test whether Stemberger's asymmetries would generalize to a different error inducing task, I followed the tongue twister paradigm of Shattuck-Hufnagel (1983, 1987, 1992), also used in Experiment 1. A sample twister for the consonant pair /z/-/s/, which fall within the voicing category, is shown in (102).

(102) sit zap zoo sip

All of the tongue twisters in this experiment consisted of four monosyllabic words with the target pair as word onsets. In general, each word in the tongue twister had a different vowel, but if a vowel was repeated it was never repeated on adjacent words.[20] Syllable codas were not strictly controlled, but were picked to be similar as much as possible in order to increase the overall error rate (Dell 1984). The words in the tongue twisters were balanced for frequency, which sometimes affected the choice of available vowels and coda consonants.

For each consonant pair, I constructed four tongue twisters. Two of these twisters formed a complementary set, where the error outcomes for one twister were used to create the second twister. The two word outcome twisters for /z/-/s/ are:

---

[20] Using repeated vowels would actually increase the error rate (Dell 1984 and others), but different vowels were used for two reasons. First, using different vowels allows an error on the word initial consonant to be unambiguously classified as a single C error and not a CV or entire word error. Subjects often interrupt themselves when they make errors, and having distinct vowel transitions for the different consonant pairs often clarifies whether the interrupted error is just a C or whether it is a larger unit in these cases. Second, since asymmetries based on underspecification are controlled only for consonants, confounding effects due to the entire word or the following vowel were to be avoided.

(103) sit zap zoo sip
zit sap sue zip

The other two twisters for each consonant pair had error outcomes which were not words. The two twisters of this type for /z/-/s/ are given in (104).

(104) sung zone Zeus seam
zig suck sank zilch

Each pair of twisters uses two balanced patterns of the word onsets, /s/-/z/-/z/-/s/ and /z/-/s/-/s/-/z/. A complete list of the 88 stimuli for this experiment is given in section 9.4 at the end of the chapter.

In the pair of stimuli where the error outcomes were words, the words and their corresponding error outcome were also balanced for frequency. The other pair, with non-word outcomes, represent a second test of the robustness of Stemberger's asymmetry findings. In the SLIPS experiments in Stemberger (1991a) the error outcomes were always words. The presentation of the results below shows that this is a crucial difference, as the non-word outcome twisters show little evidence of asymmetry, indicating that the presence of asymmetries in Stemberger's experiments are crucially dependent upon the particular task and stimuli used, rather than inherent differences between the consonants. I present an alternative account below.

### *9.2.2 Method.*

The method for this experiment is a four word tongue twister paradigm nearly identical to Experiment 1. The randomization procedure was slightly different. The stimuli were broken into two groups. Half of the stimuli for each consonant pair were placed in each group, one stimulus for each consonant pair with word outcomes and one with non-word outcomes. Within each category (palatal vs. alveolar, labial or dorsal vs. coronal, fricative vs. stop, nasal vs. stop, voiced vs. voiceless) the stimuli were balanced for the order of specified and underspecified target consonants. The stimuli were randomized within each group so that no target consonants were repeated in adjacent stimuli. Half of the subjects read one group of stimuli first and half of the subjects read the other group first.

### *9.2.3 Subjects.*

The subjects were 21 Northwestern undergraduates, primarily from introductory classes in linguistics and cognitive science. All of the subjects were monolingual speakers of English. Subjects were paid for their participation in the experiment.

### *9.2.4 Tabulation.*

Error tabulation followed the same procedure as Experiment 1.

### *9.2.5 Results and discussion.*

Table 9.2 shows error totals aggregated across consonants pairs and subjects for each category based on the type of error outcome: word or non-word. Some of these totals are broken down in more detail in the analysis below. In Table 9.2 and elsewhere, exchanges are marked as being on the specified or underspecified consonant based on the first error in the exchange, as was done in Experiment 1 (see chapter 6).

**Table 9.2**: Aggregate error counts for each category in word and non-word outcomes

| Word outcome: | Target | | |
|---|---|---|---|
| Category | Specified | Underspecified | Total |
| Palatal | 55 | 88 | 143 |
| Place | 41 | 47 | 88 |
| Fricative | 34 | 83 | 117 |
| Nasal | 36 | 51 | 87 |
| Voicing | 58 | 95 | 153 |
| Total | 224 | 364 | 588 |
| Non-word outcome: | Target | | |
| Category | Specified | Underspecified | Total |
| Palatal | 58 | 50 | 108 |
| Place | 22 | 30 | 52 |
| Fricative | 41 | 49 | 90 |
| Nasal | 31 | 41 | 72 |
| Voicing | 37 | 41 | 78 |
| Total | 189 | 211 | 400 |

In examining Table 9.2 for asymmetries, the contrast between the word outcome condition and the non-word outcome condition is striking.[21] While the word outcome condition

[21] The reader might first notice the overall greater number of errors in the word outcome case than in the non-word outcome case. This 'lexical bias effect' has been demonstrated in naturally occurring error corpora (Dell & Reich 1981) and in the SLIPS paradigm (Motley & Baars 1975). Given the results I present below, it would be ideal to examine both word outcomes and non-word outcomes in speech error experiments to insure generality. I abstract away from the differences in overall counts between the two outcome types below and focus only on the presence and absence of asymmetries.

does show apparent asymmetries in the palatal vs. coronal, fricative vs. stop, and voiced vs. voiceless categories, there are no asymmetries for any category in the non-word outcome case.

The data for each outcome were tested within each category for statistical significance using two factor ANOVA with error type (specified target consonant or underspecified target consonant) and subject as factor, and with error type and stimulus item as factors. The tests for asymmetry were based on the combined substitution and exchange errors. If exchanges are not included, comparable results are obtained. For word outcomes, the palatal asymmetry is significant with subject ($F(1,20) = 13.4$, $p = 0.002$) and item ($F(1,5) = 5.1$, $p = 0.07$) as factors. The fricative asymmetry is also significant with subject ($F(1,20) = 30.6$, $p = 0.001$) and item ($F(1,7) = 7.4$, $p = 0.03$) as factors. The voicing asymmetry is significant with subject as factor ($F(1,20) = 25.2$, $p = 0.001$), but surprisingly it was not significant with item as factor ($F(1,9) = 2.7$, $p = 0.14$). I return to this below. The nasal asymmetry was close to significant with subject as factor ($F(1,20) = 2.7$, $p = 0.10$) and with item as factor ($F(1,7) = 3.59$, $p = 0.11$). However, the nasal asymmetry is close only when exchanges are included (there were 10 exchanges where the first error targeted a stop and only 2 where the first error targeted a nasal). If the exchanges are not included, the asymmetry is far from significant.

The lack of significance for the voicing asymmetry with consonant pair as factor in the word outcome case is explained when the voicing contrast data are examined in greater detail. Table 9.3 shows error totals aggregated across subjects in the word outcome condition for each consonant pair in the voicing category. Notice that nearly all of the asymmetry for the category comes from the single consonant pair /z/-/s/. There were 50 substitution errors where /z/ replaced /s/ in the word outcome condition, while there were only 9 errors where /s/ replaced /z/. Among the remaining pairs there were 51 errors where the specified consonant replaced the underspecified one, and 43 errors where the underspecified consonant replaced the specified one, a nonsignificant asymmetry.

**Table 9.3**: Errors for each consonant pair in the word outcome voicing category

| | Target | | |
|---|---|---|---|
| Pair | Specified | Underspecified | Total |
| p/b | 14 | 14 | 28 |
| f/v | 7 | 15 | 22 |
| t/d | 8 | 13 | 21 |
| s/z | 9 | 50 | 59 |
| k/g | 14 | 9 | 23 |
| Total | 52 | 101 | 153 |

The disparity between the word and non-word outcome condition provides decisive evidence against underspecification as the source of asymmetry in the palatal and fricative categories, between /z/ and /s/, and in the results of Stemberger (1991a). If underspecification were involved, the status of the error outcome as a word or non-word should be irrelevant, as underspecification is a property of the individual segment and its prosodic position. In addition, asymmetry was found in the word case only for palatals vs. alveolars, fricatives vs. stops, and /z/ vs. /s/. No support for an asymmetry was found for labial or dorsal vs. coronal place, for nasals vs. stops, or for the other voiced vs. voiceless pairs. The contrast between my results and Stemberger's suggests that the asymmetries he found are at least dependent upon the error inducing task and the particular stimulus words used. A similar result was found by Levitt & Healy (1985), who tested the palatal bias using nonsense syllables in a SLIPS paradigm and were unable to find an effect. Their finding agrees with the non-word outcome data in my experiment.

Since underspecification is not responsible for the asymmetries in error rate found by Stemberger, what is? The status of the error outcome as a word or non-word is apparently crucial. Stemberger (1991a) informally characterizes the effects as 'anti-frequency' effects, though that cannot be the account given the lack of asymmetry in the non-word outcome case, and the fact that frequency effects have been demonstrated and replicated (Motley, Baars & Camden 1975; Levitt & Healy 1985). Note also that the fact that the error must be a word outcome accounts for why Stemberger (1991a) found only a few significant asymmetries in his corpus of naturally occurring errors. These errors would be a mix of word and non-word outcomes, and thus any effect of asymmetry in the cases which were word outcomes would be diluted.

### *9.3 Similarity Neighborhoods in Speech Production*

I propose that the 'anti-frequency' effects found by Stemberger (1991a) are the result of the effects of the SIMILARITY NEIGHBORHOODS (Luce 1986) of the words involved in the error. The similarity neighborhood of a word is the set of words which are phonologically similar to that word. The similarity neighborhood of a word is typically determined by the SINGLE PHONEME SUBSTITUTION RULE: any word which can be generated by substituting a single phoneme for a phoneme in the word, or by adding or deleting a single phoneme to the word is a member of the word's neighborhood (Luce, Pisoni, & Goldinger 1990). The words in the similarity neighborhood of a target word are presumably activated when the word is processed. Words which have many neighbors are in 'dense' neighborhoods, while words which have few neighbors are in 'sparse' neighborhoods. It has been shown in speech perception that words in dense neighborhoods are harder to identify correctly in noise than words in sparse neighborhoods (Luce 1986; Goldinger, Luce, & Pisoni 1989; Luce, Pisoni, & Goldinger 1990). Similarity neighborhood effects have also been demonstrated in lexical decision tasks. Non-words that lie within a dense neighborhood of words are more difficult to identify as non-words than non-words in a sparse neighborhood of words (Luce 1986).

In the word outcome case of the tongue twister experiment, I suspect that the error rates were affected by the relative similarity neighborhoods of the target word and the intruding word. I assume that both the target word and the intruding word are activated by the tongue twister task, along the lines of the Dell (1986) spreading activation model. Let's take /z/-/s/ as a concrete example. Since /s/ is a very high frequency phoneme, it presumably has denser word neighborhoods than /z/. There are many words that begin with /s/. There are relatively few words that begin with /z/. Thus, in general, the similarity neighborhoods of words beginning with /s/ are denser than the similarity neighborhoods of words beginning with /z/. If a word in a sparse neighborhood is easier to produce than a word in a dense neighborhood (parallel to the perception case), then the outcome which begins with /z/ is at a distinct advantage, both due to its own neighborhood characteristics and due to the neighborhood characteristics of its competitor, /s/.

The neighborhood account of the asymmetry in speech errors is supported by a post-hoc analysis of the stimuli used in Experiment 2. I retrieved the neighborhood densities for all words in the experimental stimuli that could be found in the on-line Webster's Pocket Dictionary used in Luce (1986), known as the Hoosier Mental Lexicon (Nusbaum, Pisoni, & Davis 1984). Words which were not found in the on-line dictionary are marked with an asterisk in the presentation of the stimuli in section 9.4. The mean difference in neighborhood density between the words containing 'specified' and 'underspecified' consonants in each of the categories is given in Table 9.4. Large differences in word density are found for all but the place of articulation contrasts. This is consistent with the results of the experiment, as some evidence of asymmetry was found in Experiment 2 in every category except place of articulation.

**Table 9.4**: Difference in neighborhood density between 'specified' and 'underspecified' stimuli in experiment 2.

| Category | Mean ΔDensity |
|---|---|
| Palatal | 8.00 |
| Place | -2.43 |
| Fricative | 8.54 |
| Nasal | 5.62 |
| Voice | 6.54 |

The influence of similarity neighborhoods also accounts for the difference between word and non-word outcome cases in Experiment 2. In the case of non-word outcomes, there is no competing word which has been activated by the onset consonant pattern in the tongue twister task. In that case, some neighborhood effects might still emerge, but they would be weaker than the word outcome case, in which a double effect of neighborhood density exists. In fact, non-significant asymmetries were found in the non-word outcome case in Experiment 2.

Conversely, the SLIPS task used by Stemberger (1991a) involves priming the error outcome in advance of the stimulus. This priming may have enhanced the neighborhood effects by activating other members of the neighborhood prior to the stimulus presentation. The priming pairs presumably heightened the activation of the competing neighborhoods, and this may have enhanced the asymmetries in his experimental results. Thus, the fact that Stemberger found more asymmetries than I did in Experiment 2 are also explained under the neighborhood hypothesis, but not with the underspecification hypothesis. Priming should have no effect on underspecification.

Neighborhood effects in speech production have been the subject of two recent papers (Vitevich 1996a, 1996b). In a study of a small corpus of naturally occurring errors, Vitevich (1996a) found that the neighborhood density of words which contained targets of consonant exchange errors was greater than the neighborhood density of an equivalently sized sample of control words taken randomly from the dictionary. Vitevich (1996b) found that phonological whole word substitutions (so-called MALAPROPISMS) also occurred on target words with higher neighborhood density than a set of randomly selected controls. Finally, Vitevich (1996a) also presents the results of a SLIPS experiment which controlled for the neighborhood density of the word pairs. In contrast to the above results, he found that stimuli which contained words in sparse neighborhoods had a higher error rate than stimuli which contained words in dense neighborhoods.

In Vitevich's SLIPS experiments, both of the words in the stimulus pair were either high or low density. By comparison, in Experiment 2 and in Stemberger's SLIPS experiments, the stimuli apparently contained a mix of high density and low density words, roughly corresponding to the 'underspecified' and 'specified' status of the onset consonants. Thus, the results of Vitevich's SLIPS experiment do not contradict the analysis given above, as the stimuli are not comparable. The results of his studies of naturally occurring errors support my analysis by showing that words in dense neighborhoods have higher error rates. A complete analysis of neighborhood effects in speech production is beyond the scope of this thesis, and is left here as an outstanding research problem.

### *9.4 Stimuli from Experiment 2*

Words which were not found in the Hoosier Mental Lexicon (Nusbaum, Pisoni, & Davis 1984) and thus were not included in the post hoc analysis of similarity neighborhoods in section 9.3 are marked with an asterisk.

| Category | Targets | Outcome | Stimulus | | | |
|---|---|---|---|---|---|---|
| palatal | s/ʃ | N | sulk | shop | shirk | sink |
| palatal | s/ʃ | N | shape | soup | soil | shard |
| palatal | s/ʃ | W | seat | shack | short | sin |
| palatal | s/ʃ | W | sheet | sack* | sort | shin |
| palatal | t/tʃ | N | teach | chat | chirp | temp* |
| palatal | t/tʃ | N | chimp | tarp* | toss | chase |
| palatal | t/tʃ | W | tore* | chin | chuck | tip |
| palatal | t/tʃ | W | chore | tin | tuck | chip |
| palatal | d/dʒ | N | dad | judge | jilt | doll |
| palatal | d/dʒ | N | jinx | duck | dial | join |
| palatal | d/dʒ | W | dunk | jock* | Jean* | deaf |
| palatal | d/dʒ | W | junk | dock | dean | Jeff* |
| | | | | | | |
| nasal | t/n | N | Turk | nuke | niche | tax |
| nasal | t/n | N | nymph | tough | tooth | net |
| nasal | t/n | W | tape | nap | nick | ton |
| nasal | t/n | W | nape | tap | tick | none |

| | | | | | | |
|---|---|---|---|---|---|---|
| nasal | d/n | N | dove | nerve | neat | death |
| nasal | d/n | N | nurse | dose | dork* | nod |
| nasal | d/n | W | deck | nice* | knave | done |
| nasal | d/n | W | neck | dice | Dave* | nun |
| nasal | p/m | N | peach | mush | moot | pert |
| nasal | p/m | N | mob | peep | purse | mask |
| nasal | p/m | W | poll | mill* | mike | peak |
| nasal | p/m | W | mole | pill | pike | meek |
| nasal | b/m | N | bow | mine | mint | band |
| nasal | b/m | N | mist | bus | baste | moose |
| nasal | b/m | W | bake | muck | match | bar |
| nasal | b/m | W | make | buck | batch | mar |
| | | | | | | |
| fricative | s/t | N | talc | soak | sift | task |
| fricative | s/t | N | sow* | tare | torn | sum |
| fricative | s/t | W | tight | saint | silt | tag |
| fricative | s/t | W | sight | taint | tilt | sag |
| fricative | z/d | N | dome | zinc | Zack* | dusk |
| fricative | z/d | N | zest | dike | dank | zag* |
| fricative | z/d | W | doom | zing | Zen | deal |
| fricative | z/d | W | zoom* | ding* | den | zeal |
| fricative | f/p | N | patch | fudge | fund | perk |
| fricative | f/p | N | fiend | pint | peck | fake |
| fricative | f/p | W | pat | faint | fawn* | pin |
| fricative | f/p | W | fat | paint | pawn | fin |
| fricative | v/b | N | beige | vouch | voice | bush |
| fricative | v/b | N | vague | bog | bump | verb |
| fricative | v/b | W | bane | veer | vow | bat |
| fricative | v/b | W | vein | beer | bough | vat |
| | | | | | | |
| voice | t/d | N | tort | date | dump | top |
| voice | t/d | N | dance | taut | tint | daze |
| voice | t/d | W | till | duel | dart | tote |
| voice | t/d | W | dill | tool | tart | dote |
| voice | p/b | N | pot | bean | bum | pant |
| voice | p/b | N | bank | pink | pork | bask |
| voice | p/b | W | post | beak | bunt | pounce |
| voice | p/b | W | boast | peek | punt | bounce |
| voice | s/z | N | sung | zone* | Zeus* | seam |
| voice | s/z | N | zig* | suck | sank | zilch* |
| voice | s/z | W | sit | zap* | zoo* | sip |
| voice | s/z | W | zit | sap | sue | zip* |
| voice | f/v | N | foil* | valve | verse | far |

| | | | | | | |
|---|---|---|---|---|---|---|
| voice | f/v | N | vex | fix | farce | vice |
| voice | f/v | W | file* | veil | vole | feign* |
| voice | f/v | W | vile | fail | foal | vain |
| voice | k/g | N | court | golf | gear | kin |
| voice | k/g | N | goose | kite | cult | gasp |
| voice | k/g | W | coal | gill | gate* | cap |
| voice | k/g | W | goal | kill | Kate* | gap |
| | | | | | | |
| place | p/t | N | taunt | perch | pitch | tusk |
| place | p/t | N | pinch | tent | tomb | pal |
| place | p/t | W | tack | pipe | pole | tear |
| place | p/t | W | pack | type | toll | pear |
| place | k/t | N | tire | cast | kiss | touch |
| place | k/t | N | keep | time | tinge | case |
| place | k/t | W | toast | cook | coil | toad |
| place | k/t | W | coast | took | toil | code |
| place | b/d | N | daft | boost | babe | dive |
| place | b/d | N | burp | dorm | ditch | boot |
| place | b/d | W | doubt | belt | bare | dale |
| place | b/d | W | bout | dealt | dare | bail |
| place | g/d | N | dunce | geese | good | deed |
| place | g/d | N | gold | damp | dead | guise |
| place | g/d | W | dune | game | gone | door |
| place | g/d | W | goon* | dame | dawn | gore |
| place | f/s | N | serve | false | faith | soft |
| place | f/s | N | fond | silk | sock | fog |
| place | f/s | W | soot | feed | fell | sir |
| place | f/s | W | foot | seed* | sell | fur |
| place | m/n | N | nudge | mesh | Max* | nook |
| place | m/n | N | mast | naught | north | mouth |
| place | m/n | W | nail | might | maim | node |
| place | m/n | W | mail | night | name | mode |

# CHAPTER 10
# Frequency Effects and Underspecification in Phonotactics

In the analysis of OCP-Place in Arabic, the stochastic constraint model assumes that the relative frequency of a form is a function of its acceptability. The stochastic constraint model of a linguistic constraint provides a more accurate description of the OCP-Place data than traditional categorical constraints. This model makes the necessary step of incorporating effects of frequency into the phonology, making the phonological formalism more psycholinguistically plausible. One goal of this chapter is to provide additional evidence for frequency effects in phonotactics.

In the previous two chapters, I argued against analyses of phonological speech errors using radical underspecification. Phonotactic constraints have also been provided as evidence against underspecification theory (Yip 1989, Clements 1988, Mester & Itô 1989, see McCarthy & Taub 1993 for a review). In a derivational phonological theory as in *SPE*, a phonotactic constraint like OCP-Place places restrictions on the abstract, underlying form of a morpheme. Thus, the constraint is applicable to forms before redundancy rules and default rules apply (Stanley 1967,Yip 1989). As discussed in chapter 5, Pierrehumbert (1993) used contrastive underspecification to account for differences in OCP-Place effects in Arabic at different places of articulation. Frisch, Broe, & Pierrehumbert (1995) accounted for the same effects without underspecification. Davis (1991) presents a case for radical underspecification of coronal place in English, also based on the OCP-Place constraint. The second goal of this chapter is to demonstrate that Davis's evidence for underspecification of [Coronal] in English is, in fact, a frequency effect. Thus, I conclude there is no evidence for underspecification in the OCP-Place data.

## *10.1 The Role of Frequency in Phonotactics*

There is a growing body of literature which demonstrates that statistical factors are relevant in phonotactics. It is often assumed that a phonotactic constraint is valid, even if it has one or two exceptions (e.g. Fudge 1969, 1987; Clements & Keyser 1983). When McCarthy (1988) originally proposed the OCP-Place constraint, it was put forth as a statistical tendency, rather than a categorical rule, though the difference was not made formally explicit at that time. Pierrehumbert (1994), which I review in more detail below, demonstrated that the vast majority of occurring word-internal triconsonantal clusters in English are those which are most likely to occur based on frequency, and that native speakers are aware of these frequencies. Finally, Kessler & Treiman (1996) and Treiman et al. (1996) show that English has syllable internal statistical tendencies to which native speakers are sensitive. A frequency based approach to phonotactics therefore has a number of advantages over the traditional model, which assumes that statistical tendencies are irrelevant to the phonology.

*10.1.1 Review of Pierrehumbert (1994).*

Pierrehumbert (1994) examines the extent to which constraints on consonant clusters in English syllables can predict constraints on word medial clusters of three or more consonants. There are two aspects of Pierrehumbert's study which I highlight here. First, she found that frequency was the single greatest factor in predicting the presence or absence of a cluster. Second, she found a word medial OCP effect, which prohibits identical consonants from occurring in first and third position in a triconsonantal cluster.

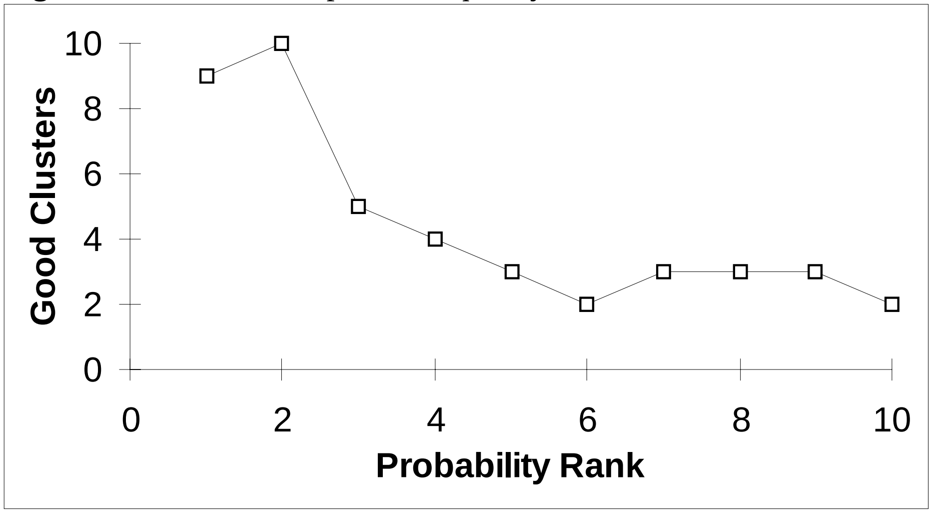
Pierrehumbert (1994) first considered consonant sequences at the beginnings and ends of words, to get an independent estimate of the expected frequency of medial combinations. While possible word onsets are the same as possible syllable onsets, word endings may contain coronal appendices which are not considered part of the syllable. She assumes that the set of word endings with coronal appendices removed represents the set of possible syllable codas (see Pierrehumbert 1994: 170 for details on how she determined which sequences were appendices). There is also a well known constraint against geminates morpheme internally in English. Taking the cross product of the set of word onsets with the set of syllable codas and then removing all instances which create geminates, there are a total of 8708 theoretically possible medial clusters of three or more consonants. In a search of the on-line Collins English dictionary, Pierrehumbert (1994) found that only 50 clusters actually occur morpheme internally.

Pierrehumbert (1994) found that the best predictor of a cluster's occurrence was its expected frequency: 45 of the 50 occurring clusters are among the 200 with highest expected frequency. Figure 10.1, adapted from Pierrehumbert (1994), shows the very strong effect of frequency on medial cluster occurrence. In Figure 10.1, categories on the x-axis are groups of 20 expected clusters, ranked by expected. In other words, clusters were grouped in order of their expected frequency. The first group is the 20 clusters with the highest expected frequency, the second group is the 20 clusters with the highest expected frequency among those remaining, and so on. The y-axis shows the number of clusters found for each group. The rate of occurrence of clusters decreases as expected frequency (represented by the probability ranking) decreases.

Pierrehumbert (1994) proposes five additional constraints to account for the set of 50 actually occurring clusters among the most frequent 200. Among these constraints, she proposes a long distance OCP effect to rule out clusters with identical first and third consonants. The set of clusters in the most frequent 200 which violate this constraint are shown in (105). The OCP as formalized in autosegmental phonology cannot account for long distance identity constraints of this kind (Pierrehumbert 1993, 1994).

| | | | | | | |
|---|---|---|---|---|---|---|
| (105) | /lfl/ | /lkl/ | /lpl/ | /lbl/ | /lgl/ | /lsl/ |
| | /tst/ | /tstr/ | /ntn/ | /ndn/ | /nsn/ | /ksk/ |

**Figure 10.1**: Effect of expected frequency on medial cluster occurrence in English.




Pierrehumbert's results demonstrate that both syllable structure constraints and word structure constraints play a role in English phonotactics. In particular, the strong effect of syllable onset frequency and syllable coda frequency in determining the occurrence of word medial clusters shows that the syllable onset and syllable coda function as units in the phonotactics of English. Pierrehumbert proposes that the syllable grammar is stochastic, so extremely low expected frequency clusters do not have to be ruled out by linguistic constraints. Instead, they are not found solely due to their low expected frequency. At the word level, the presence of an OCP effect between the syllable coda and the following onset shows that some phonotactic constraints apply across syllables.

A closer look at Pierrehumbert's results reveal an effect of very low expected frequency which was not discussed in the paper. The lack of any very low expected frequency combinations is surprising, given that there are a large number of very low expected frequency clusters, some of them would be expected to occur by chance. In other words, if all of the low expected frequency consonants are considered together, the statistically expected number of such clusters is large.[22] In order to demonstrate this fact, I have to make a few assumptions about the data which are not specified in Pierrehumbert (1994).

Recall that 45 of the 50 occurring clusters are found among the 200 possible clusters with highest expected frequency. I call this group the 'high frequency' clusters. The remaining 8508 possible clusters are the 'low frequency' clusters. Pierrehumbert (1994:175) notes that the 40 cluster types with highest expected frequency have expected frequency above 1/2,000. The other 160 high frequency clusters have expected frequency above 1/10,000 (Pierrehumbert 1994:174). For the 200 high frequency clusters then, the mean expected frequency of a cluster is at least 0.00018 = (40/2,000 + 160/10,000)/200. The low frequency clusters have expected frequency no

[22] This fact was originally noticed by Janet Pierrehumbert. My thanks to her for pointing it out to me, as the results in this chapter follow from investigating the same pattern in the Arabic and English OCP-Place data.

greater than 1/10,000. I assume for the sake of discussion that the mean expected frequency of low frequency clusters is 1/100,000. Let us also assume that the dictionary contains approximately 10,000 monomorphemic words. We can determine the aggregate expected number words containing high and low frequency clusters with the following formula:

(106) Expected = Cluster Frequency × Number of clusters × Number of monomorphs

Given this formula, there are 360 expected words containing high frequency clusters (0.00018 × 200 × 10,000). There are 851 expected words containing low frequency clusters (1/100,000 × 8508 × 10,000). Pierrehumbert (1994:177) gives the 15 words containing the low frequency clusters in her dictionary. There are 45 occurring high frequency clusters, so there are at least 45 words containing high frequency clusters in the dictionary. The calculation above is summarized in Table 10.1. The relative rate of occurrence for high frequency clusters (O/E) is at least six times the rate of occurrence for low frequency clusters.[23]

**Table 10.1**: Relative frequency of medial clusters of high and low frequency.

| Expected Frequency | Observed | Cluster Frequency | × | Num clusters | × | Num mono-morphemes | = | Expected | O/E |
|---|---|---|---|---|---|---|---|---|---|
| High | >45 | ~0.00018 | × | 200 | × | ~10,000 | = | 360 | >0.125 |
| Low | 15 | ~0.00001 | × | 8508 | × | ~10,000 | = | 851 | 0.0176 |

Low frequency clusters are found far less often than aggregated expected frequency predicts. This can be accounted for if we assume that very low expected frequencies do not arbitrarily aggregate in the stochastic syllable grammar. For example, if we assume that clusters which are too improbable to appear on their own do not contribute to aggregate expected frequencies, then any cluster with an expected frequency of less than 1/20,000 is not expected to appear in a lexicon of 10,000 monomorphemes (the expected number of words for a cluster of that frequency is less than 0.5, which rounds to 0 expected words since words can only be found in integer increments). In section 10.4, I replicate Pierrehumbert's study by examining the cooccurrence of complex onsets and codas in English syllables.

From a cognitive standpoint, we can interpret this effect as an influence of the set of existing words on the creation of new words and maintenance of the use of old words. Once a high expected frequency form exists, which is very likely to occur due to its expected frequency,

---

[23] This estimate is crucially dependent on the assumed expected frequency of low frequency medial clusters. If this frequency is well below 1/100,000, then the result may not stand. Note, however, that the O/E for high frequency clusters is actually higher than what is shown, since the expected frequency of high frequency clusters was estimated conservatively and the number of words containing high frequency clusters is well above 45 as many of the clusters are found in more than one word.

it becomes an exemplar of an admissible word upon which other forms can be created by analogy. By contrast, very low expected frequency combinations never ‘get their foot in the door’, as even one occurrence of a low expected frequency combination is rare. Overall, then, the actual number of tokens of a high expected frequency combination can become disproportionately high (greater O/E), while lower frequency combinations are less likely to become actually occurring exemplars (lower O/E) when all other factors are equivalent. The overall result is a ‘rich get richer’ effect where high expected frequency combinations have lots of occurring exemplars which reinforce their pattern, and contribute to additional words being created following that pattern. This in turn increases their expected frequency further. Low expected frequency forms have no such reinforcement and are thus less likely to serve as models for new words, reducing their long term expected frequency further.

### *10.1.2 Frequency effects in OCP-Place in Arabic.*

Pierrehumbert (1994) showed that employing a frequency sensitive grammar reduced the set of possible medial clusters in English considerably, simplifying the set of phonotactic constraints. Further analysis of her results found low expected frequency clusters to be disproportionately underrepresented when all low expected frequency combinations are aggregated together. In this section, I demonstrate analogous frequency effects in the verbal roots of Arabic.

The stochastic constraint model of the Arabic verbal roots presented in chapter 7 predicts the acceptability of a triliteral root based on the similarity of the consonant pairs in the root and their expected frequency by random cooccurrence. Model fits were presented for data aggregated by similarity, with extremely good fit ($R^2 > 0.99$). However, the frequency model, which incorporated no OCP-Place effects whatsoever, accounted for the majority of the variation in the data ($R^2 = 0.85$). Like Pierrehumbert’s analysis above, the assumption that the grammar of Arabic triliteral roots is stochastic predicts systematic gaps based on frequency. For example, /Z/ is rarely found in the verbal roots in combination with labial consonants (see Tables 5.1 and 5.2 in chapter 5). The extremely low frequency of /Z/ in Arabic accounts for the low cooccurrence of /Z/ with labial consonants, which are not subject to an OCP-Place restriction. There is no need for an independent constraint against the cooccurrence of /Z/ with a labial.

Within consonant pairs in Arabic that have roughly equivalent similarity, expected frequency disproportionately influences which pairs are actually found. This is exactly parallel to the effect in Pierrehumbert’s (1994) stochastic syllable grammar discussed above. Table 10.2 shows the cooccurrence of adjacent consonant pairs in Arabic, aggregated by similarity. In addition, each aggregate similarity group is divided into a ‘high frequency’ category and a ‘low frequency’ category. The high frequency consonant pairs for each similarity level are the consonant pairs which have expected frequency above the mean expected frequency for that similarity level. The low frequency consonant pairs have less than mean expected frequency. For seven of the nine similarity groups, O/E is less for the low frequency group than the high frequency group, with one tie. This asymmetry is significant on a binomial test ($p = 0.036$). This asymmetry is surprising, given that the O/E measure factors out relative frequency by dividing observed occurrences by expected frequency of cooccurrence.

**Table 10.2**: Frequency effects in the Arabic OCP-Place constraint.

| Similarity | Expected Frequency | Observed | Expected | O/E |
|---|---|---|---|---|
| 0 | Low | 1009 | 773.7 | 1.30 |
| | High | 1969 | 1575.6 | 1.25 |
| 0-0.1 | Low | 129 | 110.1 | 1.17 |
| | High | 322 | 255.1 | 1.26 |
| 0.1-0.2 | Low | 109 | 194.7 | 0.56 |
| | High | 383 | 355.9 | 1.08 |
| 0.2-0.3 | Low | 20 | 78.1 | 0.26 |
| | High | 131 | 182.1 | 0.72 |
| 0.3-0.4 | Low | 5 | 35.9 | 0.14 |
| | High | 24 | 95.3 | 0.25 |
| 0.4-0.5 | Low | 2 | 38.7 | 0.05 |
| | High | 12 | 141.5 | 0.08 |
| 0.5-0.6 | Low | 0 | 6.8 | 0 |
| | High | 3 | 34.1 | 0.09 |
| 0.6-1 | Low | 0 | 37.1 | 0 |
| | High | 0 | 53.1 | 0 |
| 1 | Low | 0 | 48.9 | 0 |
| | High | 1 | 150.7 | 0.01 |

Frequency plays a role in determining exactly which consonant pairs occur among ones which are equally acceptable under the gradient OCP-Place constraint. Recall from above that extremely low expected frequency medial clusters are rarely found in English, even though such clusters are expected to occur in large numbers when considered in aggregate. The difference in O/E between low expected frequency and high expected frequency consonant pairs in Arabic shows the same effect. If extremely low expected frequency pairs do not occur solely based on frequency, and high expected frequency pairs occur disproportionately, regardless of the OCP-Place constraint, then O/E for low expected frequency pairs will be lower than O/E for high frequency pairs. This may be an effect of the existence of exemplars which violate OCP-Place, as discussed above.

## *10.2 Radical Underspecification in OCP-Place in English*

Overall, OCP effects provide evidence against underspecification, as OCP constraints are found, cross-linguistically, to refer to traditionally underspecified features. Davis (1991) presents a case of an OCP-Place constraint in English which appears to treat the coronal place of articulation as underspecified. In this section, I review evidence for the OCP-Place constraint in English and Davis's evidence for underspecification of coronal place. In the next section, I present an alternative account of Davis's data based on the frequency effects shown above. I conclude that frequency, not underspecification, accounts for the exceptional behavior of coronal place (cf. Newman et al. 1996).

### *10.2.1 OCP-Place effects in English.*

Berkley (1994a, b) demonstrates that English has an OCP-Place effect analogous to Arabic. Berkley (1994a) found statistical restrictions between onset and coda consonants in English monomorphemic monosyllables. For example, while words like *palm*, *fib*, *sit*, *dot*, *king*, *skunk*, *leer* and *run* are found, they are statistically underrepresented. Her corpus consisted of the monomorphemic monosyllables of sufficient frequency to appear in the MRC Psycholinguistic Database. Observed and expected frequency of words containing homorganic onset and coda consonant pairs separated by a single segment are shown in Table 10.3 (adapted from Berkley 1994a: 60). Berkley computed expected frequency by considering the expected frequency of each major place of articulation class in onset and coda.

**Table 10.3**: Cooccurrence of homorganic English onsets and codas in monosyllables.

| Class | Observed | Expected | O/E |
|---|---|---|---|
| Labial | 26 | 64.9 | 0.40 |
| Coronal Obstruent | 67 | 91.8 | 0.73 |
| Coronal Sonorant | 94 | 163.0 | 0.58 |
| Dorsal | 10 | 23.3 | 0.42 |

Note that the English constraint appears to be dependent on natural class similarity, much like Arabic. The coronal consonants split into two cooccurrence classes based on manner, which Berkley concludes is a similarity effect. Cooccurrence between the major coronal classes is greater than the cooccurrence within each class (Berkley 1994a). Note that, as in Arabic, the effect of manner on similarity depends on the system of contrasts. Berkley found no evidence for a major split within the small classes of labials or dorsals. The split between the labial obstruents and sonorants in Arabic is quite subtle, as it can be seen only in the non-adjacent consonant pairs (Pierrehumbert 1993).

Berkley (1994b) demonstrates that English OCP-Place effects weaken with distance, parallel to the case of Arabic. Homorganic consonants in onset and coda separated by two segments, for example *vibe* with a diphthong or *cling* with an initial cluster and short vowel, are underrepresented, but less so than consonants separated by a single segment. Homorganic consonant pairs separated by three segments, for example *crusk*, *scald*, *trite*, and *spout*, are only marginally underrepresented. Berkley (1994b) finds no evidence for an OCP-Place effect between onset and coda among the few monomorphemic monosyllables with onset and coda consonants separated by four segments. These distance effects provide additional evidence that the OCP-Place constraint in English is also grounded in similarity (Pierrehumbert 1993, and discussed in chapter 5).

### *10.2.2 Underspecification of coronal in OCP-Place.*

Following Fudge (1969), Clements & Keyser (1983), and Davis (1984), Davis (1991) discusses restrictions on the consonant that flanks the vowel in *sCVC* sequences. He states that there are "virtually no monomorphemic forms in English that have the sequence *sCVC* where the two C's are either both labial or both velar" (Davis 1991: 57). Given the results discussed above, this is not surprising, as Berkley (1994a, b) showed English has a general OCP-Place effect between onset and coda, and Pierrehumbert (1994) has found a constraint against identical first and third consonants in a three consonant sequence. We can account for these patterns with a single OCP-Place constraint which can apply across intervening segments.

The *sCVC* data have received previous attention in the literature as the cooccurrence restriction is nearly categorical. Davis presents *skunk* as the only example which contains homorganic non-coronal C's in an *sCVC* sequence that he found in the on-line Webster's Pocket Dictionary (Nusbaum, Pisoni, & Davis 1984). He presents *spam*, *spumoni*, *spoof*, and *spiffy* as other exceptions which come to mind. Notably, there are no words like **spup* or **skuk* in English with identical non-coronal consonants. Once again, these gaps are easily accounted for by a similarity based OCP-Place constraint. Words like **spup* or **skuk* contain maximally similar consonants, while *spoof* and *spiffy* have moderately similar stops and fricatives, and *spam*, *spumoni*, and *skunk* have dissimilar obstruents and sonorants.

Davis points out that, while **spup* and **skuk* are non-occurring in English, there are plenty of examples similar to **stut*, including *stout*, *state*, *stet*, *stoat*, and *stat*. In addition, there are many words with /st/ in the syllable onset and a non-identical coronal obstruent in the coda. Examples include *stud*, *stood*, *steed*, *stash*, and *stitch*. If coronal place is underspecified, then *stout* does not violate OCP-Place, as the onset and coda /t/ have no place of articulation, and are thus transparent to the operation of the constraint.

Berkley (1994a, b) demonstrated OCP-Place effects between onset and coda in $C_{[Coronal]}VC_{[Coronal]}$ monomorphemic monosyllables. There are two possible accounts of the difference between $stVC_{[Coronal]}$ and $C_{[Coronal]}VC_{[Coronal]}$. First, it might be that [Coronal] is only underspecified in /s/-stop clusters. However, this formal maneuver has no explanatory value. Second, it may be that underspecification is irrelevant in both cases, and something else is special about /st/ onsets. I take the second approach, asserting below that $stVC_{[Coronal]}$ words occur due to their high expected frequency.

There are two other well known cooccurrence restrictions between the onset and coda which do apply to coronal consonants (Fudge 1969, Clements & Keyser 1983, Davis 1984). First, forms like **snun* are not found. However, this constraint generalizes to /m/, and combinations of /n/ with /m/, and /n/ or /m/ in onset with /ŋ/ in coda as well. Forms like **smum*, **snum*, **smun*, **snung*, and **smung* are also absent. Davis (1991) thus assumes that this constraint applies to nasals in general, and does not refer to the underspecified coronal place of /n/. The fact that this generalization appears to apply across place of articulation implies that it is not an OCP-Place effect. I shall have more to say about this case below.

The second constraint restricts liquids in onset and coda, when the onset is a cluster. Thus, it has been claimed that forms like **slul* and **brur* are not found. Since this constraint refers to the class of liquids, Davis again assumes that the constraint does not rely on the presence of the [Coronal] feature. This constraint is not as strict as the nasal constraint, as it has at some marginal exceptions; I found *flail, drear,* and *crore* in the CELEX dictionary. Apparently, this constraint also does not apply as strictly between /l/ and /r/, for example *braille*, *broil*, *blur*, *floor*, *frail*, and *growl*. I claim this constraint can be treated more parsimoniously as special case of OCP-Place which requires reference to the underspecified [Coronal] feature.

To summarize the above discussion, English has OCP-Place effects between homorganic consonants separated by another segment. These effects are found in medial clusters, in *CVC* sequences, and most noticeably in *sCVC* sequences. English OCP-Place effects are apparently based on similarity. Surprisingly, $stVC_{[Coronal]}$ appear to be unaffected, and there is a second constraint, apparently unrelated, against *sNVN* sequences where *N* is any nasal.

### *10.3 Frequency Effects in OCP-Place in English*

In the previous sections I discussed a number of cooccurrence constraints between English onsets and codas. In this section, I unify these constraints by showing that the exceptional status of /t/ in *sCVC* sequences and the apparent constraint against any *sNVN* sequence are due to frequency effects in phonotactics. Thus, the special status of $stVC_{[Coronal]}$ sequences is not evidence for underspecification. Also, a second constraint against nasals in *sCVC* sequences is not required, as such sequences are already expected to be rare by frequency alone.

Berkley (1994a, b) established that there are OCP-Place effects in monomorphemic monosyllables in English. Others (e.g. Sanchez 1990, Lamontagne 1993) have assumed that no such constraint exists. Informally, it is easy to think of a seemingly endless list of counter examples to a long distance OCP: *peep, pip, pep, pap, pop, Pope, poop, pup, bib, babe, Bob, boob, bub,* etc. In fact, identical singleton onset and coda consonants are found in monomorphemic monosyllables at near expected frequency. It is much more difficult to construct such a list for maximally similar, but not identical segments. Broe (1995) discusses a similar pattern of cooccurrence in Ngbaka. In Ngbaka, roots containing repeated identical consonants are common, but roots which contain repeated consonants which are distinguished by a single feature difference are not found. Because of the lack of restriction against identical consonant pairs, which I present in detail below, OCP-Place effects in English were not easy to detect.

The OCP-Place constraint in English has a complicated pattern. I do not analyze all aspects of the constraint in this thesis, but the following facts became apparent to me while investigating Davis's arguments for underspecification of [Coronal] in English. First, as mentioned above, OCP-Place operates in English much as it does in Arabic. Berkely's results point to a similarity effect which weakens with distance. Second, identical consonant pairs separated by a single segment *are* subject to a strong OCP effect. This effect is seen in triconsonantal clusters, in *sCVC* words, and in *(C)lVl* and *(C)rVr* words. Third, there is a weakening of OCP-Place effects for identical singleton onset and coda consonants in *stressed syllables*. This weaking is gradient and is conditioned by the sonority of the segments. Stops are most strongly affected, and thus are able to cooccur as singleton onsets and codas in stressed syllables most frequently. Fricatives and nasals cooccur in stressed syllables slightly less frequently than stops. The liquids are affected least, and have minimal cooccurrence in stressed syllables. This level of cooccurrence is comperable to their cooccurrence in unstressed syllables. Fourth, and finally, cooccurrence in English OCP-Place effects is influenced by frequency much like Arabic. Among onset and coda pairs with equivalent similarity, high frequency pairs occur disproportionately more frequently than expected.

In the remainder of this section, I present a preliminary discussion of the particular details of OCP-Place effects in English discussed above. I first review my data and methods of tabulation. Then I present evidence for OCP-Place effects in English, replicating in part the results of Berkley (1994a, b). Next, I demonstrate the effect of stress on identical singleton onset and coda consonant cooccurence. Finally, I show that frequency accounts for the occurrence of *state*, *stoat*, and so on, and the non-occurrence of **spup* and **skuk*.

### *10.3.1 Materials.*

To generate detailed statistics of the English lexicon, I used the on-line CELEX dictionary. The CELEX dictionary is a large dictionary of British English which contains phonological, morphological, syntactic, and semantic information. The dictionary contains phonemic transcriptions and has been syllabically parsed. In addition, it has been morphologically parsed. Thus, it was possible to automatically extract onset and coda consonant pairs from monomorphemic words for analysis of OCP-Place effects. The CELEX dictionary was altered by me in only one respect. In order to obtain additional data on the high sonority liquids, I inspected the transcriptions of every word with orthographic *r* and added coda /r/ where appropriate according to my own Midwestern American English pronunciation.

### *10.3.2 Tabulation.*

In chapter 7, I showed that word initially, the OCP-Place constraint in Arabic is stronger than it is later in the word. In order to obtain the strongest possible OCP effects, I primarily consider data from monosyllabic words, and from the first syllable of multisyllabic words. Actual occurrences of onset and coda pairings are compared to expected rates of cooccurrence computed as if onsets and codas combined to form syllables at random, based on the actual frequency of onsets and codas in the first syllable of monomorphemic words. Data for unstressed syllable cooccurrence were generated by examining unstressed syllables among the first three syllables of monomorphemic words.

### *10.3.3 Methods 1: OCP-Place effects in English.*

Support for Berkley's results are found by examining maximally similar (but not identical) singleton onset and coda pairs in stressed syllables among the first three syllables of monomorphemic words in the CELEX dictionary. For this table, I consider consonants to be maximally similar if they differ by a single major feature. Thus, p/b, t/s, g/ŋ are maximally similar, p/v, d/s, and k/ŋ are not. The coronal sonorants {l, r, n} were all considered to be maximally similar. Table 10.4 shows aggregated observed, expected, and O/E in the major OCP-Place classes.

**Table 10.4**: OCP-Place effects among maximally similar consonants in stressed syllables.

| Syllable 1 | Observed | Expected | O/E |
|---|---|---|---|
| Labial | 16 | 71.8 | 0.22 |
| Coronal Obstruent | 42 | 86.1 | 0.49 |
| Coronal Sonorant | 70 | 151.8 | 0.46 |
| Dorsal | 10 | 38.2 | 0.26 |
| Syllable 2 | Observed | Expected | O/E |
| Labial | 5 | 19.1 | 0.26 |
| Coronal Obstruent | 65 | 88.9 | 0.73 |
| Coronal Sonorant | 100 | 159.1 | 0.63 |
| Velar | 4 | 11.2 | 0.36 |
| Syllable 3 | Observed | Expected | O/E |
| Labial | 3 | 3.1 | 0.98 |
| Coronal Obstruent | 20 | 26.7 | 0.75 |
| Coronal Sonorant | 32 | 47.9 | 0.67 |
| Velar | 1 | 0.9 | 1.08 |

Berkley's results are clearly replicated: there is an OCP-Place effect in English. In addition, it provides corroborating evidence for the word onset effect on OCP-Place in Arabic. In English, the O/E in syllable one is much less than the O/E in syllable three. OCP-Place effects in English weaken as the word is processed from left-to-right.

*10.3.4 Methods II: Effects of stress on the OCP.*

Identical consonant pairs behave differently in English than they do in Arabic. In Arabic, identical consonant pairs are maximally restricted. In English, identical singleton onsets and codas in stressed syllables are found more frequently than the minimally distinct consonant pairs examined above. Half of the data in this section are monomorphemic words with identical onset and coda consonants in initial stressed syllables. These syllables had the strongest effect for minimally distinct onsets and codas shown above. The other half of the data are identical onset and coda consonants in unstressed syllables, in the first three syllables of monomorphemic words.

Table 10.5 shows identical consonant cooccurrence for word initial stressed syllables. The consonants are grouped by sonority classes, as the rate of cooccurrence (O/E) varies significantly among classes. The hypothesis that all consonant pairs cooccur at the same constant rate can be rejected by a chi-square test (cells with expected frequency below two are not included, $\chi^2(11) = 28.6$, $p = 0.002$). A chi-square test within each sonority class cannot reject the hypothesis that consonants within each sonority class have the same rate of relative cooccurrence.

In Table 10.5, the stops have the highest rate of cooccurrence, followed by the nasals and fricatives. The liquids are the only group that appear to have a strong cooccurrence restriction. The rate of cooccurrence of identical stops, fricatives, and nasals in word initial stressed syllables is much higher than the rate of cooccurrence of maximally similar, but not identical, consonants in the word initial stressed syllable in the previous section. Thus, it appears that for all but the high sonority liquids, identical onsets and codas have only a weak cooccurrence restriction in stressed syllables in English.

**Table 10.5**: Identical singleton onset and coda in word initial stressed syllables.

| Stops | | | | Fricatives | | | |
|---|---|---|---|---|---|---|---|
| | Observed | Expected | O/E | | Observed | Expected | O/E |
| /p/ | 11 | 8.8 | 1.25 | /f/ | 2 | 3.0 | 0.66 |
| /b/ | 6 | 3.6 | 1.68 | /v/ | 0 | 1.2 | 0 |
| /t/ | 13 | 13.8 | 0.94 | /s/ | 7 | 9.1 | 0.77 |
| /d/ | 4 | 6.0 | 0.67 | /z/ | 1 | 0.7 | 1.44 |
| /k | 11 | 17.0 | 0.65 | /ʃ/ | 2 | 0.9 | 2.19 |
| /g/ | 3 | 2.3 | 1.32 | /ʒ/ | 0 | 0.0 | 0 |
| Total | 48 | 51.3 | 0.93 | Total | 12 | 15.0 | 0.80 |
| Nasals | | | | Liquids | | | |
| | Observed | Expected | O/E | | Observed | Expected | O/E |
| /m/ | 15 | 15.2 | 0.99 | /l/ | 4 | 16.3 | 0.25 |
| /n/ | 9 | 16.0 | 0.56 | /r/ | 4 | 24.6 | 0.16 |
| Total | 24 | 31.2 | 0.77 | Total | 8 | 40.9 | 0.20 |

By contrast, identical singleton onsets and codas in unstressed syllables have a strong cooccurrence restriction. Table 10.6 shows the relative cooccurrence of identical onsets and codas for each sonority class in the unstressed syllables. Unlike the stressed syllables, O/E is low for all sonority classes, and there is not a great deal of variation between classes. The relative rate of cooccurrence here is slightly stronger than what was found between maximally similar onsets and coda in stressed word initial syllables in the previous section.

**Table 10.6**: Cooccurrence by sonority classes in unstressed syllables.

| Class | Observed | Expected | O/E |
|---|---|---|---|
| Stop | 10 | 31.4 | 0.32 |
| Fricative | 5 | 11.0 | 0.45 |
| Nasal | 10 | 30.3 | 0.33 |
| Liquid | 5 | 33.3 | 0.15 |
| Total | 30 | 105.9 | 0.28 |

The interaction of stress and sonority with OCP-Place effects provides more evidence for the gradient nature of the OCP-Place constraint. The existence of this interaction also suggests a possible account. There is evidence that, under stress, segments are hyper-articulated (de Jong 1995; cf. Linblom 1983, 1990). This hyperarticulation enhances differences between segments in onset and coda position. For example, stops in particular are acoustically very different in onset position versus coda position in stressed syllables. Stressed onset stops in English have characteristic aspiration which is not found in unstressed syllables. In addition, the abrupt amplitude increase in the release of a stresed onset stop is perceptually highly salient compared to the amplitude decrease into a stop closure post-vocalically (see Silverman 1995 for reference and discussion). For an unstressed onset, there are less dramatic differences between onset and coda position. Thus, I propose that the perceived similarity of onset and coda consonants, particularly those of low sonority, is reduced by the positionally dependent allophonic variation in stressed syllables. Given this suggestive account, an articulatory and acoustic analysis of the differences between consonants of different degrees of sonority in onset and coda position of stressed and unstressed syllables is needed to verify or disprove the hypothesis. Such an analysis is beyond the scope of this thesis, so the account is left here as merely suggestive. Additional evidence for this account is discussed in the next section.

#### *10.3.5 Methods III: Effects of frequency on OCP-Place in clusters.*

The existence of the *sCVC* constraint, the studies of Berkley (1994a, b), and Pierrehumbert's (1994) constraint against identical first and third consonants in a medial cluster all suggest that OCP-Place applies to consonants within a cluster in the onset or coda as well as to singleton onset and coda consonants. In this section, I analyze the cooccurrence of identical consonants separated by a vowel in clusters as well as singletons.

The data for the singletons are the same as used above. The data for clusters come from the stressed syllable in the first three syllables of monomorphemic words. The data for clusters in unstressed syllables is extremely sparse. Apparently, stress aids in licensing complex syllable onsets and codas in English. Expected rates of cooccurrence for clusters were computed by treating the cluster as a single unit, with a single frequency, in other words just like the cluster was a singleton.

Data for identical consonant cooccurrence in stressed syllables, combined with the data from above for singletons in stressed and unstressed syllables, is presented in Table 10.7. In this table, singleton stops, fricatives, and nasals in onset and coda of stressed syllables are aggregated together into the 'stressed consonantal single-single' category. Singletons in unstressed syllables are aggregated together with stressed liquids into the 'other single-single' category. Clusters are aggregated together by position. Cluster onsets with identical singleton codas, for example *flail* and *state*, are the 'cluster-single' category. Singleton onsets with identical consonants in the coda cluster, for example *lilt*, *nonce*, and *coax*, are the 'single-cluster' category. Finally words with both onset and coda cluster with identical consonants, for example **spapt*, **trark*, and **shrarp*, are not found, but are listed with their expected frequency in the 'cluster-cluster' category.

**Table 10.7**: Cooccurrence of identical consonants in onset and coda separated by a vowel in English monomorphemes.

| onset-coda type | actual | expected | O/E |
|---|---|---|---|
| stressed consonantal single-single | 84 | 97.6 | 0.86 |
| other single-single | 38 | 146.8 | 0.26 |
| cluster-single | 12 | 100.1 | 0.12 |
| single-cluster | 18 | 57.1 | 0.32 |
| cluster-cluster | 0 | 46.7 | 0 |

The rate of cooccurrence between identical consonants in the 'cluster-single' and 'single-cluster' categories are reasonably comperable to the rate of cooccurrence in the 'other single-single' case. There is a clear difference between these cases and the 'cluster-cluster' category. There are no instances of identical consonants in onset and coda clusters separated by a vowel. However, the expected frequency of any individual 'cluster-cluster' combination is extremely low. The highest expected frequency in this category is 0.67, for words like **krart*. The mean expected frequency is 0.09 words per combination. Thus, the non-occurrence of any of these clusters need not be attributed specifically to OCP-Place, but instead follows from the results of Pierrehumbert (1994). Extremely low frequency combinations do not occur, even though some are expected to occur when combinations are considered in aggregate.

While the 'single-cluster' and 'cluster-single' cases do not have cooccurrence rates which are very different from the 'other single-single' case, the asymmetry in cooccurrence rates between these two cases is statistically significant ($\chi^2(1) = 7.26$, $p = 0.008$). The 'cluster-single' case is much more restricted than the 'single-cluster' case. This asymmetry can be accounted for in the same manner as the difference in cooccurrence rates between consonantal singletons in stressed and unstressed syllables. As part of an onset cluster, for example the /t/ in *state*, acoustic differences between onset and coda allophones are reduced compared to the stressed singleton onset case. Once again, I hypothesize that perceived similarity is higher between identical onset and coda when the onset is part of a cluster, due to the loss of allophonic contrasts, than it is in the stressed singleton case.

One might then suspect that the rate of cooccurrence in the ‘single-cluster’ case would be higher that it is, comperable to the ‘stressed consonantal single-single’ case. However, there is another factor to consider. Expected frequency also affects the strength of the OCP-Place constraint. Words in the class ‘cluster-single’ have a much higher expected frequency than words in the class ‘single-cluster’. The mean expected frequency for ‘single-cluster’ combinations is low, 0.53 words per cluster. The maximum expected frequency is 3.94, for *sVst*, which actually occurs in the word *cyst*. By contrast, the mean expected frequency for ‘cluster-single’ combinations is 3.16 words per cluster, with maximum expected frequency of 6.88, for words of the form *stVt*.

There are two interacting factors at work in the cooccurrence of identical consonants in clusters. Onset clusters reduce allophonic variation, and thus increase perceived similarity, which decreases cooccurrence (O/E). Clusters are of generally low frequency, which also decreases cooccurrence. Thus, ‘cluster-single’ cases, which are otherwise comperable to ‘other single-single’ cases, are disproportionately underrepresented due to somewhat low onset cluster frequency. The ‘single-cluster’ cases, which are otherwise comperable to ‘stressed consonantal single-single’ cases, have very low coda cluster frequency, and are even more strongly underrepresented. Finally, the ‘cluster-cluster’ cases are affected by of these influences simultaneously. They have such low frequency onsets and codas that even the most probable combination, **krart*, is only marginally likely to occur (0.67 expected). In addition, these combinations have allophonically similar consonant pairs, which subjects them to a strong OCP-Place constraint. As a result, no combinations of this type are found.

Note that, in the discussion above, I mentioned that the most likely combination in the ‘cluster-single’ class is *stVt*. I am now in a position to offer an alternative to Davis’s underspecification account of the special status of $stVC_{[Coronal]}$. The extremely high expected frequency of *stVt* makes it the most likely violation of OCP-Place to be found among *sCVC* words. The CELEX dictionary contains five examples: *state, stet, Stetson, stoat,* and *stout*.

Recall from the previous section that the OCP-Place constraint is affected by frequency. For example, consider the cooccurrence of adjacent consonant pairs like /s-d/, /z-d/, /S-d/, /Z-d/, and /ʃ-d/ in Arabic. There are 13 roots which begin with *sd*, 12 roots that begin with *Sd*, and 9 roots that begin with *ʃd*. No roots begin with *zd* or *Zd*. While {s, S, ʃ} are less similar to /d/ than {z, Z}, the difference in similarity is not so great as to predict this much variation in cooccurrence. It is also the case that {s, S, ʃ} are higher in frequency than {z, Z}. Thus, based on the combined factors of higher frequency and lower similarity, at least one /s-d/, /S-d/, and /ʃ-d/ consonant pair is likely to occur. Once one combination occurs, it is an exemplar of a valid consonant combination that can be used in creating novel roots.

This argument generalizes to all data of the form $stVC_{[Coronal]}$. The onset cluster /st/ is the most frequent onset cluster in the CELEX dictionary. In general, coronal coda consonants are also of high frequency. Labial and dorsal consonants occur with much lower frequency in coda position. The relatively large number of words like *stud, steed, stash, stain,* and *stare* are due to the high expected frequency of the /st/ onset and coronal coda combinations. The low frequency of $spVC_{[Labial]}$ and $skVC_{[Dorsal]}$ words is a result of the relatively low expected frequency of the combination. OCP-Place effects have a greater impact on the low frequency forms than the high frequency forms.

Finally, let me address the existence of a **sNVN* constraint in English. Table 10.8 shows relative cooccurrence for nasals in the the stressed first syllable of monomorphemes in the CELEX dictionary. The left side of Table 10.8 shows s*NVN* combinations, and the right side shows comperable *NVN* combinations. Rows for combinations which are subject to the OCP-Place constraint are shaded.

**Table 10.8**: Cooccurrence of *(s)NVN* in English monomorphemes.

| sNVN | Observed | Expected | O/E | NVN | Observed | Expected | O/E |
|---|---|---|---|---|---|---|---|
| smVm | 0 | 1.3 | 0 | mVm | 15 | 15.2 | 0.99 |
| smVn | 0 | 2.1 | 0 | mVn | 27 | 20.1 | 1.34 |
| smVŋ | 0 | 0.7 | 0 | mVŋ | 9 | 6.3 | 1.43 |
| snVm | 0 | 2.4 | 0 | nVm | 6 | 6.6 | 0.91 |
| snVn | 0 | 3.7 | 0 | nVn | 9 | 16.0 | 0.57 |
| snVŋ | 0 | 1.2 | 0 | nVŋ | 1 | 3.2 | 0.31 |

There is relatively high cooccurrence among non-homorganic pairs in stresses *NVN* syllables on the right side of Table 10.8. The lowest O/E is for *nVŋ* which has the lowest expected frequency. Among *sNVN* forms on the left side of the table, all are of low expected frequency. The form *snVn*, which has highest expected frequency, is presumably subject to the strong OCP-Place constraint for ‘cluster-single’ consonant combinations discussed above. Thus, finding no occurrences of this form is not surprising.

Whether or not we adopt a **sNVN* constraint depends on whether the zero occurrences of the other *sNVN* forms is significant or due to random variation in the data. The highest expected frequency ‘cluster-single’ combination which is non-occurring in the CELEX dictionary is *grVr*, which has expected frequency 6.17. However, this form violates OCP-Place. Among the 26 non-occurring combinations with higher expected frequency than *snVm* or *smVn*, only 4 cannot be attributed to some degree of an OCP-Place effect. OCP-Place once again appears to have a strong effect on the cooccurrence of onset and coda consonants in English, eliminating a number of otherwise possible clusters. The clusters which do not violate OCP-Place which are more likely to occur than *smVn* or *snVm* are given in (107).

(107) tVv swVk tʃVŋ bVrt

Other combinations which do not occur, and contain no OCP-Place violations, but are more frequent than *snVŋ* are:

(108) gVtʃ vVŋ dʒVf krVf spVg ʃVb pVld pVsk nVmp kwVs

There are a number of consonant pairs which are as frequent as *snVm* and *smVn* which also do not occur. These other pairs have no obvious pattern which suggests a parsimonious constraint, or a connection to **sNVN*. Since there is also no indication of a general *NVN* cooccurrence restriction, I conclude that the evidence for a **sNVN* constraint is not sufficiently compelling to consider the gap anything other than a frequency effect.

### *10.4 Frequency in English Syllable Phonotactics*

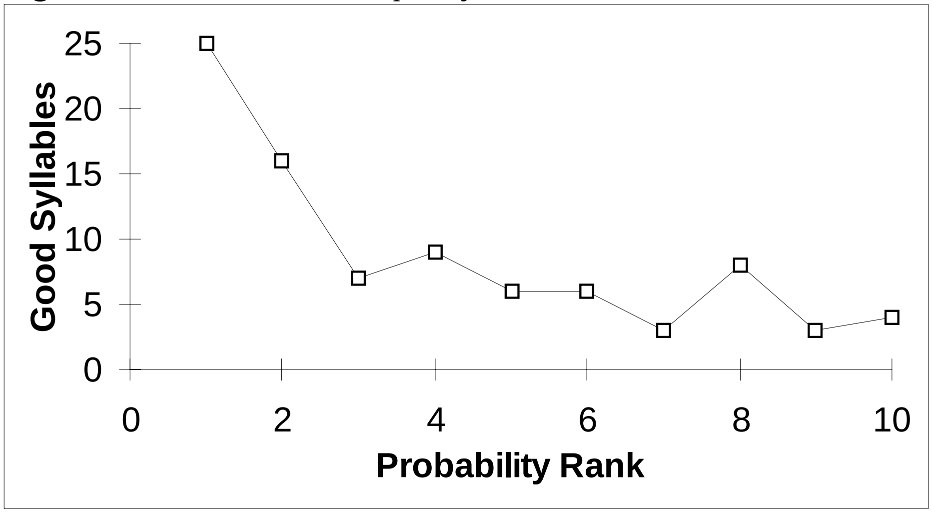
In this section, I present a novel analysis of English syllable structure which is intended as a replication of the results for the medial cluster study reviewed in section 10.1.1 (Pierrehumbert 1994). Pierrehumbert (1994) found frequency to be the primary predictor of medial consonant cluster occurrence, and proposed that the syllable grammar is stochastic. In the analysis of OCP-Place effects in clusters above, I showed that 'cluster-cluster' combinations which violate OCP-Place, like **spapt*, **trark*, and **krart* are not found. The low frequency of such 'cluster-cluster' pairs suggests that the analysis of complex syllables in English monomorphemes may provide additional evidence that the syllable grammar is stochastic.

I examined all word initial stressed syllables of the form *(s)*CCVCC, in other words with both onset and coda clusters. In order to examine syllable structure and not word structure, I removed coronal appendices found in word final position. Appendices were removed if there were no other non-coronal clusters which followed an analagous pattern in the dictionary. For example, post-vocalic /ld/ and /ls/ clusters were assumed to be codas, with no appendices, given the existence of post-vocalic /lb/ and /lf/ clusters. However, post-vocalic /n/ could only be followed by coronals, so these coronals were considered to be appendices and were not included. There are a total of 37 unique onset clusters and 37 unique coda clusters in stressed word initial syllables in monomorphemes in the CELEX on-line dictionary, so there are 1369 possible unique complex syllables.

Among the 1369 possible complex syllables, only 100 actually occur, in 137 unique words. As was found in Pierrehumbert (1994), the vast majority of these complex syllables were among those with highest expected frequency. The 400 most frequent complex syllables contain 90% of the occurring words (87% of the occurring cluster combinations). Pierrehumbert (1994) found that the most frequent 200 medial combinations out a possible 8708 combinations contained 90% of the occurring clusters. The frequency effect between onset and coda, which are separated by a vowel, is weaker than the frequency effect in a medial cluster, where no vowel intervenes.

Figure 10.2 graphically displays the effect of frequency on onset and coda cluster occurrence, analagous to the figure from Pierrehumbert (1994) presented above. In this figure, possible syllables are aggregated in frequency order in groups of 40. Probability rank one is the 40 most frequent syllable types, probability rank two is the next 40 most frequent, and so on. The y-axis marks the number of occurring syllables types for each group.

**Figure 10.2**: The effect of frequency on cooccurrence of onset and coda clusters.




There was some evidence in Pierrehumbert’s data that very low expected frequency clusters were disproportionately underrepresented, an effect also found in the Arabic OCP-Place effects and English OCP-Place effects examined above. Low expected frequency combinations of clusters in the onset and coda of stressed syllables in monomorphemes are also disproportionately underrepresented. Table 10.9 presents O/E for complex syllables, divided into high and low expected frequency groups. Combinations among the 400 with highest expected frequency are considered ‘high frequency’, and the remainder are considered ‘low frequency’. This table replicates the tentative computation based on Pierrehumbert (1994) for medial clusters, implying that the above result was not an artifact of the assumptions used in making that computation.

**Table 10.9**: Relative frequency of complex syllables of high and low expected frequency.

| Expected Frequency | Observed | Expected | O/E |
|---|---|---|---|
| High | 124 | 147.5 | 0.84 |
| Low | 13 | 32.2 | 0.40 |

In summary, I conclude that there is additional evidence that the syllable grammar in English is stochastic, based on the predictive power of frequency on the occurrence of complex syllables which have onset and coda clusters. We have also seen that the occurrence of verbal roots in Arabic can, for the most part, be predicted by frequency. The addition of the OCP-Place constraint to Arabic and to English accounts for the underrepresentation and non-occurrence of a large number of combinations with moderate expected frequency ($R^2 > 0.99$ in chapter 7). For example, these two factors taken together account for nearly all of the variation in the Arabic data. Finally, there is a disproportionate underrepresentation of low expected frequency combinations and overrepresentation of high expected frequency combinations in English complex syllables, medial clusters, and Arabic verbal roots. I claim that this is due to the status of existing words as exemplars which act as word templates. High expected frequency combinations are likely to occur, and once they do they serve as models of acceptable words. Additional phonotactically acceptable words can be created along the same pattern, disproportionately increasing the frequency of high expected frequency combinations. Low expected frequency combinations are less likely to have that critical first exemplar, and thus have a disproportionately reduced level of occurrence.

### *10.5 Conclusion: Frequency and Underspecification*

In this chapter, I have shown that frequency plays a major role in the phonotactics of English and Arabic. Following Pierrehumbert (1994), I account for the effects of frequency by assuming that the grammar is frequency sensitive, and combinations of forms occur in proportion to their expected frequency. Additional constraints, like OCP-Place, influence the relative cooccurrence of forms. The stochastic constraint model of Frisch, Broe, & Pierrehumbert (1995) complements the stochastic syllable grammar, as the stochastic constraint model makes the connection between a linguistic parameter relevant to the constraint and the relative frequency of a form as predicted by the stochastic syllable grammar. Forms which violate the stochastic constraint are disproportionately underrepresented, and may be non-occurring if their predicted rate of occurrence is very low.

The effect of frequency on cooccurrence accounts for the seemingly anomolous presence of a large number of $stVC_{[Coronal]}$ forms in English. The frequency account provides an alternative to an underspecification analysis. Coronals, and /t/ in particular, are special due to their high frequency, rather than an abstract formal property like underspecification. A similar conclusion was reached in the previous chapter, where neighborhood effects, which are linked to frequency, provided an alternative to an account of so-called anti-frequency effects in speech production.

There is also evidence for the special status of high frequency coronals in speech perception. Newman et al. (1996) found that high frequency coronals did not display the same sensitivity to lexical neighborhood effects that lower frequency consonants did in a phoneme identification task. The sensitivity to neighborhood effects was not due to underspecification, as low frequency coronals behaved like other low frequency consonants. Rather, high frequency /t/ and /s/ were the only phonemes examined which were not influenced by neighborhood effects.

In general, underspecified features are of higher frequency than specified features so that frequency and underspecification are correlated properties (Stemberger & Stoel-Gammon 1991). However, the predictions of underspecification are categorical, while frequency is a gradient and context dependent property. The results of this thesis show that frequency provides a more appropriate account of some effects of underspecification. In combination with the formal difficulties inherent in the use of underspecification (Broe 1993), and the move toward constraint based phonological theories, alternative accounts to underspecification phenomena are necessary.

Keating (1991) suggests that the variety of possible coronal articulations contributes to their cross-linguistically high frequency. In addition, the corner of the alveolar ridge provides a definitive landmark for coronal place of articulation targets, suggesting that coronal articulation is articulatorily more robust. Such factors may also contribute to the high frequency of coronals within a single language. If analagous factors exist for voicelessness of obstruents and stop manner of articulation, then the frequency of these features is also expected to be higher. Since frequency provides one alternative to underspecification, factors which increase the frequency of these features contribute to an alternative account for underspecification phenomena.

The three features [alveolar], [stop], and [voiceless] are each a particular contrast in a larger range. Articulatorily, these features may all have advantages over the other features in their respective contrasts. Among the places of articulation, there are additional anatomical reasons why alveolars may be preferred (Zemlin 1988). First, alveolar articulations are made with the tongue tip, which is perhaps the best articulator in the vocal tract. The forward two thirds of the tongue is innervated by the 12th cranial nerve (the hypoglossal), and is the only articulator which doesn’t share motor innervation with another articulator. The tongue has tactile sensation all along its forward surface and also has proprioceptive feedback of overall position and stretch. Second, the alveolar ridge has a particularly well developed membrane that presents a series of wrinkles (the palatal rugae) as a landing site. Thus, the amount of feedback in an alveolar articulation is perhaps greater than for any other articulation in the vocal tract.

Analogously, a stop articulation may have articulatory advantages over a continuant. A stop closure is the discontinuity at the end of a continuous range of approximation. Thus, it is impossible to ‘over articulate’ a stop, as a greater than the intended amount of closure still results in a stop. In addition, a closure provides tactile feedback. Achieving frication is a relatively more difficult task, requiring a constriction narrow enough to cause turbulence, but short of a stop (Stevens 1972; Mrayati, Carre, & Guerin 1988). The same argument holds for the voicing contrast. Voicelessness is relatively simple, and can be achieved by strong glottal abduction or adduction (Pierrehumbert 1995), both endpoints on a continuum of possible articulations (cf. Sawashima & Hirose 1983). Voicing requires close approximation of the vocal folds, so that the cycle of pressure build up and elastic recoil is achieved.

In each case, articulation involving the underspecified feature is in some sense simpler or more stable, and hence more robust. It is possible that the reduced demands on the production system are in part responsible for the relatively reduced error rate within the unmarked categories found by Stemberger (1991b) for underspecified consonants. These functional factors influence language and may become grammaticized, resulting in the patterns of defaults and transparency found in some languages which have been taken as evidence in support of underspecification. Finally, these factors influence the cross-linguistic and within language robustness of segments, and thus the frequency with which they are encountered.

# CHAPTER 11
# Discussion and Conclusion

In this chapter, I compare the phonological frameworks of Optimality Theory and Declarative Phonology along with the conception of the phonology as an emergent property from a connectionist network. The linguistic frameworks each have some insight into the similarity effects presented in this thesis, but no linguistic framework yet proposed can capture all of the results presented here in a natural way. Assuming an underlying connectionist formalism provides a psychologically plausible model of phonological theory which can accommodate the gradient effects of frequency and similarity presented in this thesis (Dell 1996).

## *11.1 Summary of Results*

In chapter 2, I introduced structured specification (Broe 1993), which is a representation of the segment inventory that explicitly represents natural classes and redundancy relations between natural classes. In chapter 3, structured specification was incorporated into a metric of segmental similarity which, as a result, is sensitive to contrast and redundancy relationships among features. In chapter 4, I tested the similarity metric on two corpora of naturally occurring consonantal segment errors. The natural classes similarity model was shown to be superior to previously used measures of segmental similarity in predicting consonantal error rate. This similarity metric, in combination with the stochastic constraint model of a gradient linguistic constraint (Frisch, Broe, & Pierrehumbert 1995) was applied to the similarity based OCP-Place effect in Arabic (Pierrehumbert 1993). The natural classes similarity model, in combination with the stochastic constraint model, provided superior predictions to the original Pierrehumbert (1993) similarity model and to the autosegmental formulation of the OCP-Place constraint (McCarthy 1994).

In chapters 6 and 7, I showed that the word onset position was particularly sensitive to similarity, in both speech errors and the OCP-Place constraint. In modeling the effect of word position in OCP-Place, I implemented a model of stochastic constraint combination. The model of gradient constraint combination used the product of stochastic constraints. It is a fuzzy logic version of simultaneous categorical constraint satisfaction.

In chapters 8 and 9, I presented alternative accounts of effects attributed to underspecification in language production (Stemberger 1991a, b). I demonstrated that the similarity metric based on structured specification gives a superior prediction of speech error rates in comparison to a similarity metric based on underspecified representations. In addition, I showed that asymmetries in speech error rates attributed to underspecification (Stemberger 1991a) are dependent on the status of a speech error outcome as a word or non-word, and thus are not effects of underspecification. I propose instead that neighborhood density effects, which are correlated with phoneme frequency and hence underspecification, provide an alternative analysis of the asymmetries found.

In chapter 10, I discuss the influence of frequency on phonotactics. Frequency is a strong predictor of occurrence of medial clusters in English (Pierrehumbert 1994), and of cooccurrence of clusters in syllable onset and coda in English. In addition, frequency effects provide the foundation for the stochastic constraint analysis of OCP-Place effects in Arabic. Frequency accounts for a large portion of the variation in the Arabic data. Frequency also plays a role in OCP-Place effects in English. Violations of OCP-Place are especially frequent among coronals. However, coronals combinations have high expected frequency, and are thus more likely to occur overall. Finally, I showed that low expected frequency combinations are disproportionately underrepresented in medial clusters in English, in syllables with both onset clusters and coda clusters, and among consonants which are predicted to be equally affected by violations of OCP-Place in both English and Arabic. Conversely, high expected frequency combinations are disproportionately found using the O/E measure, which factors out expected frequency. I account for the difference between high and low expected frequency combinations by considering the effect that the existence of an actual word has on the grammar. When a word of a particular form exists, it serves as an exemplar of a phonotactically acceptable word. Other words following the same pattern as the exemplar are acceptable as well, so productive use of existing forms will overrepresent occurring combinations. Since high expected frequency items are more likely to occur than low expected frequency items, high expected frequency items are more likely benefit from the existence of exemplars.

## *11.2 Similarity, Frequency, and Phonotactic Constraints in Phonology*

The use of similarity, frequency, and the logistic function as a model of a gradient linguistic constraint is a non-traditional approach to phonotactics. Traditional phonotactic constraints are categorical statements of well-formedness. The stochastic constraint model takes a continuous variable as input and produces a continuous acceptability value as output, and thus can model gradient effects. The stochastic constraint model can also be parameterized to create categorical effects. The use of the stochastic constraint is a necessary extension of formal linguistic theory, which cannot model the frequency and similarity effects presented in this thesis.

The traditional conception of a categorical phonotactic constraint represents the extreme on a much broader field of gradient phonotactics. Violations of the OCP-Place constraint do not disqualify a form absolutely. Instead, violations by different forms occur to different degrees, and forms which violate the constraint more seriously are found less frequently than forms which are minor violations. The fundamental dimension in determining the degree of violation is similarity. In combination with expected frequency, actual occurrence of forms can be predicted.

A phonotactic constraint can thus be seen as a function which takes as input a word form, and produces as output a degree of acceptability. The domain of the constraint is the set of actual, possible, and impossible word forms. The range of the function is acceptability, which is reflected in frequency in the lexicon. Thus, the implicit knowledge of the native speaker in this model of phonotactics includes the knowledge of the set of actual words and the frequency of those words. The constraint is an abstraction or generalization of this knowledge which can be applied to other cases. This model of phonotactics is no different from a model of categorization, in which the category is the set of ‘good words’ (Frisch, Broe, & Pierrehumbert 1995). The stochastic constraint model seems exceptionally appropriate as the logistic function upon which it is based can be used as a model of the ogival curves characteristic of categorical perception (e.g. Kuhl & Miller 1978).

These results provide indirect evidence against a cognitive distinction between a linguistic object and a linguistic constraint. Identifying forms which violate a constraint is parallel to the categorical perception of linguistic objects. The additional effects of existing exemplars on occurrence and cooccurrence provide additional evidence that linguistic constraints are abstractions over the set of lexical items (Pierrehumbert 1994, Beckman & Edwards 1996). No current formal phonological theory appropriately models this conception of grammar. The theory which comes closest to this position is Declarative Phonology (Scobbie 1993). In this framework, constraints are underspecified objects (or conversely, an object is a fully specified constraint). Simple abstractions are combined via the formal mechanism of unification to create complete complex objects.

The evidence presented in this thesis suggests that native speakers have specific linguistic knowledge about both the individual lexical items of their language and the generalizations over those items, the constraints. Recall that in structured specification (Broe 1993), the phoneme inventory is represented as a multidimensional hierarchical structure. The phonemes themselves appear at the lowest level of the hierarchy. Successively more general natural classes, which contain the phonemes, are also explicitly represented. Thus, both the individual members of the classes, as well as the generalizations, are explicitly learned and represented. A phonology which is built along the same model, which thus encodes a great deal of redundant information, allows a great deal of room for individual, dialectal, and cross-linguistic variation in the representation of phonological patterns. Over and under generalization are natural phenomena where the domain of a phonological process is mis-generalized. The lattice model of phonology also admits explanatory cognitive factors for cross-linguistic tendencies. For example, most natural kind classes have a privileged, or so-called ‘basic level’ which is acquired at an early age (Rosch 1975). Information about categories at the basic level is also easiest to access. Basic level categories are at a moderate level of abstraction (e.g. ‘cat’), between very general superordinate categories (e.g. ‘animal’) and very specific subordinate categories (e.g. ‘Manx’). The hierarchy of phonological generalizations may also have a basic level in the middle of the range of possible generalizations: at the segmental and the phrasal level phonology. Generalizations at lower levels of phonetic detail, or higher levels of supra-phrasal structure are likely to more difficult to learn and most susceptible to obliteration by individual and contextual variation.

The lattice model of phonology can account for language specific patterns. Universal patterns will also emerge. It is commonly assumed that generalizations found in many languages reflect some form of universal constraint (e.g. Lindblom 1983, 1990). I propose that physical and cognitive universals constrain the space in which language particular phonology can operate.

## *11.3 Connectionism, Optimality, and Constraints*

There are a number of existing formal phonological theories. They vary in the naturalness with which they can incorporate the results of this thesis. I concentrate here on the two current constraint based models of phonology which have been most influential to my work: Optimality Theory and Declarative Phonology. Neither framework can account for the data in this thesis. I suggest instead that a phonology grounded in connectionist formalism provides a psychologically plausible model of phonology which can account for frequency and similarity effects.

Optimality Theory is a notable exception to traditional linguistic theories in that violations of constraints are permitted, when higher ranked constraints intervene. However, Optimality Theory focusses on the mapping from input to output forms. Phonotactic patterns among underlying forms are considered epiphenomenal and result from cases where the ranked hierarchy of constraints produces a many-to-one mapping of underlying forms to surface forms (Pierrehumbert 1996). In this situation, a mechanism of lexicon optimization (Prince & Smolensky 1993) uses the most harmonic input to output mapping to determine the actually occurring input. Thus, there are no patterns among underlying forms which are not based on the constraint hierarchy.

Optimality Theory cannot account for statistical patterns of underrepresentation through constraint ranking in a strict dominance hierarchy (Berkley 1994a, b). Any constraint ranking which eliminates a phonotactic pattern in one form will eliminate that pattern in all equivalent forms.

In addition, Optimality Theory cannot account for cumulative interactions of gradient effects (Pierrehumbert & Nair 1995). In this thesis, the similarity effects of OCP-Place were found to interact with word position, expected frequency, and stress. The degree of violation of OCP-Place depends on the position of the combination with respect to the left edge of the word, syllable stress, and on the expected frequency of the combination.

The effects of high versus low expected frequency, accounted for using occurring words as exemplars of phonotactically acceptable existing words, is especially problematic for Optimality Theory. The existence of underlying forms other than the particular input under consideration is not relevant to the set of Optimality Theoretic constraints. A parallel to the exemplar problem occurs on the output side of the Optimality Theoretic grammar. Paradigm uniformity constraints (Kiparsky 1972, Flemming 1995, Steriade 1996) require reference to the set of existing output forms to determine well-formedness. Specific information about individual inputs and outputs are not part of the formalism and thus these data cannot be captured.

The constraints in Optimality Theory are proposed to be universal and not abstractions over the lexicon. Regular patterns are the domain of Universal Grammar only. Exemplar effects and paradigm conditions thus are doubly difficult to model in this system. Since the lexicon is language particular, it plays no role in the Optimality Theoretic constraint hierarchy. Preliminary experimental research has found evidence that exemplars do influence phonological judgement. Cole, Dell, & Guest (1996) found that subjects who were trained on the stress pattern of nonsense words were equally likely to predict the stress of novel forms by constraint satisfaction (utilizing a deduced generalization about constraint interaction) as by template matching to learned exemplars with the same number of syllables.

While Optimality Theory is overly restrictive in its architecture, a model of phonotactic constraints based on connectionism is relatively under structured (see Dell et al. 1993). Connectionist models are inherently quantitative, as activation levels of the nodes in a connectionist network are typically gradient. Connectionist models are also strongly associative. A connectionist model can, in principle, learn to associate any two nodes in the network. The more frequently two nodes are activated together in the network, the stronger their association becomes. Thus, connectionist models are inherently able to model effects of frequency of cooccurrence (Dell 1996). In chapter 3, I showed that the natural classes similarity model made comparable predictions to an activation network model of similarity. Exemplar effects, like the effect of expected frequency within similarity classes, are a natural part of connectionist models. I also argued in chapters 6 and 7 that effects of interference caused by activation in a network model can account for the extra sensitivity of word onsets to similarity. What is unclear, in a connectionist model, is the status of representations and constraints. Since every level of linguistic structure is an abstraction to some degree, deciding on the fundamental units from which the generalizations are emergent is a challenging problem, and, according to Dell (1996), a crucial one.

In Declarative Phonology there is a natural hierarchy of constraints, as in the lattice based model of phonology sketched above. Constraints are partial descriptions of well-formed structures, and can thus be hierarchically ordered by their specificity (Bird 1995). Constraints are combined by enforcing their requirements simultaneously. In Declarative Phonology, constraints are generally surface true. In cases of constraint conflict, constraints are ranked by the elsewhere condition and the most specific constraint applies. The Declarative Phonology framework must be extended to allow gradient constraint models and gradient constraint combination in order to capture the OCP-Place effects in this thesis.

In Optimality Theory, constraints are universals, and are arbitrarily ranked. In Declarative Phonology, constraints are type abstractions and are ranked only by the elsewhere condition. I believe the true form of the grammar is somewhere in between. It is likely that universal constraints, based on physiology and psychology, influence the form of language particular constraints. In cases of conflicting universal constraints, the conflict may be resolved arbitrarily in favor of one factor or another. This degree of arbitrariness is necessary if cross-linguistic variation is to be permitted in the face of universal constraints. Universal constraints are necessary if we are to ever truly explain patterns which are seen over and over cross-linguistically.

The patterns in the OCP-Place constraint examined in this thesis show how a constraint based on universal cognitive factors such as similarity and frequency can show language particular variation. Differences were found between OCP-Place effects among identical consonants in English and Arabic. In Arabic, identical consonant pairs are rarely found. In stressed syllables in English, they are frequent. This language particular variation is connected to differences between Arabic and English morphology. In the non-concatenative morphology of Arabic, the abstract consonantal root is a morpheme. The consonant sequence can be found in different prosodic positions. In English, however, there is less variation in the realization of the segments of a morpheme in onset and coda position, and less variation in the stress of a particular syllable in a morpheme. These prosodic factors consistently influence the perceived similarity of segments, while the prosodic factors in Arabic are more easily abstracted away from as contextual variation.

### *11.4 Conclusion*

In this thesis, I have presented a case study for a phonology which is grounded in physical and cognitive factors. In order to model the effects of frequency and similarity on the lexicon, I have presented a formal model of a gradient linguistic constraint and gradient constraint combination. I propose that gradient linguistic constraints can be incorporated into a lattice model of the phonology which encodes individual exemplars of occurring words as well as generalizations across exemplars. This is a psycholinguistically plausible model of phonology which has much in common with a connectionist approach to linguistic knowledge. My results suggest that careful examination of the statistical patterns of language can reveal a great deal about the mental representation of implicit linguistic knowledge.

## References


Archangeli, D. (1984). *Underspecification in Yawelmani phonology and morphology*. New York: Garland.

Archangeli, D. (1988). Aspects of underspecification theory. *Phonology* 5: 183-208.

Archangeli, D. & D. Pulleyblank. (1989). Yoruba vowel harmony. *Linguistic Inquiry* 20: 173-217.

Archangeli, D. & D. Pulleyblank (1994). *Grounded phonology*. Cambridge, MA: MIT Press.

Beckman, M. & J. Edwards (1996). Lexical frequency effects on young children's imitative productions. Manuscript submitted for publication. Ohio State University, Columbus, OH.

Bender, M. L. & H. Fulass (1978). *Amharic verb morphology*. East Lansing, MI: Michigan State University African Studies Center.

Berkley, D. (1994a). The OCP and gradient data. *Proceedings of FLSM V. Studies in the Linguistic Sciences*. 24.1/2: 59-72.

Berkley, D. (1994b). Variability in obligatory contour principle effects. In *Papers from the 30th regional meeting of the Chicago Linguistics Society*. Chicago: University of Chicago Press. 1-12.

Bird, S. (1995). *Computational phonology: a constraint-based approach*. Cambridge: Cambridge University Press.

Bod, R. (1995). *Enriching linguistics with statistics: performance models of natural language*. ILLC Dissertation Series 1995-14. Amsterdam: ILLC.

Broe, M. (1992). Introduction to feature geometry. In G. Docherty & D. Ladd (eds.), *Papers in laboratory phonology II: gesture, segment, prosody*. Cambridge: Cambridge University Press. 149-165.

Broe, M. (1993). Specification theory: the treatment of redundancy in generative phonology. Unpublished Ph.D. dissertation, University of Edinburgh.

Broe, M. (1995). Specification theory and Ngbaka co-occurrence constraints. Paper presented at the Mid-Continental Workshop on Phonology. Columbus, OH.

Browman, C. & L. Goldstein (1986). Towards an articulatory phonology. *Phonology Yearbook* 3: 219-252.

Browman, C. & L. Goldstein (1990). Tiers in articulatory phonology, with some implications for casual speech. In J. Kingston & M. Beckman (eds.), *Papers in laboratory phonology: between the grammar and physics of speech*. Cambridge: Cambridge University Press. 341-376.

Buckley, E. (1993). Tigrinya root consonants and the OCP. Paper presented at the 16th annual Penn Linguistics Colloquium. Philadelphia, PA.

Chomsky, N. (1957). *Syntactic structures*. The Hague: Mouton.

Chomsky, N. (1993). A minimalist program for linguistic theory. MIT occasional papers in linguistics.

Chomsky, N. & M. Halle (1968). *The sound pattern of English*. New York: Harper and Row.

Clements, G. N. (1985). The geometry of phonological features. *Phonology Yearbook* 2: 225-252.

Clements, G. N. (1987). Phonological feature representation and the description of intrusive stops. In *Papers from the 23rd regional meeting of the Chicago Linguistic Society*. Chicago: University of Chicago Press. 29-50.

Clements, G. N. (1988). Toward a substantive theory of feature specification. In *Proceedings of NELS 18*. Amherst, MA: GLSA. 79-93.

Clements, G. N. & S. J. Keyser (1983). *CV phonology: a generative theory of the syllable*. Cambridge, MA: MIT Press.

Cole, J., G. Dell, & D. Guest (1996). Testing the transitivity hypothesis in OT. Paper presented at the 2nd Midcontinental Workshop on Phonology. University of Illinois, Urbana, IL.

Cole, J. & L. Trigo (1989). Parasitic harmony. In H. van der Hulst and N. Smith (eds.) *Features, segmental structure and harmony processes*. Dordrecht: Foris. 19-38.

Coleman, J. (1992). The phonetic interpretation of headed phonological structures containing overlapping constituents. *Phonology* 9(1): 1-44.

Cowan, J. (1979). *Hans Wehr: a dictionary of modern written Arabic*. Wiesbaden: Otto Harrasowitz.

Davis, S. (1984). *Topics in syllable geometry*. New York: Garland.

Davis, S. (1991). Coronals and the phonotactics of nonadjacent consonants in English. In C. Paradis & J.-F. Prunet (eds.), *Phonetics and phonology 2: the special status of coronals*. New York: Academic Press. 49-60.

DeJong, K. (1995). The supra-glottal articulation of prominence in English: linguistic stress as localized hyperarticulation. *JASA* 97(1): 491-504.

Dell, G. (1984). Representation of serial order in speech: evidence from the repeated phoneme effect in speech errors. *JEP: Learning, Memory, and Cognition* 10: 222-233.

Dell, G. (1986). A spreading activation theory of retrieval in sentence production. *Psychological Review* 93: 283-321.

Dell, G. (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. (1990). Effects of frequency and vocabulary type on phonological speech errors. *Language and Cognitive Processes* 5(4): 313-349.

Dell, G. (1996). Counting, connectionism, and lexical representation: a commentary. Manuscript submitted for publication. University of Illinois, Urbana, IL.

Dell, G., C. Juliano, & A. Govindjee (1993). Structure and content in language production: a theory of frame constraints in phonological speech errors. *Cognitive Science* 17: 149-195.

Dell, G. & P. Reich (1980). Toward a unified theory of slips of the tongue. In V. Fromkin (ed.), *Errors in linguistic performance: slips of the tongue, ear, pen, and hand*. New York: Academic Press. 273-286.

Dell, G. & P. Reich (1981). Stages in sentence production: an analysis of speech error data. *Journal of Verbal Learning and Verbal Behavior* 19: 624-634.

Devore, J. L. (1987). *Probability and statistics for engineering and the sciences*, 2nd ed. Monterey, CA: Brooks/Cole.

Eriksen, C. & D. Schutze (1978). Temporal factors in visual information processing: a tutorial review. In J. Requin (ed.) *Attention and Performance VII*. Hillsdale, NJ: Earlbaum.

Fant, G. (1973). *Speech sounds and features*. Cambridge, MA: MIT Press.

Flemming, E. (1995). Auditory representations in phonology. Unpublished Ph.D. dissertation, UCLA.

Frisch, S., M. Broe, & J. Pierrehumbert. (1995). The role of similarity in phonotactic constraints. Unpublished manuscript. Northwestern University, Evanston, IL.

Fromkin, V. (1971). The non-anomalous nature of anomalous utterances. *Language* 47(1): 27-52.

Fudge, E. (1969). Syllables. *Journal of Linguistics* 5: 253-287.

Fudge, E. (1987). Branching structure within the syllable. *Journal of Linguistics* 23: 359-377.

Garret, M. (1980). The limits of accomodation: arguments for independent processing levels in sentence production. In V. Fromkin (ed.), *Errors in linguistic performance: slips of the tongue, ear, pen, and hand*. New York: Academic Press.

Gazdar, G., L. Klein, G. Pullum, & I. Sag (1985). *Generalized phrase structure grammar*. Oxford: Blackwell.

Gentner, D. (1989). The mechanisms of analogical learning. In S. Vosniadou & A. Ortony (eds.), *Similarity, analogy, and thought*. New York: Cambridge University Press. 199-241.

Gluck, M. & G. Bower (1988). From conditioning to category learning: an adaptive netword model. *JEP: General* 117:227-247.

Goldinger, S., P. Luce, & D. Pisoni (1989). Priming lexical neighbors of spoken words: effects of competition and inhibition. *Journal of Memory and Language* 28: 501-518.

Goldinger, S., D. Pisoni, & P. Luce (1996). Speech perception and spoken word recognition: research and theory. In N. Lass (ed.), *Principles of experimental phonetics*. St. Louis: Moseby. 277-327.

Goldsmith, J. (1979). *Autosegmental phonology*. New York: Garland.

Goldsmith, J. (1994). *Autosegmental and metrical phonology*. Cambridge: Basil Blackwell.

Goldstone, R. (1994a). Similarity, interactive activation, and mapping. *JEP: Learning, Memory, and Cognition* 20: 3-28.

Goldstone, R. (1994b). The role of similarity in categorization: providing a groundwork. *Cognition* 52: 125-157.

Goldstone, R., D. Gentner, & D. Medin (1989). Relations relating relations. *Proceedings of the 11th annual Conference of the Cognitive Science Society*. Hillsdale, NJ: Earlbaum. 131-138.

Goldstone, R., D. Medin, & D. Gentner (1991). Relational similarity and the nonindependence of features in similarity judgments. *Cognitive Psychology* 23: 222-262.

Goodman, N. (1972). Seven strictures on similarity. In N. Goodman (ed.), *Problems and projects*. New York: Bobbs-Merrill.

Greenberg, J. (1950). The patterning of root morphemes in Semitic. *Word* 6: 162-181.

Halle, M. (1959). *The sound pattern of Russian: a linguistic and acoustical investigation*. The Hague : Mouton.

Halle, M. & K. Stevens (1971). A note on laryngeal features. *QPR* 101: 198-213. Research Laboratory of Electronics, MIT.

Hayes-Roth, B. & F. Hayes-Roth (1977). Concept learning and the recognition and classification of exemplars. *Journal of Verbal Learning and Verbal Behavior* 16: 321-338.

Hayward, K. & R. Hayward. (1989). “Guttural”: arguments for a new distinctive feature. *Transactions of the Philological Society* 87: 179-193.

Itô, J. & R. A. Mester (1986). The phonology of voicing in Japanese. *Linguistic Inquiry* 17: 49-73.
Jakobson, R., G. Fant & M. Halle (1952). *Preliminaries to speech analysis*. Cambridge: MIT Press.
Keating, P. (1984). Phonetic and phonological representation of stop consonant voicing. *Language* 60: 286-319.
Keating, P. (1991). Coronal place of articulation. In C. Paradis & J.-F. Prunet (eds.), *Phonetics and phonology 2: the special status of coronals*. New York: Academic Press. 29-48.
Kelso, J., E. Saltzman, & B. Tuller (1986). The dynamic perspective on speech production: data and theory. *Journal of Phonetics* 14: 171-196.
Kessler, B. & R. Treiman (1996). Syllable structure and phoneme distribution. Manuscript submitted for publication. Stanford University, Stanford, CA.
Kenstowicz, M. (1994). *Phonology in generative grammar*. Cambridge, MA: Blackwell.
Kiparsky, P. (1972). Explanation in phonology. In S. Peters (ed.), *Goals of linguistic theory*. Englewood Cliffs, NJ: Prentice-Hall. 180-227.
Kiparsky, P. (1982). Lexical phonology and morphology. In I. S. Yang (ed.), *Linguistics in the morning calm*, vol. 2. Seoul: Hanshin. 3-91.
Kisseberth, C. (1971). Cyclic rules in Klamath phonology. *Linguistic Inquiry* 3: 3-33.
Koskinen, K. (1964). Kompatibilitat in den dreikonsonantigen hebraishen wurzeln. *Zeitschrift der Deutschen Morgenlandischen Gesellschaft* 144: 16-58.
Kosko, B. (1990). Fuzziness vs. probability. *International Journal of General Systems* 17: 211-240.
Kuhl, P. & J. Miller (1978). Speech perception by the chinchilla: Identification functions for synthetic VOT stimuli. *JASA* 63: 905-917.
Ladefoged, P. & I. Maddieson (1996). *Sounds of the world's languages*. Oxford: Blackwell.
Lamontagne, G. (1993). Syllabification and consonant cooccurrence restrictions. Unpublished Ph.D. dissertation, University of Massachusetts-Amherst.
Lane, E. (1863). *An Arabic-English lexicon*. London: Williams and Norgate.
Leben, W. (1973). Suprasegmental phonology. Unpublished PhD dissertation, MIT.
Levelt, W. (1989). *Speaking: from intention to articulation*. Cambridge: MIT Press.
Levitt, A. & A. Healy. (1985). The roles of phoneme frequency, similarity, and availability in the experimental elicitation of speech errors. *Journal of Memory and Language* 24: 717-733.
Liberman, A., K. Harris, H. Hoffman, & B. Griffith. (1957). The discrimination of speech sounds within and across phoneme boundaries. *Journal of experimental psychology* 54: 358-368.
Lightener, T. (1973). Against morpheme structure rules and other things. In *Issues in phonological theory*. The Hague: Mouton. 53-60.
Lindblom, B. (1983). Economy of speech gestures. In P. MacNeilage (ed.), *The production of speech*. New York: Springer-Verlag. 217-245.
Lindblom, B. (1990). Explaining phonetic variation: a sketch of the H&H theory. In W. Hardcastle & A. Marchal (eds.), *Speech production and speech modeling*. Dordrecht: Kluwer. 403-439.
Lindblom, B. & J. Sundberg (1971). Acoustic consequences of lip, tongue, jaw, and larynx movements. *JASA* 50: 1166-1179.

Lombardi, L. (1991). *Laryngeal features and laryngeal neutralization*. New York: Garland.

Lombardi, L. (1996). Postlexical rules and the status of privative features. *Phonology* 13: 1-38.

Luce, P. (1986). Neighborhoods of words in the mental lexicon. Unpublished Ph.D. dissertation, Indiana University.

Luce, P, D. Pisoni, & S. Goldinger (1990). Similarity neighborhoods of spoken words. In G. Altmann (ed.), *Cognitive models of speech processing*. Cambridge: MIT Press. 148-172.

MacKay, D. (1969). Forward and backward masking in motor systems. *Kybernetik* 2: 57-64.

MacKay, D. (1970). Spoonerisms: the structure of errors in the serial order of speech. *Neuropsychologica* 8: 323-353.

MacKay, D. (1972). The structure of words and syllables: evidence from errors in speech. *Cognitive Psychology* 3: 210-227.

MacKay, D. (1982). The problem of flexibility, fluency, and speed-accuracy trade of in skilled behavior. *Psychological Review* 89: 483-506.

Marslen-Wilson, W. (1984). Function and process in spoken word-recognition. In H. Bouma & D.G. Boudwhuis (eds.), *Attention and performance X: control of language processes*. Hillsdale, NJ: Lawrence Earlbaum.

Marslen-Wilson, W. (1987). Functional parallelism in spoken word-recognition. In U. Frauenfelder & L. Tyler (eds.), *Spoken word recognition*. Cambridge: MIT Press.

Marslen-Wilson, W. & A. Welsch (1978). Processing interactions and lexical access during word recognition in continuous speech. *Cognitive Psychology* 10:29-63.

Massaro, D. (1970). Retroactive interference in short-term recognition memory for pitch. *Journal of Experimental Psychology* 83: 32-39.

McCarthy, J. (1979). *Formal problems in Semitic phonology and morphology*. New York: Garland.

McCarthy, J. (1986). OCP effects: gemination and antigemination. *Linguistic Inquiry* 17(2): 207-263.

McCarthy, J. (1988). Feature geometry and dependency: a review. *Phonetica* 43: 84-108.

McCarthy, J. (1994). The phonetics and phonology of Semitic pharyngeals. In P. Keating (ed.), *Papers in laboratory phonology III: phonological structure and phonetic form*. Cambridge: Cambridge University Press. 191-283.

McCarthy, J. & A. Taub (1993). Review of C. Paradis and J.-F. Prunet, eds., The special status of coronals. *Phonology* 9: 363-370.

Medin, D., R. Goldstone, & D. Gentner (1993). Respects for similarity. *Psychological review* 100: 254-278.

Medin, D., & M. Schaffer (1978). Context theory of classification learning. *Psychological Review* 85: 207-238.

Mester, R. A. (1986). Studies in tier structure. Unpublished Ph.D. dissertation, University of Massachusetts-Amherst.

Mester, R. A. & J. Itô (1989). Feature predictability and underspecification: palatal prosody in Japanese mimetics. *Language* 65: 258-293.

Meyer, A. (1988). *Phonological encoding in language production: a priming study*. PhD thesis, Nijmegen University.

Mohanon, K. P. (1990). On the basis of radical underspecification. *Natural Language and Linguistic Theory* 9: 285-325.

Morton, J. (1982). Disintegrating the lexicon: an information processing approach. In J. Mehler, E. Walker, M. Garrett (eds.), *On mental representation*. Hillsdale, NJ: Erlbaum. 89-109.

Motley, M. & B. Baars. (1975). Encoding sensitivities to phonological markedness and transition probability: evidence from spoonerisms. *Human Communication Research* 2: 351-361.

Motley, M., B. Baars, & C. Camden (1983). Experimental verbal slips studies: a review and an editing model of language encoding. *Communication Monographs* 50: 79-101.

Mrayati, M., R. Carré & B. Guerin (1988). Distinctive regions and modes: a new theory of speech production. *Speech Communication* 7: 257-286.

Newman, R., J. Sawusch, & P. Luce (1996). The influence of underspecification and phoneme frequency in speech perception. Manuscript submitted for publication. University at Buffalo. Buffalo, NY.

Nosofsky, R. (1992). Exemplar-based approach to relating categorization, identification, and recognition. In F. Ashby (ed.), *Multidimensional models of perception and cognition*. Hillsdale, NJ: Erlbaum.

Noteboom, S. (1969). The tongue slips into patterns. In *Nomen society, Leyden studies in linguistics and phonetics*, A. Sciarone, *et al* eds. The Hague: Mouton. 144-156.

Nusbaum, R., D. Pisoni, & C. Davis (1984). Sizing up the Hoosier mental lexicon: measuring the familiarity of 20,000 words. Research on Speech Perception Progress Report no. 10. Speech Research Laboratory, Indiana University.

Ohala, J. (1990). Alternatives to the sonority hierarchy for explaining segmental sequential constraints. *Papers from the 26th Regional Meeting of the Chicago Linguistic Society*. Chicago: University of Chicago Press. 319-338.

Padgett, J. (1991). OCP subsidiary features. *Papers from the annual meeting, North Eastern Linguistics Society* 22.

Padgett, J. (1995a). *Stricture in feature geometry*. Stanford: CLSI.

Padgett, J. (1995b). Partial class behavior and nasal place assimilation. In *Proceedings of the Arizona phonology conference: workshop on features in Optimality Theory*. *Coyote Working Papers*. Tucson: University of Arizona.

Padgett, J. (1995c). Feature classes. Unpublished manuscript. University of California, Santa Cruz, CA.

Paradis, C. & J.-F. Prunet. (1991). *The special status of coronals*. San Diego: Academic Press.

Pierrehumbert, J. (1990). Phonological and phonetic representation. *Journal of Phonetics* 18: 375-394.

Pierrehumbert, J. (1993). Dissimilarity in the Arabic verbal roots. *Proceedings of NELS* 23. Amherst: GLSA Publications.

Pierrehumbert, J. (1994). Syllable structure and word structure. In P. Keating (ed.), *Papers in laboratory phonology III: phonological structure and phonetic form*. Cambridge: Cambridge University Press. 168-190.

Pierrehumbert, J. (1995). Prosodic effects on glottal allophones. In O. Fujimura & M. Hirano (eds.), *Vocal fold physiology: voice quality control*. San Diego, CA: Singular Publishing Group, Inc. 36-60.

Pierrehumbert, J. (1996). Cooperation and competition in the lexicon. Paper presented at SWOT. University of California, Irvine, CA.

Pierrehumbert, J. & M. Beckman (1988). *Japanese Tone Structure. Linguistic Inquiry Monograph*. 15. Cambridge: MIT Press.

Pierrehumbert, J. & S. Frisch (1996). Synthesizing allophonic glottalization. In J. van Santen, R. Sproat, J. Olive, & J. Hirschberg (eds.), *Progress in speech synthesis*. NY: Springer-Verlag.

Pierrehumbert, J. & R. Nair (1995). Word games and syllable structure. *Language and speech* 38: 77-114.

Pinker, S. & A. Prince (1988). On language and connectionism: analysis of a parallel distributed processing model of language acquisition. *Cognition* 28: 73-193.

Pisoni, D. (1973). Auditory and phonetic memory codes in the discrimination of consonants and vowels. In *Perception and Psychophysics 13*. 253-260.

Pollard, C. & I. Sag (1994). *Head-driven phrase structure grammar*. Chicago: CSLI.

Prince, A. & P. Smolensky (1993). *Optimality theory*. RUCCS-TR2. Rutgers University, New Brunswick, NJ.

Repp, B. (1984). Categorical perception: issues, methods, findings. In N. Lass (ed.), *Speech and language: advances in basic research and practice*, vol 10. 243-335.

Rosch, E. (1975). Cognitive representations of semantic categories. *JEP: Human Perception and Performance* 1: 303-322.

Rosch, E. & C. Mervis (1975). Family resemblance: studies in the internal structure of categories. *Cognitive Psychology* 7: 573-605.

Rosch, E., C. Mervis, W. Gray, D. Johnson, & P. Boyes-Braem (1976). Basic objects in natural categories. *Cognitive Psychology* 8: 382-439.

Sagey, E. (1986). The representation of features and relations in non-linear phonology. Ph.D. dissertation, MIT.

Sanchez, N. (1990). Of *pups* and **spups* in English. *CUNY Forum* 15: 7-18.

Saltzman, E. & K. Munhall (1990). A dynamical approach to gestural patterning in speech production. *Ecological Psychology* 1: 333-392.

Sawashima, M. & H. Hirose (1983). Laryngeal gestures in speech production. In P. MacNeilage (ed.), *The production of speech*. New York: Springer-Verlag. 11-38.

Scobbie, J. (1993). Constraint violation and conflict from the perspective of Declarative Phonology. *The Canadian Journal of Linguistics* 38:2, 155-168.

Sevald, C. & G. Dell (1994). The sequential cueing effect in speech production. *Cognition* 53: 91-127.

Shattuck-Hufnagel, S. (1979). Speech errors as evidence for a serial order mechanism in sentence production. In W. Cooper & E. Walker (eds.), *Sentence processing: psycholinguistic studies presented to Merrill Garrett*. Hillsdale, NJ: Lawrence Earlbaum.

Shattuck-Hufnagel, S. (1987). The role of word onset consonants in speech production planning. In *Motor and sensory processing in language*. E. Keller & M. Gopnick, eds. Hillsdale, NJ: Erlbaum.

Shattuck-Hufnagel, S. (1992). The role of word structure in segmental serial ordering. *Cognition* 42: 213-259.

Shattuck-Hufnagel, S. & D. Klatt. (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.
Silverman, D. (1995). Phasing and recoverability. Unpublished Ph.D. dissertation, UCLA.
Stanley, R. (1967). Redundancy rules in phonology. *Language* 43: 393-436.
Stemberger, J. (1982). *The lexicon in a model of language production*. PhD dissertation, University of California-San Diego.
Stemberger, J. (1991a). Apparent anti-frequency effects in language production: the addition bias and phonological underspecification. *Journal of Memory and Language* 20:161-185.
Stemberger, J. (1991b). Radical underspecification in language production. *Phonology* 8: 73-112.
Stemberger, J. & B. MacWhinney (1990). Form-oriented inflectional errors in language processing. *Cognitive Psychology* 18: 329-354.
Stemberger, J. & C. Stoel-Gammon (1991). The underspecification of coronals: evidence from language acquisition and performance errors. In C. Paradis & J.-F. Prunet (eds.) *The special status of coronals*. New York: Academic Press. 181-199.
Stemberger, J. & R. Treiman. (1986). The internal structure of word-initial consonant clusters. *Journal of memory and language* 25: 163-180.
Steriade, D. (1987). Redundant values. In *Papers from the 23rd regional meeting of the Chicago Linguistics Society*. Chicago: University of Chicago Press. 339-362.
Steriade, D. (1995). Underspecification and markedness. In J. Goldsmith (ed.), *The handbook of phonological theory*. Oxford: Blackwell.
Steriade, D. (1996). Paradigm uniformity and the phonetics-phonology boundary. Manuscript submitted for publication. UCLA, Los Angeles, CA.
Stevens, K. (1972). The quantal nature of speech: evidence from articulatory-acoustic data. In E. David & P. Denes (eds.), *Human communication: a unified view*. New York: MacGraw Hill. 51-66.
Stevens, K. (1989). On the quantal nature of speech. *Journal of Phonetics* 17: 3-45.
Stevens, K. & S. J. Keyser (1989). Primary features and their enhancement in consonants. *Language* 65(1): 81-106.
Thompson, W. (1995). Phonological similarity and speech errors. Review of J. Stemberger's "Radical underspecification in language production." Unpublished manuscript. Northwestern University, Evanston, IL.
Treiman, R., B. Kessler, S. Knewasser, & R. Tinkoff (1996). Adults' sensitivity to phonotactic probabilities in English words. Manuscript submitted for publication. Wayne State University, Detroit, MI.
Treisman, M. (1978). Space or lexicon? The word frequency effect and the error frequency effect. *Journal of Verbal Learning and Verbal Behavior* 17: 37-59.
Tukey, J. (1997). *Exploratory data analysis*. Reading, MA: Addison-Wesley.
Tversky, A. (1977). Features of similarity. *Psychological Review* 84: 327-352.
Tversky, A. & I. Gati (1978). Studies of similarity. In E. Rosch and B. Lloyd (eds.), *Judgement under uncertainty: heuristics and biases*. Hillsdale, NJ: Earlbaum.
Tversky, A. & I. Gati (1982). Similarity, seperability, and the triangle inequality. *Psychological Review* 89 (2): 123-154.

Van den Broeke, M., & L. Goldstein (1980). Consonant features in speech errors. In V. Fromkin (ed.), *Errors in linguistic performance: slips of the tongue, ear, pen, and hand*. New York: Academic Press.

Vitevich, M. (1996a). Spoonerisms, SLIPS, and the influence of phonological similarity neighborhoods. Manuscript submitted for publication. Language Perception Laboratory: University at Buffalo, Buffalo, NY.

Vitevich, M. (1996b). The neighborhood characteristics of malapropisms. Manuscript submitted for publication. Language Perception Laboratory: University at Buffalo, Buffalo, NY.

Watkins, M. (1975). Inhibition in recall with extralist “cues”. *Journal of Verbal Learning and Verbal Behavior* 14: 294-303.

Yip, M. (1988). The OCP and phonological rules: a loss of identity. *Linguistic Inquiry* 19(1): 65-100.

Yip, M. (1989). Feature geometry and cooccurrence restrictions. *Phonology* 6: 349-374.

Zadeh, L.A. (1965). Fuzzy sets. *Information and Control* 8: 338-353.

Zemlin, W. (1988). *Speech and hearing science: anatomy and physiology*. 3rd edition. Englewood Cliffs, NJ: Prentice Hall.

# Vita

## Stefan Frisch

**Born:** St. Paul, Minnesota, July 8th, 1968.

**Education:**

| | |
|---|---|
| Dec 1996 | Ph.D. in Linguistics, Northwestern University. |
| May 1993 | M.A. in Linguistics, University of Wisconsin at Madison. |
| 1990-1992 | Graduate studies in Mathematics, University of Wisconsin at Madison. |
| May 1990 | B.S. in Mathematics with high departmental honors, University of Illinois at Urbana-Champaign. |

**Publications:**

Frisch, S. (1997). The change in negation in Middle English: a NEGP licensing account. *Lingua* 101 (1/2): 21-62.

Pierrehumbert, J. & S. Frisch (1996). Synthesizing allophonic glottalization. In J. van Santen, R. Sproat, J. Olive, & J. Hirschberg (eds.), *Progress in speech synthesis*. New York: Springer-Verlag. 9-26.

Frisch, S., M. Broe, & J. Pierrehumbert. (1995). The role of similarity in phonology: explaining OCP-Place. *Proceedings of the 13th International Conference of the Phonetic Sciences*. vol. 3, 544-547.

Frisch, S. (1995). Evidence for economy of projection in historical change. In Jill Beckman (ed.), *Proceedings of the 25th meeting of the North Eastern Linguistic Society*. Amherst: GLSA Publications. vol. 2, 191-203.

Frisch, S. (1994). Reanalysis precedes syntactic change: evidence from Middle English. *Proceedings of the 5th meeting of the Formal Linguistics Society of Midamerica. Studies in the Linguistic Sciences* 24.1/2. 187-201.

Pierrehumbert, J. & S. Frisch. (1994). Source allophony and speech synthesis. In J. Olive, J. Hirschberg, J. van Santen, & R. Sproat (eds.), *Proceedings of the second ESCA/IEEE workshop on speech synthesis*. 1-4.

Frisch, S. (1994). The change in negation in Middle English. *Penn Review of Linguistics* 18: 29-44.

**Conference presentations:**

(unless listed under publications)

| | |
|---|---|
| Jan 1997 | "Alternatives to underspecification in language production." Paper presented at the 71st annual Meeting of the Linguistic Society of America. Chicago, IL. |
| Nov 1996 | "Against underspecification in speech errors." Paper presented at the 2nd annual Mid-Continental Workshop on Phonology. University of Illinois, Urbana-Champaign, IL. Submitted for publication. |

July 1996 “Temporally organized lexical representations as phonological units.” Paper presented at the 5th Conference on Laboratory Phonology. Northwestern University, Evanston, IL. Submitted for publication.

Jan 1996 “Similarity between OCP-Place effects and phonological speech errors.” Paper presented at the 70th annual Meeting of the Linguistic Society of America. San Diego, CA.

Nov 1995 “Effects of structured similarity in phonological speech errors.” Paper presented at the Mid-Continental Workshop on Phonology. Ohio State University, Columbus, OH.

Jan 1995 “Parallels among functional projections.” Paper presented at the 69th annual Meeting of the Linguistic Society of America. New Orleans, LA.

Jan 1993 “Unexploding exploded INFL.” Paper presented at the 4th annual Midwinter Student Conference in Linguistics. University of Wisconsin, Madison, WI.

**Professional Academic Experience:**

09/96-present Post-doctoral Fellow, Speech Research Laboratory, Department of Psychology, Indiana University. NIH Training Grant No. DC00012.

06/95-08/96 Research Assistant to Janet Pierrehumbert, Department of Linguistics, Northwestern University. NSF Grant No. BNS 9022484.

01/95-06/95 Teaching Assistant, Department of Linguistics, Northwestern University.

06/94-12/94 Research Assistant to Janet Pierrehumbert, Department of Linguistics, Northwestern University. NSF Grant No. BNS 9022484.

08/90-05/93 Teaching Assistant and Lecturer, Department of Mathematics, University of Wisconsin at Madison.

**Academic teaching experience:**

Lecturer

Ling C08: Scientific tools in the analysis of language. National High School Institute, Northwestern University.

Math 112: College Algebra. Department of Mathematics, University of Wisconsin at Madison.

Teaching Assistant

Cog Sci B10: Introduction to cognitive science (vision, language, and memory). Department of Cognitive Science, Northwestern University.

Ling B06: The structure of words and sentences. Department of Linguistics, Northwestern University.

Math 221: Calculus I; Math 222: Calculus II; Math 210: Finite Mathematics; Math 114: Algebra and Trigonometry. Department of Mathematics, University of Wisconsin at Madison.