Graded Constraints on English Word Forms

James L. McClelland and Brent C. Vander Wyk

Department of Psychology and
Center for the Neural Basis of Cognition
Carnegie Mellon University

Abstract

We present evidence that graded constraints determine the occurrence rates of the different rhyme types found in the ensemble of simple uninflected words in the English language. The rhyme types are defined in terms of vowel length (long vs. short), presence of particular post-vocalic elements, and their place of articulation. The rhyme types in the corpus (uninflected monosyllabic lemmas found in CELEX, which uses Southern British 'Received Pronunciation') conform to a template defined by a small number of absolute or categorical constraints. Among those forms consistent with the template, several graded constraints are identified, including constraints favoring short vowels, fewer segments, coronal places of articulation, and, when stops are present, absence of voicing. Such constraints induce a partial ordering over the expected rates of occurrence of different rhyme types; these have as special cases a pattern of implications for whether or not a form occurs at all (if form X occurs, then form Y should occur; if Z does not occur, then W should not occur). The constraints can be incorporated into a monotonic function characterizing the expected frequencies of occurrence of different rhyme types. Observed occurrence rates are better explained by a linear accumulation of constraints than by a multiplicative accumulation function. We also find that the constraints favoring coronals and short vowels are amplified when combined with other constraints, and are stronger in words of higher token frequency.

NOTE: This is a working draft created in February, 2006. It will be updated before submission. Not for Quotation; Comments Welcome by email to: jlm@psych.stanford.edu.

One goal of linguistic inquiry is to characterize the pattern of occurrence of linguistic forms. It has been common to treat the occurrence of candidate forms of a given type as an all-or-nothing affair: a form-type either is or is not acceptable, and so forms of the given type either can or cannot occur. Under this approach, the theorist attempts to provide a formal system, consisting of a structural framework and a set of rules or constraints on possible forms, which makes it possible to provide a perspicuous account of which types of forms are acceptable and which are not. The constructs used within the framework (e.g., the rules or constraints) operate in a categorical fashion: They rule candidates in or out, without allowing for graded differences in the relative degree of acceptability of different forms. For example, Harris (1994)provides a structural framework for characterizing the rhymes of English word-final stressed syllables and a set of statements cast within that framework that specify constraints on which rhymes can occur. One such statement is the following: "for rhymes of the form /VVlX/, where X is a stop, X must be coronal." This accounts for the non-occurrence of rhymes like /i:lg/ as in 'fielg' while allowing rhymes like /i:ld/ as in 'field' (the notation VV refers to any of a set of English vowels that Harris calls 'long', among which /i:/ is included).

In the present article, we make the case that considering only whether a form can or cannot occur ignores systematic facts about language, namely facts about differences in rates of occurrence of different forms. Indeed, we suggest that forms have a graded tendency to occur, conceived as a continuous underlying variable. We base our argument on data from a small subdomain, namely the rhymes occurring in monomorphemic monosyllabic word lemmas found in Southern British English. We rely on a count of the number of words containing rhymes of each possible type, based on the CELEX lemma corpus (Baayen, Piepenbrock, & van Rijn, 1993), with proper nouns and forms of questionable morphological status (such as *first* and *stealth*) removed. We present evidence that there are systematic gradations in the observed rates of occurrence of different types of rhymes, strongly suggesting graded differences among forms in their underlying acceptability or tendency to occur.

We then examine the factors that appear to govern the pattern of variation we see in the rate of occurrence of different forms. We find that much of the structure in these patterns can be accounted for by positing a small number of graded constraints. The constraints create strong implications for the relative rates of occurrence of candidate forms; the binary categorization of forms into those that do occur vs. those that do not occur are seen as consequences of the cumulative impact of the graded constraints.

We begin with an informal inspection of a subset of the data. Table 1 shows the average rate of occurrence of several types of rhymes containing at least one stop consonant in the corpus. Each rhyme type encompasses the set of rhymes all containing the indicted set of coda consonants and one of a set of vowels classified (following Harris, 1994) as 'long' or 'short'. Although no two vowels are strictly equivalent, a division into long (VV) and short (V) vowels has been shown to be useful in capturing a number of features of English rhymes (e.g. (Hammond, 1999; Harris, 1994).[1]

The table distinguishes between rhymes on the basis of vowel length, voicing of the stop and any other obstruents, and the place of articulation of the coda stop. For each combination of these variables, the rate of occurrence when the stop occurs alone is given in the first column, followed by its rate of occurrence with what we will call an embellishment, either a pre-stop liquid (designated l_), homorganic nasal (n_), coronal fricative (s_ or z_) or a post-stop coronal fricative (_s or _z) or a second coronal stop (_t or _d). The rates of occurrence are given as the average number of words with the given coda per vowel of the indicated type. Thus, the entry for the unvoiced coda /t/ occurring alone with a short vowel is 22.6, indicating that rhymes of this type, namely /Vt/, occur 22.6 times per short vowel in our corpus of monosyllabic English lemmas.

There are several things apparent in the table. First, there is a wide range of variation in the rate of occurrence of different forms. The rhyme /Vt/ occurs 22.6 times per vowel, whereas the rhyme /VVks/ occurs only 0.2 times per vowel (in the two words 'hoax' and 'coax'). This emphasizes the fact that taking notice only of binary distinctions between forms that do and do not occur would miss a hundred-fold range of variation in the rates of occurrence of different rhyme types.

Second, the variation in rate of occurrence is systematic. For example, both coda voicing and vowel length have an effect on how often a given rhyme is used: holding other factors constant, voiced codas and long vowels tend to result in lower rates of occurrence. In addition, holding other factors equal, coronal consonants tend to occur at higher rates than their non-coronal counterparts. The presence of any one of the indicated embellishments also tends to reduce the rate of occurrence, compared to the corresponding unembellished form.

Third, among the factors that influence the systematic differences in occurrence rate, no single constraint is dominant, and it is apparent that each constraint adds an additional penalty. For example, compared to a given unvoiced short vowel rhyme (/Vp/, /Vk/ or /Vt/), both the voiced short-vowel counterpart (/Vb/ et al.) and the unvoiced long vowel counterpart (/VVp/ et al.) tend to occur at lower rates. Furthermore, the rhymes that combine a voiced coda with a long vowel (/VVb/ et al.) are even less common than either the voiced short-vowel rhymes or the unvoiced long-vowel rhymes. Within all combinations of coda voicing and vowel length, embellished forms occur at lower rates than their unembellished counterparts. Finally, in many cases, rhymes containing coronal stops occur more frequently than their non-coronal counterparts. The overall pattern indicates that graded penalties against long vowels, voiced-codas, non-coronal places of articulation, and each of the different types of embellishments accumulate, with each violation contributing to the total penalty, thereby reducing the rate of occurrence of any given rhyme type relative to its counterparts violating fewer of these constraints.

---

[1] Not all English vowels contribute to the entries in the table, and forms with a consonant preceded or followed by a fricative other than /s/ have also been excluded. The vowel restrictions and the few other cases involving fricatives are discussed in the fuller analysis presented below.

This informal analysis, though far from complete, demonstrates several of the fundamental points that will be explored in the remainder of the paper. There is systematic variation in the occurrence rates of word forms. And moreover, this variation can be fruitfully described using a small set of graded parameters, with no single parameter showing absolute dominance over others.

Relation to Other Work

We know of little formal theory directed at the explanation of graded patterns of different forms occurrence rates. However, differences in forms' characteristics have been shown to produce graded effects in a variety of linguistic and non-linguistic tasks such as goodness judgments (Coleman & Pierrehumbert, 1997; Frisch, Broe, & Pierrehumbert, 2004; Frisch, Large, & Pisoni, 2000), nonword repetition (Vitevitch & Luce, 1998), speech errors (Goldrick, 2004), phoneme identification (Pitt & McQueen, 1998), and recognition memory (Frisch et al., 2000). It is true that there is some work in which differences in rates of occurrence have been discussed (Harris, 1994; Kessler & Treiman, 1997). But attempts to develop a formal framework that characterizes which forms can and cannot occur have not fully integrated these graded differences, and the role of the graded accumulation of constraint violations has not been explicitly addressed. For example, (Harris, 1994) often alludes to what he calls *preferences*, e.g. for coronal relative to non-coronal rhyme types, but does not systematically consider whether the patterns of occurrence vs. non-occurrence he describes could be explained by the accumulated weight of a set of such preferences.

It is important to note that we do not intend to be critical of Harris' efforts. Indeed, his analysis along with the Optimality-Theory based approach of Hammond (1999) provides a crucial starting point for our work. We see our effort as building on their insights into the factors that influence occurrence and non-occurrence, and indeed differences in rates of occurrence (Harris's preferences) of English rhyme-types. Some of their observations, e.g. that coda complexity is more restricted with long vowels, and the special status of coronal consonants, are fundamental to our further analysis. Our effort here is to suggest that it may be profitable to place primary theoretical focus on the graded differences in occurrence rates. This perspective may allow us to characterize the preferences Harris has pointed to in formal terms, and to explain much of the data on which forms do and do not occur in terms of the cumulative consequences of graded constraints that also explain differences in relative rates of different form's occurrence.

Several constraints were mentioned above, including vowel length, coda voicing, and coronality. These constraints themselves are not new with us, and indeed have been prominently considered in other approaches, including Harris (1994) and Hammond (1999). As do many other authors, we remain agnostic on the exact source and nature of the constraints. It may be that they arise from articulatory, perceptual, and communicative considerations (Lindblom, MacNeilage, & Studdert-Kennedy, 1984; Redford, Chen, & Miikkulainen, 2001). Historical considerations may also play a role. Our goal here is to investigate how the constraints combine to determine the relative rates of occurrence of forms, while leaving the determination of their basis to other investigations.

Our approach shares some common elements with Optimality Theory (OT) based approaches to phonology (Boersma, 1998; Hammond, 1999; Prince & Smolensky, 2004), and we see it as extending the line of thought introduced in the earliest writings on OT (Prince & Smolensky, 2004), which applied the connectionist principle of graded constraint satisfaction (Rumelhart, Smolensky, McClelland, & Hinton, 1986; Smolensky, 1986) to the analysis of language structure. Specifically, OT makes use of constraints to determine the acceptability of forms. Both the constraints in OT and those of the current proposal can be thought of as having graded strengths. However, there are two important differences. First, in OT, the constraint strengths are assumed to be assigned in such a way as to create a strict dominance hierarchy: Whether a form is acceptable is determined wholly by the highest ranked constraint violation. Furthermore, acceptability in OT is not a matter of degree: The acceptability of a candidate form is an all or none matter. This is in clear contrast to the way that constraint violations are accumulated in the current proposal. In our approach, each constraint violation imposes a graded penalty, resulting in a graded and continuous constraint violation score. Constraint violations can build up to the extent that a form will not occur at all; further violations cannot, of course, decrease the rate of occurrence any further.

As in OT, constraints in our approach can have different strengths. This means that some constraints can have a larger impact than others. In principle, continuous strength values can be assigned over a wide enough range to create the effect of strict dominance. Thus, we are not claiming that dominance-like constraint relations do not and cannot exist. We are only claiming that strict dominance must be relaxed to capture the pattern of differential occurrence rates, and can contribute to a perspicuous account of patterns of occurrence and non-occurrence. As we shall see, the relative magnitudes of the weights derived for the constraints needed to account for the graded patterns seen in our corpus are not large enough to approximate strict dominance.

We are not the first to suggest that a collection of graded constraints can conspire together to determine whether a form can or cannot occur. Indeed, our approach is strongly influenced by the work of Burzio (2000). Coming from a starting point within OT, Burzio has already pointed out how combinations of constraints can work together to determine, for example, the phonological shape of the past-tense inflection of an irregular English verb. Burzio suggests that morphological processes, especially those of limited productivity, appear to reflect graded constraint summation. Another important effort stressing graded constraints can be found in the work of Kessler and Trieman (1997), which clearly demonstrates that there are indeed graded differences in the occurrence rates of different forms that cannot be attributed to chance variation. A concept very similar to graded constraints plays a role in a version of OT called Functional Phonology (Boersma, 1998; Boersma & Hayes, 2001) that addresses certain probabilistic patterns of alternation in the realization of certain word-forms. This approach could potentially be adapted to address some of the phenomena we consider here.

The Graded Constraint Theory

The informal examination above showed that there are systematic differences in the rates of occurrence of different rhyme types, and it appears that these systematic differences can be described by a small set of graded constraints. These constraints capture differences in relative rates of occurrence of the various candidate forms and imply binary existence entailments between related forms. The goal here is to provide a formal framework for representing the constraints and for relating them to entailments for relative rates of occurrence of different forms, including entailments concerning which forms will or will not occur at all. We now introduce such a framework, called the Graded Constraint Theory, hereafter GCT.

The Tenets of GCT are as follows:

1. Phonological forms vary in the strength of their tendency to be used in lexical items. Each form F_i has a strength S_i.

2. The strength of a form F_i is a uniformly decreasing function (represented by D) of each of the members of the set of constraints $\{C\}_i$ that the form violates, where each constraint violation serves to reduce the strength of F_i.

   $$S_i = D(\{C\}_i)$$

3. We assume that each form F_i has a non-negative tendency to be employed in a language. This tendency is represented by a Poisson rate variable R_i. While more complex relationships are possible, we consider the simple case in which R_i is simply set equal to S_i when S_i greater is than 0, and to 0 otherwise:

   $$R_i = [S_i]+$$

Several of the concepts introduced in 1-3 require further discussion.

The Concept *Uniformly Decreasing Function*

The concept *uniformly decreasing function* draws on definitions introduced by Williams (1986). The starting place is his definition of the concept *uniformly monotonic function of a given argument*. Formally, a function *M* is uniformly monotonic in a given argument *a* if for all possible combinations of values of the function's other arguments, the value of *M* is a monotonic function of *a*, and the sense of this monotonicity (increasing or decreasing) is the same for all of these combinations. The concept *uniformly decreasing function of a given argument* involves the further restriction that the sense of the monotonicity is decreasing (here we use 'decreasing' as a shorthand for 'monotonically decreasing'). By a similar extension of a further definition offered by Williams, we say a function D is *uniformly decreasing* if it is uniformly decreasing in each of its arguments.

One example of a function that is a uniformly decreasing function of all members of a set of constraints C_j is:

$$S_i = \sum_j w_j C_j + \beta, \; w_j < 0 \text{ for all } j$$

Additional conjunctive terms reflecting joint effects of constraints can be added (e.g. $w_{jj'} C_j C_{j'}$) as long as the weights are all negative. For simplicity we will restrict our analysis to cases in which the constraints C_j take values 1 or 0 to indicate when the form violates (1) or does not violate (0) constraint j. However, note that this function would still fit the definition if the constraints themselves were graded ($C_j \geq 0$ for all j). As long as the constraints remain non-negative and the weights remain negative, the function above, even with conjunctive terms added, remains uniformly decreasing.

Another example is:

$$S_i = \alpha \prod_j \omega_j^{C_j}, \; 0 < \omega_j < 1.$$

Again additional factors can be included for joint effects of constraints ($\omega_{jj'}^{C_j C_{j'}}$), as long as all the ω's lie in (0,1). The function is in fact a monotonic transformation of the previous function, with the ω's here corresponding to the logs of the w's there, so that a w of 0 corresponds to an ω of 1 and increasingly negative values of w correspond to values of ω getting closer and closer to 0. Once again graded constraints ($C \geq 0$) are consistent with the formulation: That is the function remains uniformly decreasing.

Note that other functions obeying (2) are also possible.

Function Relating Strength to Rate

We have specified a very simple relation between a form's strength S and its rate R. We do not intend to reject the possibility that there may be a more complex relationship than is captured by the threshold function $[.]^+$. However, some more complex choices would be difficult to identify separately from the function D itself. Indeed we can see the simple relationship we have specified to be one that forces all the work in explaining variation in form's occurrence rates into the function that determines the form's strength. We will consider an alternative in the General Discussion.

Hypothetical Status of Underlying Rate Variables R_i

It is important to understand that the above apparatus introduces a formal theory containing hypothetical constructs, namely Constraints, Weights, Strengths, and Rates, none of which are directly observable *per se*. In particular, it is important to be clear that the rates R_i associated with particular forms F_i are not seen as equal to the actual frequencies of occurrence. Instead, they are viewed as governing a random process through which the particular ensemble of words we have in English was established. This process is subject to a wide range of factors, including historical use in precursor languages, borrowings from other languages, and invention (*hound*, *beige*, and *fax* are examples of each type). Our theory does not address these factors, and simply treats them as unpredictable, so that the ensemble of specific forms that are actually used cannot be fully specified. What can be provided are probabilistic limits on the rates of occurrence of forms of different types, as discussed below.

From Underlying Rates to Observed Average Occurrence Rates

In GCT, the observed rates of occurrence of forms obey the Poisson distribution, a distribution that has been used fruitfully to describe many processes in which the number of events of a given type appears to be governed by a random process with a given rate. The Poisson distribution is usually written:

$$p(k) = (R^k/k!)e^{-R}$$

This function gives, for every possible value of $k \geq 0$, its probability of being observed in a sample taken from the distribution, under the assumption that the underlying rate of occurrence is R. Figure 1a shows the Poisson distribution function for several values of R (1, 2, 5, 10, and 20). There are two points to notice. First, when R is small, there is a marked probability of observing an actual occurrence rate of 0. As R grows larger, however, we can see that the probability of observing 0 occurrences gets very small very quickly. For R of 5 or greater, the probability of observing 0 occurrences becomes smaller than 1 in 100. The second point to notice is that both the mean of the Poisson and its spread or standard deviation increase with R. In fact the mean is simply equal to R, and the standard deviation is equal to $\sqrt{R}$. Note that R can be any positive real number; it is not restricted to integer values, even though the observed rates are integers.

One virtue of the Poisson is the fact that the sum of several independent Poisson processes is itself a Poisson process whose rate is equal to the sum of the component process rates. So, from the assumption that the word-specific rates are Poisson processes, it follows that the rates of occurrence of words with a particular rhyme are also Poisson processes. This is an extremely useful result since the word-specific rates are too small for reliable observation. Accordingly, it is necessary to aggregate. One possibility would be to aggregate over all words with a given rhyme (say, _/it/). Because the rates for individual rhymes tend still to be quite small in most cases, it appears necessary to further aggregate over word-rhymes to examine *rhyme types*. This aggregation further reduces the sampling variability so as to allow for the possibility of statistically reliable evaluations, as described below. Inspired by Harris (Harris, 1994), a rhyme type specifies a type of vowel (long, represented VV, or short, represented V) plus a specific sequence of post-vocalic consonants. For example, /Vst/, /VVnd/ and /VVmp/ are three different rhyme types. The observed average occurrence rate of monomorphemic monosyllabic word lemmas a_i with rhymes of type i, will be the primary data for evaluating the theory:

$$a_i = 1/n_i \sum_j o_{ij}$$

Here n_i is the number of specific rhymes of the given type. The summation ranges over all of these rhymes, and o_{ij} represents the actual number of occurrences of the rhyme j of type i.

The value of a_i for all rhymes that occur in monomorphemic monosyllabic English word lemmas can be found in Table 2. The a_i are treated as arising from n_i independent samples from a set of Poisson distributions with average rate R_i. Note that the different vowels within a type need not have equivalent rates under this formulation; all that is needed is that the average of their rates is R_i.

The distribution for the average of n Poisson variables with mean rate R is considerably tighter than the distribution of a single Poisson with rate R. Specifically, while the mean is the same, the standard deviation is reduced by a factor of $1/\sqrt{n}$. The effect of this for average rates R = 1, 5, 10 and 20 are shown for n = 5 and 10 in Figure 1b and c. Note among other things that when n is 5 or 10, forms with average rates of 1 or greater are unlikely to be completely missing from the corpus. For forms with average rates much less than one but still actually greater than 0, it may not be unlikely to observe 0 occurrences.

Testing the Graded Constraint Theory

The assumptions above provide a way of exploring the extent to which the actual occurrence rates of particular forms are consistent with an instantiation of GCT. This instantiation would involve a set of constraints C_j, and either a specific function D or a weaker formulation in which the exact form of D is unknown and it is simply assumed to be uniformly decreasing in each of the constraints. We can refer to these different types of instantiations as *parametric* and *nonparametric* instantiations of the theory. In our analyses below we will consider both types of formulations. Here we consider how the predictions of such models might be tested against available data.

In a parametric instantiation, the theory will predict rates R_i of occurrence of each of the form types within some data set, based on a specific choice of the function D and a specific set of values assigned to the parameters of the function. We can then consider whether a particular observed value of a_i is consistent with the specified underlying value of R_i. To address this question, we rely on the logic of hypothesis testing used in standard statistical inference. According to this logic, we ask, what is the region within which a_i is likely to fall with some fairly high probability *p*, given R_i. When a_i falls outside this region, we will become suspicious about the hypothesis that R_i is the correct value. The choice of *p* reflects how confident we require ourselves to be before we begin to become suspicious. While .95 is often the minimum standard used in statistical hypothesis testing, we choose a smaller value, *p*=.9, as a threshold for becoming suspicious since we do not wish to miss possible violations of the theory. The region is chosen so that half of the excluded area under the PDF of R_i is above the region and half of it is below the region. Thus in this case if we find a value of a_i less than the lower boundary of the critical region, the probability of observing a value which is as small as or smaller than the observed value when R_i is actually the correct rate is less than 0.05. It should be noted that with *p*=.90, there is a 10% chance that even in case R_i really is the correct underlying value, we would obtain an observation that would arouse our suspicions against it. Given that we will typically be testing multiple predictions, this means that we should expect to have our suspicions raised about 10% of them even if the theory is fully correct. Because of this, whenever we observe a value outside of the 90% confidence interval, we report the exact probability of observing a value as far or further from R in the same direction. The smaller this probability is, the more confident we can be that the predicted value of R is not correct.

In a non-parametric instantiation of the theory, we are agnostic about the exact form of the function D or the weights of the various constraints, but we can still make ordinal predictions about the strengths of various forms. Doing so makes use of the following observation. Given assumption (2) of GCT, which states that the strength of a form is a uniformly decreasing function of the set of constraints the form violates, it follows that:

$$\text{If } \{C\}_{i'} \subset \{C\}_i, \text{ then } S_{i'} > S_i$$

In words, if the set of constraints violated by form i' is a proper subset of the set of constraints violated by form i, then the strength of form i' will be greater than the strength of form i. The symbol $\subset$ denotes a proper subset relation. Given a set of constraints and a set of rhyme types, each of which violates a different subset of the constraints, this observation immediately establishes a large number of predictions about the relative rates of occurrence of forms. However, it is still necessary to take into account that the observed average occurrence rates of such forms are samples from Poisson distributions that depend on the forms' underlying strengths. Given this, we now consider what stance we can take with respect to the hypothesis $S_i' > S_i$, based on the obtained values of a_i' and a_i, the observed average counts of the corresponding form types. We might be somewhat suspicious whenever $a_i' < a_i$; if in reality $S_i' > S_i$, then we should expect $a_i' > a_i$ at least half of the time, so the hypothesis is more likely to be wrong than not whenever we observe a reversal. But we cannot be certain that S_i' is not greater than S_i, since sampling variability would be expected to produce an apparent reversal of the relative strengths of the forms some of the time.

Intuitively, of course, the larger the difference between a_i' and a_i the more certain we can be that this reflects the true underlying relation. To increase the precision of this reasoning, we can generate an estimate of the likelihood of the observed difference, $a_i'-a_i$ under the hypothesis that $S_i'=S_i$, then use this to reason further about the likelihood of the observed difference under the hypothesis that $S_i'>S_i$. Conceptually, a distribution with a rate of $(a_i'+a_i)/2$ (an average of the observed rates) is constructed, from which we imagine repeatedly drawing two samples (of sizes n_i' and n_i) at random. This process will produce a distribution of differences in average observed rates whose PDF can be derived analytically, under the hypothesis that the underlying rate of occurrence of each of the two form types is actually the same. We can then determine the *p*-confidence interval for the value of this difference. If $a_i' - a_i$ is in the wrong direction under our hypothesis (i.e., it is negative when the hypothesis specifies $S_i'>S_i$) and falls below the lower bound, we should become quite suspicious about the hypothesis. The probability of a difference falling below the lower bound would be only p = (1-P)/2 even if S_i‘ and S_i were equal, and the probability would be even smaller under the hypothesis that S_i’>S_i. As before, given that multiple comparisons are being made, we expect that some fraction of them will be called into question by this method just by chance, so we present the actual value of p for the difference obtained under the hypothesis S_i’=S_i. This is an upper bound on how unlikely it would be to obtain a value of a_i’-a_i as or more negative under the hypothesis S_i’>S_i.

In practice, we use two different statistical tests to assess the statistical reliability of the a_i’-a_i difference. One such test relies on the sum, not the average, of the counts of the rates of occurrence of the rhymes contributing to types i and i’. Using these summed counts, the probability of a difference equal to or greater than that observed can be calculated, using the exact binomial test or its normal approximation (Lowry, 1998-2005). The second test is Student’s *t*-test, with the vowel-specific frequencies treated as the individual observations used to estimate the variability associated with the difference between the sample averages. This latter approach has the advantage of automatically providing an indication of the degree to which the pattern observed is consistent across vowels, something that the first approach lacks. Also, for pairs of rhyme types that had the same vowel length (e.g. /Vp/ vs. /Vb/), we are able to perform paired *t*-tests, where the members of the pair share the same vowel; this maximizes sensitivity to possible differences due to the constraint under consideration. When the vowel type is different, i.e., one is long and the other short, this pairing is not possible, so an unpaired *t*-test was performed. It should be noted that the probability estimates provided by the *t*-test is based on an assumption that the variation within the items in each set that is being compared is normally distributed with the same variance, and this is unlikely to be true with rhymes occurring with very low observed rates. The tests have different strengths and weaknesses, and are of complementary value in assessing the reliability of a violation of predictions; the *p*-values provided by either should be viewed as approximate and indicative rather than exact.

Corpus Details and Restrictions

Before proceeding to detailed analyses based on the framework above, we describe in more detail the corpus used in our analysis. The specific forms used in the analyses were drawn from the spoken British English CELEX lemma database. The lemmas were filtered to remove the multisyllabic forms, as well as any either author considered to be even marginally morphologically complex. Removed forms included those carrying specific tense or number restrictions such as *should*, *would*, etc, and all other forms with signs of derived status such as *first* and *wealth*. This left a set of 3474 monosyllabic, monomorphemic lemmas. Note that CELEX uses the Southern British English dialect sometimes called ‘Received Pronunciation’, in which rhymes containing the orthographic symbol “r” do not contain a liquid /r/ in their phonology; instead the ‘r’ affects the nature and quality of the preceding vowel, so that, for example, 'car' rhymes with 'shah'[2].

For most of our analyses, we consider two sets of vowels, one designated 'short' and the other designated 'long'. The short vowel set excludes the vowel used in most English dialects in *book*, and the long vowel set excludes the vowels in *pier, poor* and *pair*; these cases are relatively rare and vary across dialects. The surviving vowels include five that are short (those in *pit, pet, pat, putt*, and *pot*) and ten that are long. The long vowels are arguably a mixed lot, including several that occur in substantial numbers of words with orthographic /r/ (those in *barn, born,* and *burn*), and two that have been identified by Hammond (1999) as super-heavy (those in *boy* and *cow*). To ensure that our findings did not depend on either of these types, we repeated most of them using only the five remaining long vowels (those in *bay*, *be*, *buy*, *beau*, and *boo*).

A final point to note is that only the first-indicated pronunciation of each word-form was used, and this tended to be the citation form (so that the indefinite article rhymes with *bay* and the definite article rhymes with *see*). These forms are, of course, typically reduced in running speech. Our analysis does not consider such reductions, and thus should be construed as applying to the lemmas under consideration when they occur in stressed sentential positions.

Since our corpus does indeed reflect a restricted class of forms, it is worth considering what implications this restriction may have for our analyses. We consider first the fact that the analysis ignores the phonological content of word onsets. While there are differences in rates of occurrence of different word onsets, it appears that the contributions of onsets and codas to a form's relative goodness are largely independent. Indeed, classical phonological theory proposes a syllabic partitioning based on an onset and a vowel-coda rhyme (Fudge, 1969) motivated partly by the fact that, with only a few exceptions, there appear to be few systematic constraints on the co-occurrence of onsets and rhymes (see also (Harris, 1994)). Some previous analyses have found quantitative dependencies between onsets and codas (Diver, 1979), but the dependencies were mostly fairly weak and a substantial body of work supports the validity of the near-independence of onsets and rhymes (Frisch et al., 2000; Pierrehumbert & Nair, 1995; Treiman, 1988; Treiman & Kessler, 1995, 1997; Treiman, Kessler, Knewasser, Tincoff, & Bowman, 2000). A separate analysis focusing on different types of word onsets would likely yield evidence of graded constraints operating there as well, though preliminary investigations we have conducted suggest they are somewhat less systematic. In any case, we have chosen to focus on rhymes as targets for analysis rather than onsets because of the greater variety and complexity of rhyme types in English.

---

[2] In CELEX, forms ending orthographically in *nch* where usually coded as /nC/ but occasionally as /nS/, and those ending in *nge* were usually coded as /nZ/ but occasionally as /nJ/. The choice did not appear to be systematic. There may be dialectical and idiosyncractic variation in the pronunciations of some of these forms, but the two variants are never used contrastively and in all cases the affricate versions appeared at least acceptable variants, so we recoded /nS/ to /nC/ and /nZ/ to /nJ/ in every case.

There are distributional differences between multisyllabic and monosyllabic forms, and between inflected and uninflected forms. Syllables embedded in multi-syllabic forms tend to be simpler and their complexity may interact with the word's stress pattern (Hammond, 1999). On the other hand, as Harris discusses, clusters between a medial stressed syllable and a subsequent unstressed syllable appear to relax some of the constraints found in complex codas (which Harris analyzes as containing the onset of a syllable whose vowel has not been realized). Furthermore, inflected word forms can violate constraints that are not violated in uninflected forms. For instance, although the allophonic variation of the regular past-tense marker is dependent on properties of the previous phoneme, certain post-vocalic combinations occur in the rhymes of inflected past tenses that do not occur in any uninflected cases, for example the rhymes in 'beeped', /VVpt/ and 'rubbed', /Vbd/ (Burzio, 2000). It appears likely that some of the constraints that hold in monosyllabic, monomorphemic forms are relaxed in syllabically and morphologically complex forms, but it seems likely that this relaxation will turn out to be partial, and that such forms will reflect graded sensitivity to phonological constraints. The current study provides a base for future work on this issue.

Finally, the restriction to forms occurring in stressed positions can be seen as responsible for one of the non-graded constraints that applies to the forms in our corpus: In English such word forms may not contain a short vowel in isolation. In unstressed positions, of course, this constraint is relaxed, giving /@/ for *a*, /D@/ for *the*, and /t@/ for *to*, etc. This reflects a general tendency for unstressed material to be reduced, one which could be addressed with graded constraints in a future extension of our analysis.

Analyses 1: Non-Parametric Analysis

With the above considerations in place, we are ready to test predictions of the GCT for the relative occurrence rates of different types of monosyllabic, monomorphemic lemmas. We begin with non-parametric analyses where we consider whether, within each of several classes of rhyme types, an instantiation of the GCT based on a small number of constraints can provide a perspicuous account of the relative occurrence rates of the various rhyme types. The non-parametric analysis does not require a specification of the function D mentioned above. Later, we will explore several formulation of D in further analyses of a subset of rhymes found to be particularly well characterized in the non-parametric analysis.

According to the graded constraint theory, when one form C_i' violates a proper subset of the constraints violated by a form C_i, the strength of the former should be greater than the strength of the latter:

$$\text{If } \{C\}_i' \subset \{C\}_i, \text{ then } S_i' > S_i$$

Given a set of constraints, a clear partial ordering of the observed average occurrence rates of forms of particular types is implied, and a set of forms can be evaluated to determine if there are any violations of the partial ordering. There is, however, an indeterminacy introduced by the provision in the theory that strengths may range below

0, and that the underlying rates of occurrence of forms cannot. In such cases, although the underlying strengths may still obey $S_{i'} > S_i$, the associated rates of one or both forms will be 0. If only the weaker form falls below 0 the above implication still holds but if the stronger form falls below 0 then the implication is that the weaker one must fall below 0 as well, with the consequence that the observed rates of both forms will be 0. Thus, extending the above to rates, we have:

If $\{C\}_i' \subset \{C\}_i$, then $R_i' > R_i$ or $R_i' = R_i = 0$.

Here we will consider the extent to which the rhyme types of English monomorphemic monosyllables exhibit the patterns we should expect from a specific small set of graded constraints.

Our analysis is grounded within the observation that all the rhymes in the corpus can be specified within the following frame:

[V|VV] (l) (N) (F) (S) (F) (S)

In words, the rhyme consists of a short or long vowel, optionally followed by any of the following, occurring in the order indicated: an l, a nasal, a fricative, a stop, a fricative, and a stop. By convention, a form in represented by placing its constituents in the left-most positions possible, while still maintaining the correct order. Thus a word like *clasp* fills the first fricative and stop slots, while a word like *lapse* fills the first stop and second fricative slots. The final stop slot is used only when there is a second stop, as in *act*. The template encompasses all of the forms in our corpus, including the word *text*, which uses the first stop, second fricative, and final stop slots.

In addition to the restrictions placed on forms by the above, they further adhere to the following strong constraints: (1) short vowels cannot occur alone; (2) when there is a nasal it must be homorganic with the following stop or fricative; (3) the last two slots are restricted to coronals; (4) all obstruent material in any of the last four slots is always homorganic with respect to voicing; (5) a consonant cannot appear adjacent to itself (e.g. /ett/ is not allowed). Since there are no violations of these constraints, we can think of them either as very strong graded constraints or as hard or absolute constraints; we will consider graded interpretations of some of these as we proceed with our analysis.

Within the limits placed by the hard constraints identified above, the issue we now consider is the following. Can we account for the types of forms that do and do not occur, and for the relative rate of occurrence of these forms, in terms of a small additional set of graded constraints? The constraints we will consider are:

C1:	↓X	Added segments are dispreferred
C2:	↓VV	Long vowels are dispreferred
C3:	↓VO	Voiced obstruents are dispreferred
C4:	↓NC	Non-coronal segments are dispreferred

Additional constraints that apply only among coronal fricatives will be introduced when we consider such segments.

While the set of constraints is conceptually quite compact, it should be noted that constraint C1 stands for a set of constraints, each with its own strength, one for each type of added segment (liquid, nasal, fricative, or stop). Similarly, C4 stands for a set of constraints each operating against a specific non-coronal closure type (labial or velar, for example). The potential differences in strength need not be considered within our non-parametric analysis, but they are relevant to the partial ordering predictions. For example, /VVk/ and /Vk/ both violate the same instance of C4 (the constraint against velars) so the theory predicts $S_{/VVk/} < S_{/Vk/}$. On the other hand, /VVk/ and /Vp/ violate different instances of C4, which may differ in strength, preventing a clear prediction of the relative values of $S_{/VVk/}$ and $S_{/Vp/}$.

In what follows, we proceed in several steps. We first undertake a global analysis of the entire set of forms, concentrating on C1, the constraint against adding segments. We then consider several classes of English rhyme types; The first three classes will be classes that include an obstruent: those that contain at least one stop, those that contain an affricate, and those that contain a fricative. The fourth class we will consider encompasses the remainder of the possible forms consistent with the template, i.e., those that do not contain an obstruent. In each case, the question we ask is the following:

> Of the eligible candidate rhyme types, are the ensemble of types that occur at all and the relative rates of occurrence these types consistent with the partial ordering implied by the graded constraints?

Our method considers each form in relation to every other form that matches it in all ways except one that corresponds to removing a single constraint violation. Such forms are called the immediate superordinates of each target form. If a given target form exists, its immediate superordinates should exist. Furthermore, its immediate superordinates should have a higher average observed rate, subject only to sampling variability, which could produce some apparent reversals. As described above, when such reversals occur, statistical tests will be used to calculate the certainty with which we can assert a given relationship in the underlying rates based on the observed average rates.

Constraint Against Adding Material

Constraint C1 entails that if a form exists, all forms that can be made by deleting one of the constituents of the form should also exist, and and all should occur at higher rates. For example, if rhymes ending in /Vkst/ (as in *text*) are found, rhymes ending in /Vks/, /Vkt/ and /Vst/ should also exist, and all should occur more frequently that /Vkst/. These entailments hold in this case: we have words like *axe*, *act* and *fast*, and indeed there are more words of each of these types than there are words of the *text* type
.

We searched the entire tree of possible forms from the bottom (most constraints violated) up for cases of violations of the stated existence and relative occurrence rate relationships. Only one violation was found: the rhyme type /VNk/ (an average of 12 occurrences per vowel) occurs more frequently that its immediate superordinate /VN/ (an average of 9.8 occurrences per vowel). However, this difference is not significant in either the binomial test (p=0.169) or in a paired t-test (t(4)=-0.81, p = 0.462). The data are in striking conformity with the principle that the addition of material creates weaker forms. It may be noted, however, that the hard constraint prohibiting short vowels from occurring alone can be viewed as protecting our analysis from a major pattern of violations. That is, in the absence of this hard constraint, we would expect forms with a single short vowel to occur more frequently than any forms with short vowels followed by one or more subsequent consonants. Later we will consider the possibility that this pattern itself reflects a graded constraint, but one that requires a slightly more sophisticated formulation than a general bar against added material.

In evaluating the conformance of the rhymes to the remaining constraints, we will deal with several classes of rhyme types separately. The classification of the rhyme types is based on how they fill out the rhyme template above, and is partially motivated by the fact that, definitionally, the rhyme types are subjected to the constraints in different ways. For instance, rhyme types whose codas consist only of a liquid cannot be evaluated with respect to voicing or place constraints, and rhyme types containing fricatives introduce three alternative coronal fricatives, requiring the consideration of an additional constraint.

Rhymes Containing at Least One Stop

Stop rhymes are defined as those rhymes that include a stop consonant (t, d, k, g, p, or b) along with an optional embellishment of a pre-stop liquid, nasal, or fricative, and an optional post-stop fricative or second stop. Thus the present analysis applies to forms consistent with the following template:

[V|VV] (l) (N) (F) **S** (F) (S)

Our main analysis considers the embellishments listed as separate columns in Tables 1 and 2 (l_), (N_), (s_), (_s), and (_t), which nearly exhaust the relevant extant forms. The additional case of a pre-stop /f/ will be considered separately at the end of this section, and the following section shows that the analysis also extends to encompass affricates, analyzed as a stop followed by a subsequent fricative (/S/ or /Z/).

Rhymes of the eligible types are found in 1565 words, a good portion of the corpus. They have a rich immediate superordinate relationship structure which gives rise to the set of partial order structures shown in Figure 2. The figure displays the partial ordering constraints holding among a subset of the possible forms, namely those containing a vowel, a stop, and either no embellishment or any one of the possible single embellishments. Arrows represent predicted strength differences between forms based on Constraints C1-C4. For example, the arrow from /Vk/ to /Vt/ represents the fact that the strength of /Vt/ is predicted to be greater than the strength of /Vk/ based on C3

(↓NC). Forms that do not occur (and arrows to them) are shown in gray; the gray arrows represent cases where strength differences are predicted but, because the stronger item already has a rate of 0, the weaker item cannot occur less frequently. It would, however, be a potentially important violation of the theory if the weaker form occurred and the stronger form did not occur, unless the non-occurrence of the stronger form could be attributed to sample variability.

Note that the set of partial ordering predictions holding among the set of forms that contain no embellishments also holds among the set of forms that contains each of the possible embellishments (a special case arises for the forms containing a second stop; since double consonants are blocked by a strong constraint, only two small parts of the partial ordering structure are available). Furthermore, there is a partial ordering relation between each form in the base, unembellished partial ordering structure and the corresponding form in the partial ordering structure corresponding to each embellishment, based on Constraint C1 (↓X). Consistent with the overall analysis above, there are no violations of the partial ordering involving this constraint within the forms containing at least one stop. To avoid cluttering the figure, these relations are represented schematically with a single arrow to the unembellished set of forms from each of the sets of embellished forms.

The first point to note about Figure 2 is that there is not a single case in which a target form occurs and a form violating a subset of the constraints violated by a target form does not occur. Furthermore, of the 164 direct relationships, only seven pairs exhibit a violation in that their observed rates of occurrence do not match the relationship predicted by constraints C1-C4. This remarkable pattern of consistency strongly indicates the utility of these simple graded constraints in capturing differences in observed occurrence rates of rhymes of different types. However, the violations do require attention since they indicate that there may be a few meaningful deviations from the predictions from just these four constraints. As mentioned above, possible violations were tested statistically, and the results for all tests are reported in Table 3. The strongest and most clearly statistically reliable violation occurs with /Vts/ < /Vks/: binomial $p < 0.001$, $t(4) = -5.82$, $p < 0.01$. /Vts/ also occurs less frequently than /Vps/, and although this violation is not significant, it appears that /Vts/ is rarer than the theory says it should be. There are several other violations of C4, ↓NC: /Vnt/ < /VNk/ is significant by the binomial but not the t-test. /Vnt/ occurs less frequently than /Vmp/, though this is not significant. /Vg/ occurs more often than /Vd/, though this is not significant. /VVlt/ < /VVld/ and /Vld/ < /VVld/ are both significant by several tests. Thus /VVld/ appears to be stronger than two of its superordinates, producing one violation each of C2, ↓VV and C3 ↓VO. There is a remarkable consistence among the results obtained using the ful set of long vowels and the reduced set. The violations found with the full set all still hold with the reduced set, and there are only two additional violations: VVt and VVd both occur more frequently than their short vowel counterparts, Vt and Vd. The differences are not reliable, but they pattern with VVld > Vld and other violations of ↓VV that we will consider later. Possible reasons for these violations will be considered after we examine all of the relevant cases.

We now consider possible stop rhymes whose codas contain more than one embellishment. The GCT asserts that these forms would be subordinate to both of the two corresponding singly embellished forms. For example, /Vnts/ is subordinate to /Vnt/ and /Vts/. Only the following forms are attested: /Vkst/, /Vlts/, /Vlkt/, /Vlpt/, /Vmps/, /Vmpt/, /Vnts/, and /VNks/. All are quite rare (several are associated with one or two clearly foreign words). In any case, consistent with the earlier analysis, all the singly-embellished immediate superordinates of all these forms are found in the corpus and all these singly embellished forms occur more frequently than the doubly embellished forms. No triply embellished forms occur in the corpus. Thus the pattern of occurrence (or lack thereof) of multiply embellished forms is completely consistent with the partial ordering implied by C1-C4.

Although not included in Figure 2, the forms containing the additional embellishment f_ preceding a coda stop consonant are also consistent with constraints C1-C4. Considering forms with f_, only /Vft/ and /VVft/ actually occur. The former should occur more frequently than the latter, and it does; furthermore, each should and does occur less frequently than the corresponding unembellished form, because the embellishment counts as a violation of C1 (↓X). Additionally, each should occur more frequently than voiced variants or variants involving non-coronal stops, and indeed this is the case: none of these forms occur at all in the corpus. One can also compare the embellishment f_ with s_ in terms of ↓NC, but we postpone this issue until the fuller discussion of fricatives below.

As a final note in this section, it is worth pointing out that the restrictions on the occurrence of a second stop can be seen as reflecting the influences of C2-C4. The only occurring instances of such an embellishment involve _t, the least constrained possible case. No forms subordinate to this in terms of coronality or voicing are found – that is there are no voiced forms of this embellishment (*gd, *bd) and no non-coronal forms (*pk, *kp, *tk, *tp). Thus this embellishment is consistent with the influences of ↓NC and ↓OV. Finally, there are no cases in which this embellishment occurs with a long vowel, fully consistent with ↓VV.

Rhymes Containing Affricates

Affricate rhymes include the affricates /C/ or /J/, appearing in 218 words in the corpus. These phonemes can be analyzed as /tS/ and /dZ/, a stop-fricative pair. The fricative can then be viewed as another form of post-stop embellishment. This makes /tS/ subordinate to /t/ and /dZ/ subordinate to /d/. With this construal, as shown in Figure 3, the affricates can be incorporated in the above analysis of forms with post-vocalic stops without introducing any inconsistencies (we continue to postpone consideration of place-related constraints among the fricatives). Also shown in Figure 3 is the relationship of further embellishments of the affricates by either a preceding liquid or nasal. These embellished forms also satisfy the GCT, in their subordination to the base affricates and also to the embellished forms of the base stops (e.g. /VlC/ is subordinate to both /VC/ and /Vlt/).

To summarize the discussion thus far: The facts about the existence and relative occurrence rate of rhyme types including at least one stop are highly consistent with constraints C1-C4. There is not a single form attested in the corpus for which a form violating ANY subset of the constraints it violates is not also attested. Furthermore, there are only a handful of violations of the predicted relationships among the relative occurrence rates arising from these four constraints.

Rhymes Containing Fricatives

The fricative rhymes are those that contain a fricative (f, v, s, z, T, D, S, or Z) by itself or with an optional pre-fricative liquid or nasal embellishment, or a post fricative stop:

[V|VV] (l) (N) **F** (S)

From the present point of view, /Vst/, /VVst/, /Vft/ and /VVft/ appear again as fricatives with a post-fricative stop, while earlier they were classified as a stop rhyme with a pre-stop /s/ of /f/. Here again we ask how well C1-C4 account for the existence and occurrence rate relations among the various forms. The fricatives include several different coronal variants (in addition to /s/ and /z/, there is /T/ and /D/ as well as /S/ and /Z/). The relatively infrequent use of /T/ and /D/ suggests a clear but still graded constraint against the dental place of articulation:

C5a: ↓DC: Disprefer Dental Coronals

The situation with the palatalized coronals is clearly more complex but a simple corresponding constraint can be formulated for consideration:

C5a: ↓PC: Disprefer Palatalized Coronals

Figure 4 shows the partial ordering relationships among fricatives, with no or one embellishment. The diagram of the relationships among the fricative rhyme types is shallower than that of the stops since the fricatives use fewer embellishments and are less common overall. However, the diagram is wider as well, since the fricatives use more places of articulation. The majority of the relationships obey the constraints, but there are more exceptions than there were in the stops.

One important pattern is a consistent reversal of the ↓VV constraint in the voiced fricatives. Individually, some of the violations are not significant, but collectively the pattern is highly consistent. /Vv/ < /VVv/, /Vz/ < /VVz/ , and /VD/ < /VVD/ are all clearly significant by several tests. It is also true that /VZ/ < /VVZ/ although none of the statistical tests are significant. Possibly related to the relative preference for long vowels with voiced fricatives, there is also one variation of ↓OV: /VVf/ < /VVv/, though this is only significant in one of the four tests. Though non-significant, one other violation of ↓VV, /VVT/ < /VT/ should also be noted. Given the overall rarity of /T/ and the small size of the difference, this might reflect sampling variability. One additional violation arises in comparisons involving the reduced set of long vowels. Here we find that VVs>Vs, again contra ↓VV. This violation is non-significant, but patterns with other violations of ↓VV in forms involving coronals.

There is also a set of violations which relate to the place constraints. A pair of violations of ↓NC that are significant by the binomial test where /Vls/ < /Vlf/ and /Vlz/ < /Vlv/. These violations are reminiscent of the ↓NC violation /Vts/</Vks/ and /Vps/. In all these cases a pair of nearly homorganic coronal gestures, the second of which is a fricative, is less common than a non-coronal rhyme type that otherwise shares the same constraint violations. There is also /Vz/ < /Vv/, contra ↓NC, and /Vs/ < /VS/, contra ↓PC, neither of which is significant.

There are a few additional cases to consider involving fricatives together with an additional element. First, there are the forms containing a /t/, /k/, or /p/ followed by an /s/, as well as a single case containing a /d/ before a /z/ (*adz*). Such forms are all less common than the corresponding immediate superordinate forms (/Vs/, /VVs/, and /Vz/) without the preceding stop, as expected by GCT, and no forms which these forms dominate occur at all in the corpus.

One could attempt to encompass the affricates under the present analysis as well, treating the /t/ in /tS/ the /d/ in /dZ/ as pre-fricative stop embellishments. Recall that we successfully treated the /S/ and /Z/ in these forms as embellishments of a stop /t/ or /d/ above. However, treating the affricates as embellishments of the corresponding fricatives does not work (while /VS/ > /VtS/, the effect is small, and the rest of the comparisons come out in the wrong direction: /VVdZ/ > /VVZ/, /VVtS/ > /VVS/, and /VVdZ/ > /VVZ/; similar problems arise comparing liquids, and the Z and S forms with nasals may not exist at all in some dialects). Although we will not pursue it further, these facts may be consistent with the observation that, relative to the corresponding fricatives, affricates are often of shorter duration (the abruptness recognized as a stop may be a byproduct of this shortening). This suggests a treatment in which the affricate is less of a violation of ↓X than the corresponding fricative, because it is a shortened version of it, rather than an embellishment of it. For consistency this would require that a single stop is less of a violation of ↓X than a single fricative; because stops and fricatives differ in place of articulation this is in general difficult to adjudicate, but it is not inconsistent with the data (e.g. /Vt/>/Vs/).

A separate issue that arises in treating the affricates as embellished stops is that it invites a comparison of /VC/ (treated as /VtS/) with /Vts/ and of /VJ/ (treated as /VdZ/) with /Vdz/. Such a comparison would reveal a clear violation of ↓PC: The non-palatalized forms are marginal at best while the palatalized forms are quite robust. Given the clear violation of ↓NC in the fact that /Vks/ > /Vts/, it may be that some special force is operating particularly against /ts/.

In summary, our analysis of rhymes containing fricatives reveals a picture somewhat more complex than the one that holds among rhymes containing stops. First, there is clearly a graded tendency, present for all four English voiced fricatives, for the fricative to occur more often with a long vowel than a short vowel, contra the general preference for short vowels. Second, while there is clearly a strong yet still graded constraint against dental coronal fricatives relative to other types, evidence for a constraint against either palatal coronal fricatives /S/ and /Z/ or labiodental fricatives /f/ and /v/ relative to alveolar coronals /s/ and /z/ is somewhat inconsistent.

Rhymes without Obstruents

We now consider rhymes whose codas contain only liquids or nasals or no consonants at all, conforming to the following template:

[V|VV] (l) (N)

Rhymes consistent with this template occur in 1083 words in our corpus. These forms are by definition all voiced. Figure 5 shows the relationships among the rhymes that use these codas, and places them in the context of the coda, /V/ conspicuously absent from stressed monomorphemic monosyllables. As the figure indicates, the constraints ↓X and ↓VV would lead us to predict that /V/ would be the most common rhyme type in the language, but this is far from the case in fact. Indeed, these rhymes do not occur at all in stressed monomorphemic monosyllables in English. While we previously noted that a hard constraint against such forms could be specified, we will discuss below how they might reflect a relatively extreme violation of a graded constraint against rhymes that are overall too short.

In stark contrast to the non-occurrence of /V/ rhymes is the very high rate with which rhymes containing only a long vowel /VV/ occur. Such rhymes are by far the most frequently used in this subset of the language. With 36.4 occurrences per long vowel they can be described as overloaded, producing many cases in which the same phonology is used several times with completely different denotations. This is especially prominent in the case of the vowel /i:/, where /si:/ corresponds to *see, sea*, /bi:/ corresponds to *be*, *bee*, /ti:/ corresponds to *tee*, *tea*, /ki:/ corresponds to *key, quai*, /pi:/ corresponds to *pea, pee*, /fli:/ corresponds to *flea, flee*, etc.. There are similar cases with other vowels as well, e.g. *by, buy; dye, die, high, hi; lie, lie; pie, pi; too, two, to*; *so, sew*; *way, weigh;* etc. Such forms are also often used in first names and in the names of letters of the alphabet In contrast to this the most frequent rhyme type involving a consonant, /Vt/, occurs only 22.6 times per vowel, and only three cases involving a consonant (/Vt/, /Vk/, and /VVt/) have as much as half the average occurrence rate as /VV/.

The clear preference for /VV/ over /V/ is mirrored in other violations of ↓VV. We find /Vl/ < /VVl/ and /Vn/ < /VVn/, although these violations are relatively small and only show up as reliable in binomial tests. The lack of reliability in the t-tests for these cases may suggest some interactions with specific vowel contexts; as we discuss below this may help explain other violations of ↓VV involving /l/. Interestingly, however, ↓VV is honored in those forms containing non-coronal nasals; so much so in the case of velar nasals that /VVN/ is absent completely from the language. Thus, ↓VV is clearly weak at best in some contexts while strongly in evidence in others. We may also note that within the forms with nasals, there is also evidence consistent with ↓NC: Both the /V/ and /VV/, the rhymes with coronal /n/ are favored over those involving either labial /m/ or velar /N/.

Discussion

We began with an observation that the occurrence rates of rhymes in English appeared to be systematically structured by a set of constraints each applying its influence in a cumulative way. The GCT formalized this observation and in the above analysis we used a small set of graded constraints, C1-C5, to explain the relative occurrence rates among a variety of rhyme-types. Overall, the GCT faired well. Of the 363 direct pairwise comparisons among forms allowable under the stipulated hard constraints, only 20 were in violation of the predicted relationship (23 when the long vowels were restricted to a reduced set of five), and of those only 13 were found to be statistically significant violations of the GCT in one or more of the statistical tests. A much smaller number were found to be statistically reliable by both the binomial and t-tests. The pattern of violations is clearly meaningful, however. The constraints ↓X and ↓VO appear to be extremely robust; ↓X is violated in only one instance (/VN/ < /VNk/) and ↓VO is violated only in two (/Vlt/</Vld/ and /VVf/</VVv/). The constraints ↓VV and ↓NC each exhibit interesting patterns which we now consider in turn.

Violations of ↓VV. There are two sets of violations of ↓VV: One set involves forms ending in a voiced fricative: /Vv/ < /VVv/, /VZ/ < /VVZ/, /Vz/ < /VVz/, and /VD/ < /VVD. This class of violations shares the property of having a short, low-energy coda (Kent & Reed, 1992). The remaining violations ↓VV occur with coronal consonants. In addition to the voiced coronal fricative cases already listed, we find: /VT/ < /VVT/; /Vn/ < /VVn/; /Vl/ < /VVl/; and /Vld/ < /VVld/. Furthermore, in the reduced vowel set we find /Vt/ < /VVt/; /Vd/ < /VVd/; and /Vs/ < /VVs/. Only a subset of these violations is statistically reliable even in the binomial test, but the pattern is striking in that, outside the cases with voiced fricatives, there are no violations of ↓VV that involve non-coronals. In the case of the forms involving /l/, the violations may result from the fact that liquids interact with the preceding vowel, causing a lengthening or even dipthongization of what would otherwise be a short vowel monothong. Interestingly, an unvoiced stop in the coda can cause a shortening of the vowel (Reed and Kent, 2002) which would work against the liquid lengthening. This could act as a break on migration of /Vlt/ toward /VVlt/, thus explaining why /Vlt/ remains much commoner than /VVlt/ and also explaining one of the two violations of ↓VO, namely the fact that /VVld/ is more common than /VVlt/. However, the rest of the coronal violations of ↓VV do not appear to be amenable to this interpretation. It would thus appear that ↓VV is not honored in forms with simple coronal codas or voiced fricatives, but is quite strong in other cases.

Violations of ↓NC. Another class of violations includes three slight reversals of ↓NC, /Vd/</Vg/, /Vz/</Vv/, and /Vs/</VS/, along with several cases in which, while not reversed, the constraint does not appear to be operating very strongly. Cases in point include /Vt/ vs. /Vk/, /Vs/ vs. /Vf/, and /Vn/ vs. /Vm/, all of which are very close in their

rate of occurrence even though the second form in each pair violates ↓NC. The /Vd/</Vg/ violation may be attributable to exclusion of some forms that may well deserve to be considered: /Vd/ occurs in 14 irregular past tenses, including very frequent *did, said, had* and 11 other cases (this excludes 5 no-change cases like *spread*). These cases would reverse the violation if they were counted, but the resulting /Vd/ advantage would still be very slight. It would thus appear that the ↓NC constraint has only a weak impact in cases where the vowel is short and there is a single coda consonant.

The final class of violations involves a comparison between some of the forms that contain two coronals (/Vts/, /Vls/, /Vnt/ and /Vlz/) and a form that contains a non-coronal, but is otherwise the same: /Vts/ < /Vks/; /Vls/ < /Vlf/, /Vlz/ < /Vlv/; and /Vnt/ < /VNk/ , /Vmp/. Some of these cases might be attributable, at least in part, to restrictions placed on the corpus, but this cannot explain the pattern fully. If regular plural and third person inflections were included, the relationship between /Vlz/ and /Vlv/ would certainly reverse. There are six irregular past tenses forms with the /Vnt/ rhyme type, but including these would not reverse the violations involving /Vnt/ vs. /VNk/ and /Vmp/. Counting inflected forms would not be likely to reverse the greater prevalence of /Vks/ over /Vts/: both forms would gain many additions due to inflection: /Vt/->/Vts/; /Vk/->/Vks/. However, it should be noted that /ts/ is very prevalent in contractions and colloquialisms such as *it's, that's, let's, nuts, guts,* etc. The relative prevalence of /Vks/ may be partly due to words from the Germanic base of English that contained a palatal fricative, a phoneme no longer used in English, some cases of which may have migrated to /ks/. It is also possible that certain types of sequential coronal articulations are relatively difficult, leading them to be dispreferred. Consider the sequence of tongue positions required to produce the sequence /ls/. For the /l/, the tongue must assume a concave posture, to allow both the alveolar closure at the tip and the lateral opening at the sides, while for the /s/, the tongue must assume a convex posture, with the tip curled downward to match the downsloping shape of the alveolar ridge to create the fricative gap needed to produce an /s/. Such a sequence may require a greater rate of tongue movement than, for example, /lf/, in which the /l/ gesture and the /f/ gesture are made with largely independent articulators. A similar account might be attempted to explain why /ks/ is relatively more prevalent than other forms involving a consonant followed by /s/; the tongue shape required for /k/ may be rather close to that needed for /s/, allowing a relatively easy transition between the two required positions. Of course, these are speculations; systematic evidence from careful studies of articulation would be required to support them. Furthermore, they cannot readily address the violations /Vnt/</VNk/,/Vmp/.

Several of the violations of ↓VV and ↓NC occur in forms that are otherwise quite low in constraint violations: many of the violations of ↓VV occur in forms ending in single coronals, and several of the violations or close calls for ↓NC occur in forms with short vowels and single consonants. These patterns presage some interactions among constraints that will emerge the next analysis.

Analysis 2: Parametric Analysis of Stop-Rhymes

In the preceding section we analyzed a wide selection of rhyme types according to a small set of constraints. In the following section we will narrow the range of rhyme types to those containing at least one stop, while attempting to give a more detailed parametric account of the rates of occurrence of these rhyme types. As a first step in this analysis we will consider the nature of the function D which relates the actual parameterized constraints to the predicted average rate of occurrence of each rhyme type. Earlier we considered two possible forms for D, a linear sum of negatively weighted constraints, thresholded at 0:

$$S_i = \left[\sum_j w_j C_j + \beta\right]^+, \; w_j < 0 \text{ for all } j$$

Or a multiplicative combination of constraints with weights in the range (0,1):

$$S_i = \alpha \prod_j w_j^{C_j}, \; 0 < w_j < 1 \text{ for all j.}$$

For the purposes of the non-parametric analysis it was not necessary to choose between these functions since both make the same ordinal predictions. However, as we shall now see, the two forms do make different quantitative predictions, as well as different predictions concerning whether certain forms will be found at all in the corpus.

Before we begin, some elaboration of our statement of constraints is required. In Analysis 1 the relationships between embellished forms and their unembellished counterparts were all covered by the blanket constraint ↓x, which asserts that embellishments are not preferred. In a parametric analysis, however, there is no reason to suspect that the actual weight of the constraint against each embellishment will be the same. To allow for the possibility that each has a different weight, we specify a separate constraint for each of the possible embellishments separately:

C1a:	↓NAS	An added nasal segment is not preferred
C1b:	↓LIQ	An added liquid segment is not preferred
C1c:	↓preS	An added pre-stop fricative segment is not preferred
C1d:	↓postS	An added post-stop fricative segment is not preferred
C1e:	↓postT	An added post-stop stop segment is not preferred

Similarly, the strength of the constraint against a non-coronal place of articulation may differ between the labial and velar places. So constraint C4 will be treated as:

C4a:	↓NC-lab	Labial place of articulation is not preferred
C4b:	↓NC-vel	Velar place of articulation is not preferred.

The two candidate D functions above were fit to occurrence rates of all 68 of the stop-rhymes consistent with the hard constraints identified previously that could be formed by assigning a long or a short vowel, a voicing value, and a place of articulation, and either no embellishment or one of the five embellishments itemized above. For each function, values for β and *w* for each of constraints C1a-e, C2 (↓VV), C3 (↓VO) and C4a-b were found that minimized the sum of the log of the improbabilities of the obtained occurrence rates of all of the candidate stop-rhymes. Forty of the rhymes actually occur in words in English, 27 do not, and there is one, VVNk, occurring in the onomatopoetic word *oink*, which we treat as borderline. The Praxis algorithm was used for minimization (Gegenfurtner, 1992). The goal of the minimization was to find the set of parameter values that yielded the smallest value of the probability of the data given the model. The algorithm is by no means guaranteed to converge to a global optimum; the result returned can depend on the starting point and it can be thrown off by large outliers. Considerable care was taken to ensure that the results were, if not globally optimal, at least near-optimal. Each reported result is the best achieved over a wide range of different starting values.

Overall statistics reflecting the goodness of fit of the two models to the data are shown in the first two rows of Table 4. The first three measures are: -log(p), the negative log of the probability of the data given the model[3]; r^2, the square of the correlation coefficient, which corresponds to the proportion of variance among the occurrence rates that the model can account for; and N(sig), the number of significant deviations from the model's predictions. We have been quite liberal in scoring deviations as significant ($p<.1$, two tailed) in an effort to be sure to identify potential shortcomings of the model. Note that with the criterion, we would expect to find about 7 significant deviations just by chance. When we discuss individual deviations we mention its p-value so that the reader can decide how much weight to place on each specific deviation. In addition, we present an analysis of the model's success in capturing which forms do and do not occur in the corpus. For this analysis, we define a 'hit' as a case where a form occurs in the corpus and the model predicts its occurrence. A miss is a case where a form occurs in the corpus and the model fails to predict it. A false positive is a case where the form does not occur but the model predicts that it should occur; and a correct rejection is a case where the form does not occur and the model predicts that it should not. We provide a separate column for the borderline rhyme type VVNk.

---

[3]The negative log of the probability of the data is the sum of the negative log of the probability of each data point, $\Sigma_i -\log(p_i)$. p_i is the probability of data point *i* (in our case, the number of occurrences of rhyme type *i*) given the model and parameters. This summed quantity is a measure often used to compare fits of alternative models and alternative sets of parameters, and so we report this variable in our results. We found empirically, however, that the algorithm tended to find better solutions as measured by $\Sigma_i -\log(p_i)$ when it optimized a slightly different measure, namely $\Sigma_i -\log(cp_i)$. The variable cp_i is the cumulative probability of observing a number of occurrences of rhyme type *i* that is the same amount or more deviant from the predicted value. For comparability to other work, we report $\Sigma_i -\log(p_i)$ even though $\Sigma_i -\log(cp_i)$ was actually the variable optimized.

The threshold-linear model fares much better than the multiplicative model on several fronts. The probability of the data is lower (Σ -log(p) larger) under the multiplicative model. Also, the r^2 or proportion of variance explained is greater for the threshold-linear model. We also find that 36 of the multiplicative model's predictions significantly differ from the actual data, while only 23 of the additive model's predictions are significantly different from the actual occurrence rates. Furthermore, the multiplicative model does not fare well in explaining which forms do and do not occur. A part of the problem is that the model assigns a probability greater than 0 to all of the forms. In some cases the predicted rates are low enough that the non-occurrence of the form is not surprising, but in 16 cases the predicted rates for non-occurring forms are high enough that the absence of such forms is unlikely under the model ($p < .05$). On the other hand, the additive model correctly predicts the non-occurrence of 23 of the 27 forms that do not in fact occur, while predicting occurrence for all but 2 of the 40 forms that do occur.

The underlying problem with the multiplicative model is indicated by the fact that the pattern of deviations of the predicted from the obtained results clearly has a systematic curvilinear form (Figure 6). Indeed, the actual occurrence rates are better correlated with the logarithmic function shown on the figure (log((x+1)/2), truncated at 0, r^2= .850) than they are by the predictions themselves (r^2 = .826), and the logarithmically transformed predictions capture a significant portion of the variation in the occurrence rates that is not explained by the untransformed predictions (p<.00001). Since taking the log of the multiplicative model's predictions and truncating at 0 turns it into a linear threshold model, it seems clear that the latter is the superior account.

We now consider the threshold-linear model. Table 5 shows the values of the parameters. The r^2 statistic is fairly high (.850), and it only missed 2 cases of forms that actually occur, while making four false positives. Table 6 shows the details for each of these errors, as well as performance on several other forms that will become relevant later. Of the false positives, three involve non-coronals in conjunction with nasals. The model also over-predicts VVNk, which occurs in the marginal form *oink*. Of the false negatives, the model fails to predict the occurrence of Vlb and VVld. The case of Vlb seems marginal, given that only two words with Vlb are listed in the corpus, and only one of these (bulb) is in common use. VVld is a more serious failure; as previously discussed under Analysis 1, this form is more common than either VVlt or Vld, thus violating the expected partial ordering on two counts.

While these results clearly indicate some measure of success for the threshold linear model, there is, equally clearly, room for further improvement. In addition to the misses and false alarms discussed above, 15% of the variance among occurrence rates remains unexplained, and 23 of this model's predictions are significantly different from the actual occurrence rates found in the data. An analysis not shown in the table using the reduced set of five long vowels instead of the full set of 10 produced similar results: The r^2 was .871, and 22 predictions deviated significantly from the actual occurrence rates.

The remainder of this section considers whether the inclusion of interaction terms can increase the overall adequacy of the threshold linear model's account for the data. This analysis is by its nature exploratory, and the results should be treated as suggestive rather than definitive, since several variants on the formulation offered below produce quite similar results.

Initial explorations revealed that each of several alternative single interaction terms could produce an improvement in the fit, but still left many unexplained significant deviations from predictions. In an effort to understand the pattern of data, we inspected the occurrence rates of the various forms, shown in Figure 7. Consider first those forms involving either no embellishment or a nasal embellishment (the two left-most sub-panels of the upper and lower panels of the figure). It can be seen that the penalty against non-coronals is weak or even reversed in several circumstances: When there is a short vowel and no embellishment or when there is short vowel, a nasal embellishment, and the stop is unvoiced. On the other hand, when the vowel is long, or when there is a nasal together with a voiced stop, a penalty against non-coronals becomes very clear. There is also a moderate, but clear, penalty against non-coronals with both the /s/ and /l/ embellishments. Finally, velars and labials appear to behave somewhat differently: While the labials show at least some relative disadvantage compared to coronals nearly everywhere, the velars actually show an advantage over coronals in several of the simpler cases, but then show a precipitous drop in several other cases. With these considerations in mind we formulated the following specific set of interaction terms:

C7a: ↓NC x VV
C7b: ↓NC x Nas x (Vo or VV)
C7c: ↓NC x (l_ or s_)
C7d: ↓Vel x Vo x (VV or E)

When taken individually the constraints may each seem arbitrary, but they have a common theme, namely that the constraint against non-coronals is amplified when it occurs in conjunction with other constraints. C7a captures the amplification of the penalty against coronals when the vowel is long. C7b captures the amplification of ↓NC when it is combined with a Nasal and a Voiced Stop or Long Vowel. C7c captures the moderate enhancement of ↓NC in conjunction with the pre-stop embellishments l_ or s_. C7d captures the observation that forms involving velars are very rare when they appear in a voiced coda in conjunction with a long vowel or any embellishment.

The four conjunctive constraints, together with the β parameter and the 9 individual constraints considered earlier, were entered into the fitting procedure, where they were used to account for the data from the full set of 68 rhyme types. The end result was an increase in the overall r^2 to .95, and a reduction in the number of significant discrepancies between predicted and actual occurrence rates from 23 to 16. The model failed to predict the occurrence of 5 of the 40 rhymes that actually occur in English, but four of the five appear to be marginal cases (see Table 6); it predicted only one rhyme to occur that does not occur.

The predicted pattern based on the model is shown in Figure 8, along with the actual data previously shown in Figure 7; the discrepancy between the predicted and actual data is indicated by a solid error bar connecting the predicted and observed value for each point. The dashed error bars show the 90% confidence interval surrounding each prediction, so that cases where an actual value falls outside the dashed error bars represent at least a marginally significant deviation ($p<.05$ for a deviation that large in the given direction. In this section we report the one-tailed p value, but note that the probability of a deviation as large or larger in either direction, the conventional measure of statistical significance is .1). Several highly significant deviations occur for cases of two successive coronals, where the stop is unvoiced: Vnt ($p=.0027$), VVnt, Vts, and VVts (all p's<.001). Though not as marked, there is also a slight shortfall for Vlt (Pp=.030) and VVlt ($p=.013$). Interestingly the model also predicts that Vzd should sometimes occur; and it never does, thus also violating the model's predictions for occurrence rates of combinations of coronal segments. The remaining discrepancies are somewhat difficult to see, since they occur among relatively rare forms. However, the model predicts that some such forms should occur with a probability of 0, and under the model's assumptions, even one occurrence should therefore be impossible. As before, the most serious case in point is VVld, which occurs in 19 words, for an average rate of 1.9 per long vowel. The remaining cases deserve mention but may be more marginal in nature: VVsk and VVsp occur in several words, but all these cases occur in words spelled with 'a' like 'ask' and 'clasp'. These all receive a long vowel pronunciation (to rhyme with 'shah') in the Southern British dialect used in CELEX. This vowel is perhaps one of the 'shortest' of the long vowels, and the ordinary short vowel sound as in 'cat' is used for this vowel in many other dialects. Vlb only occurs in two words, only one of which appears to be in common use (bulb); VVks and VVps each occur in two words (hoax, coax; traipse, corpse). Finally, there are a number of additional discrepancies: Vkt ($p=.048$), Vps ($p=.01$), VVk ($p=.002$). Although the last two are quite reliable, the pattern of the actual data in these cases is not strikingly different from the pattern of the predictions.

Discussion

Analysis 2 described the results from several parameterized versions of the GCT. Using a simple linear model, whose weights were discovered by a regression, the model was able to capture 85% of the variance in the average rates of occurrence of the different rhyme types. It was able to reproduce the full range of occurrence rates from the most common forms to the least common ones. Moreover, it was able to account for the set of forms that do exist and the set that do not, with only 1 significant miss and 4 false alarms. Moreover, by adding terms that allowed the constraint against non-coronals to be compounded when combined with other constraints, the model was able to account for over 95% of the variance in the rates of occurrence of forms, with only a few large and reliable deviations. Thus a small set of graded constraints plus a few interactions provide a parsimonious account for most of the variation among the occurrence rates of English monomorphemic, monosyllabic word lemmas.

With the results of this analysis in hand, let us consider the interesting restriction on the occurrence of coda consonants with the vowels /aU/ and /OI/. Recall that these vowels have been identified as special in that the only appear with coronal consonants (Hammond, 1999). Hammond’s explanation for this posited that these vowels have more weight in a syllable than other long vowels or dyphthongs, while coronals have little or none. Within the GCT, these vowels could be treated as being especially long and therefore in greater violation of ↓VV. With the linear threshold model, the restriction to coronals might then arise simply because the added penalty for having a very long vowel (denoted here as VVV) pushes all non-coronal forms below threshold. However, a problem for this arises from the fact that such a constraint should also push forms ending in /nd/ and /nt/ below 0 as well, yet such forms do occur (point, joint, etc; round, sound, bound, etc). A way of explaining the occurrence of VVVnt and VVVnd in the absence of VVVk and VVVp is to appeal to constraint interactions: specifically, we suggest that the enhancement of the constraint against long vowels in the case of /aU/ and /OI/ may be accompanied by a further enhancement of the interaction ↓NC x VV.

Analysis 3: Rhyme Types in Words of Different Token Frequencies

While generally supportive of the Graded Constraint Theory, the results of Analyses 1 and 2 revealed some qualifications. In Analysis 1 we saw that the preference for coronals, ↓NC, became quite weak among forms that were otherwise quite simple, and this is confirmed by the pattern of interaction of ↓NC with other constraints seen in Analysis 2. We also saw in Analysis 1 that the constraint against long vowels, ↓VV was reversed or very weak in forms that were otherwise quite simple, namely those with single coronal consonants in the coda.

Before we consider possible interpretations of these effects, we consider evidence from a further set of analyses indicating that in fact, ↓NC and ↓VV do operate even in very simple forms. More generally, we will show in this section that these two constraints operate more strongly in forms of higher token frequency than lower token frequency.

We took three approaches to investigating this issue. The first was just to compute the correlation between the frequency of each word in our corpus and the predicted rate of occurrence of the word’s rhyme type, based on the word’s constraint violation score. The idea is that words of higher frequencies should generally be associated with lower constraint violation scores. This analysis was applied to all words with rhymes of the types considered in analysis 2, using log-transformed frequencies (log(frequency+1)). It did yield a significant correlation, as predicted, between the log frequency and the constraint violation score, although the magnitude of the correlation is relatively low (r= .0713, $p < .001$). This finding is consistent with earlier evidence that the highest frequency words tend to be found in the most densely packed lexical neighborhoods (Landauer & Streeter, 1973; Frauenfelder, Baayen, Hellwig, & Schreuder, 1993).

The second approach we took was to look at the number of forms of each of several types falling at or above each of several frequency cut-offs. This analysis is presented in Figure 9. Note that the data points at the left end of each curve include the full set of words of the indicated type that are found in the corpus, and match the numbers in Table 2. The analysis is applied to the VV rhyme type and to six sets of three rhyme types consisting of a long or short vowel and either an unvoiced stop, a voiced stop, or a nasal. These include the most frequent types overall, allowing for greater stability of any trends, and they encompass the constraints ↓VV, ↓Vo, ↓NC and their interactions. The figure indicates that the pattern of differences in relative occurrence rates of different rhyme types generally hold up at all frequency cut-offs, and some of the reversals disappear when a higher cutoff is used. For example, /Vg/ occurs slightly more frequently than /Vd/ in the total corpus, contra ↓NC, but at higher cut-offs the constraint is honored, with /Vd/ occurring in more words of relatively high token frequency than /Vg/.

What is not so clearly visible in the figure is a general tendency for both ↓NC and ↓VV to apply most strongly to words in the higher frequency ranges. We first consider ↓NC, focusing on the fraction of words containing coronal codas within different frequency bands, holding constant other characteristics of the rhyme type. For example, consider forms ending in a short vowel followed by a single unvoiced stop. We can ask whether the fraction of such forms that end in coronals decreases as a function of frequency. We performed this analysis for the six classes of rhyme types containing consonants considered in Figure 9, i.e. those defined by the possible values of vowel length (short V vs. long VV) and final consonant type (unvoiced stop, represented by *u*, voiced stop represented by *v*, nasal, represented by *n*). Linear regression analysis revealed a positive correlation between word frequency and the proportion of forms in each class containing the coronal consonant in all six cases. Five of the correlations were significant (*Vu*: r=0.6552, p=0.028; *Vv*: r=0.8341, p=0.020; *Vn*: r=0.7197, p=0.014; *VVv*: r=0.8648, p=0.013; *VVn*: r=0.8509, p=.016; all tests 1-tailed), with the sixth just missing this level of significance (*VVu*: r=0.6044, p=0.056). On visual inspection (Figure 10), two of the curves are quite different from the other four, while the other four were quite similar. The two outliers were the curves for *VVv* and *VVn*, cases in which there is a very strong, although not absolute, preference for coronals in the overall corpus. In both of these cases, there are no forms at all ending in non-coronals in words with frequencies above 100; although the proportion with coronals does drop, it remains high across all frequency bands. In the remaining cases, the proportion coronal never reaches an extreme level, but the preference for coronals is very clear in the higher frequency bands and drops to a completely neutral value for words of the lowest frequencies. Because the curves are all similar, we have combined the data across all four cases to give a clear indication of the overall relation between proportion coronal and word frequency. Variability is relatively high in higher frequency bands, but this is expected due to the relatively small number of words in each rhyme class in the highest bands; for example, in the highest frequency range the number of forms of each of the four classes contributing to the average is as follows: *Vu*, 11; *VVu*, 10 *Vn*, 9; *Vv*, 0.

Finally, we consider the influence of ↓VV as a function of frequency, when the form ends in a simple coronal coda, either /d/, /t/, or /n/. Here, if the constraint is operating, we expect the rhyme types containing short vowels to be used more often than the rhyme types containing long vowels. Since our analysis considers 5 short vowels and 10 long vowels, the proportion of short vowel rhymes should be greater than .33. In fact the overall ratio, regardless of frequency, for these three coda types combined is only .355, indicating the overall weakness of the constraint among such forms. However, when the data is broken down by token frequency, a striking pattern emerges (Figure 11): The proportion of forms that involve short vowels is very high, .708, for words with frequencies greater than 1,000, and then drops down to values insignificantly different from .333 for all other cases. The proportion with short vowels of the highest frequency words is significantly greater than .333 ($p < .00005$), and significantly different from the proportion with short vowels in each of the other frequency bins (largest p=.0034). None of the other proportions differ reliably from each other or from .333, although the difference in proportion coronal between the two lowest frequency bins (0 vs. 1-9) reaches significance in a one-tailed binomial test (z=1.889, p=.0294).

In sum, though the details differ for ↓NC and ↓VV, it appears that both constraints are honored among otherwise very simple forms, at least among words of high frequency. The implications of these findings for the GCT are considered below.

General Discussion

The graded constraint theory provides a framework in which to account for the rates of occurrence of rhymes of different types in English. The theory is supported by two main observations: First, there is massive variation in how often different rhymes are utilized in the language --- variation that is unexplained by theories that focus solely on binary distinctions specifying what can occur and what cannot occur. Second, this variation is highly structured, suggesting the presence of systematic graded constraints. The theory captures this structure by positing a set of such constraints, which when violated reduce a form's underlying strength or attractiveness, and through this affect its underlying rate. The constraints can be summarized succinctly: It is better for the vowel in a rhyme to be short, and for the coda to be simple, unvoiced, and coronal.

We presented a statistical treatment of the theory that allows us to test hypotheses about the underlying rates in the face of random factors. We found in our first analysis that the predicted partial ordering of occurrence rates of different rhyme types holds up empirically in all but a handful of cases. We then assessed the goodness of fit of several quantitative formulations of the theory to the pattern of the data, and learned that a multiplicative formulation of the theory systematically mis-predicts the data. A linear threshold formulation, while generally superior, still mis-predicts the average rates of occurrence of many rhyme types, over-predicting differences in rates of occurrence among simpler forms and under-predicting such differences among more complex forms. With the addition of interaction terms, many of these discrepancies could be resolved. The linear threshold model with interactions accounts for 95% of the variance in the relative rates of occurrence of rhymes of different types with only a handful of notable discrepancies. It also correctly identifies which forms actually occur and which forms do not, with a few relatively minor errors. The different types of constraints behave in slightly different ways. The constraints against embellishment and coda voicing appear to affect occurrence rates fairly uniformly, but we find a different pattern for vowel length and coronality: overall rates of occurrence of rhyme types are only affected by violation of these constraints when they occur in combination with each other or in conjunction with other constraints. Yet our final analysis revealed a role for coronality and vowel length even in otherwise very simple forms, among words of high token frequency.

With all these results in view, we suggest that the relationship between a form's constraint violation score and its observed occurrence rate may be somewhat more complex that we originally anticipated. The observed occurrence rate may reflect a vocabulary-wide tendency toward minimizing the overall degree of constraint, in which lexical items take rhymes of particular types according to a process that tends to minimizing the average degree of constraint violation that will be found in the word forms occurring in a large sample of the language. This average constraint violation score can be formalized as:

$$\langle cvs \rangle = (1/ \Sigma_w f_w)\ \Sigma_w f_w\ cvs_w$$

In words, the expected average constraint violation score is the token-frequency weighted sum of the constraint violation scores of the individual words indexed by *w*, divided by a normalizing term equal simply to the sum of all the words' token frequencies. Note that this average will be minimized if words with high token frequency receive the simplest rhymes, and if in general the available rhymes with the lowest constraint violation scores are used up before assigning rhymes with higher constraint violation scores. Such a policy may explain in part why forms of the simplest types are relatively fully and equally utilized, while those violating multiple constraints tend to be relatively infrequently used or never used at all.

Within an approach of this type, it seems likely to be necessary also to take into account how distinct different word forms are from each other. In the limit, without some pressure for different words to be distinct, every word form would collapse down to the same minimal sound, and this would, of course, drastically compromise successful communication. Thus it appears there must be a countervailing pressure helping to keep us from loosing phonological distinctions. Future work pursuing these ideas may lead to the formulation of a joint function of average constraint violation and overall distinctness of word forms, drawing inspiration from the effort by Lindbloom, McNeilage, and Studdert-Kennedy (1984) who proposed such a function to explain the origin of languages' combinatorial phonology. Further development of these ideas must be left to future research. For now we note one challenge that we see facing an effort of this sort: We suspect that in practice constraint violation score and distinctness may not be completely independent. For example, the addition of an embellishment could increase compression of phonetic material, resulting in some reduction in distinctness. Separating these factors may not be trivial.

It is worth acknowledging some of the many other avenues for further investigation within the effort to evaluate the explanatory utility of the GCT. For one thing, we have only attempted to treat the rhymes of English monosyllabic monomorphemic word lemmas, and within these we have mostly restricted the analysis to rhymes containing stops. The restriction to considering monomorphemic lemmas was due to the observation that certain rhyme types are found in regularly inflected word forms that do not occur in morphologically simple items (Burzio, 2000). For example, the regular past tense form 'beeped' uses a rhyme that contains a long vowel, non-labial stop, and post-stop embellishment, and this rhyme type is not used in any single morpheme forms. Similarly voiced stop-stop combinations, such as /bd/ in 'rubbed' are found in regularly inflected forms but not in morphologically simple items. It will be interesting in future work to include such items, and also to consider the inclusion of irregularly inflected items, which Burzio suggests should be viewed as phonologically regular. One suspicion we have is that both regular and irregular inflected forms will best be viewed as reflecting a relaxation of constraints operating in morphologically simple monosyllabic word lemmas, although the degree of relaxation may be greater in the former than the latter cases.

Extending the GCT to multisyllabic forms will be more complicated. Previous work in this area has shown that prosodic structure interacts with phonological constraints. That is, the distribution of phonological forms is not independent of stress or location in the word (Hammond, 1999). However, it is unclear if new constraints would be required to account for this larger set of forms, or if they could be accounted for by merely adjusting the weights on the constraint violations based on the prosodic status of the given rhyme. The analysis in Harris (1994) of word-internal consonant clusters suggests that they are generally more permissive, but otherwise subject to similar (and we would argue, still graded) constraints. For example, the sequence /mb/, unattested in word final position, no longer falls below threshold in such cases, but only if the preceding vowel is short (c.f. limber, climber, in which the /b/ is sounded after the short vowel in the former but after the long vowel in the latter).

Similarly, expanding this analysis to other languages is not a trivial matter. It is an open question to what extent the constraints discussed here would generalize to other languages, with a simple re-adjustment of the weights. It is important to be clear that the objective here was to provide evidence that the strategy of incorporating graded constraints into a theory has utility in terms of accounting for structure, and not to have identified a set of universal constraints applicable in all contexts and languages. It is likely, that the particular constraints in another language, or indeed in another portion of English, are different. For instance, a preliminary analysis suggests that onsets even in English are governed by constraints different than rhymes. There may be good reasons for this, which could be explicated in detail in future work. Suffice it to say that onsets are subject to different demands due to their role in communication and their articulatory contexts which also differ from that of rhymes, and these differences might provide impetus for a different set of constraints. Similarly, different languages use different sets of phoneme contrasts, syllable requirements, and inflectional patterns all of which interact with the phonological constraints which might manifest themselves. However, it is important to note that some of the constraints identified have been shown to have a reliable cross-linguistic impact on the acoustical realization of forms (Delattre, 1962).

We have previously noted that our work builds on two important earlier characterizations of English sound structure. Harris' (1994) analysis greatly clarifies the overall structure of English word forms and offers succinct rule-like statements characterizing what types of rhymes do and do not occur. Hammond (1999) develops tools within an OT framework that characterize subtle distributional properties, including the restriction of coda consonants occurring with the superlong vowels /OI/ and /AU/. We suggest that the graded constraints and constraint interactions introduced here can capture much of the structure these authors characterized by rules or strictly ranked constraints; of course, we note as well that both works deal with phenomena arising in word forms outside the limited range we have considered here. As indicated above, it will be interesting to examine in future work how far GCT can be extended to address this broader range of phenomena. We also note that both works presage our emphasis on graded constraints: Harris does so by introducing what he called preferences to acknowledge difference in occurrence rates of different rhyme types, and Hammond does so by introducing additional gradations in vowel length beyond Harris's long/short distinction. We hope we have repaid the debt we owe to these investigators by pointing the way toward a fuller integration of graded constraints into the formal characterization of the sound structure of English.

It is very important to stress that, despite the emphasis on graded structure and graded constraints in this paper, we cannot conclude that all structure ought to be characterized as graded. As a matter of fact, in our analyses we assumed the operation of certain more absolute constraints. For example, we acknowledged the sonority sequencing principle as a categorical restriction limiting the range of possible forms under consideration. While some of the more absolute constraints might be encompassed simply as very strong constraints with the GCT, others such as sonority sequencing may turn out to be more fundamental.

The graded constraint theory shares many assumptions in common with Optimality Theory, a theory that emerged in part from earlier work on graded constraint satisfaction (Smolensky, 1986; Rumelhart, Smolensky, McClelland, and Hinton, 1986). Because we share this tradition, we were surprised to see OT move toward a system of strict constraint ranking. Our current work grows out of the intuition that strict ranking is not always maintained. We suspect that close scrutiny of occurrence rates in other sub-domains of phonology will reveal many other cases in which graded constraints combine to determine the relative goodness of a form. Indeed, we note that Burzio (2000) has argued that many interesting patterns seen in derivational morphology reflect graded constraint satisfaction processes, and has argued that a full account of such phenomena requires a relaxation of strict constraint ranking. Yet it does appear that many constraints operating on the phonological forms of words are categorical in nature. As just one simple example, it has been noted that in Spanish, only coronals are found in morpheme-final codas (Bybee, 2001). Given the hybrid and rapidly-changing nature of English, it might be surmised that it is an aberration among languages, and reflects a system in transition rather than steady state. While there may be some partial truth to this, it is also clear language is continually changing. that frequent forms lead in these changes which then spread gradually through the language (Bybee, 2002). If this is the case, there will always be graded variation in the rates of occurrence of different phonological forms, consistent with the graded constraint theory.

Author Notes

Supported by NSF grant BNS 01-XXXXX and NIH grant P50 MH 644445. We would especially like to acknowledge Luigi Burzio for helpful discussions and pointers to relevant literature. We would also like to acknowledge useful discussions with Joan Bybee, Matt Goldrick, Lori Holt, David Plaut, Paul Smolensky and members of the CMU PDP research group.

References

Baayen, R. H., Piepenbrock, R., & van Rijn, H. (1993). The CELEX Lexical Database (CD-ROM). Philadelphia, PA: Linguistic Data Consortium, University of Pennsylvania.

Boersma, P. (1998). *Functional phonology: Formalizing the interactions between articulatory and perceptual drives*. The Hague: Holland Academic Graphics.

Boersma, P., & Hayes, B. (2001). Empirical tests of the gradual learning algorithm. *Linguistic Inquiry, 32*, 45-86.

Burzio, L. (2000). Missing Players: Phonology and the Past-tense Debate. In K. K. Grohmann & C. Struijke (Eds.), *Working Papers in Linguistics* (Vol. 10): University of Maryland.

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

Bybee, J. (2002). Word frequency and context of use in the lexical diffusion of phonetically conditioned sound change. *Language Variation and Change, 14*, 261-290.

Coleman, J. S., & Pierrehumbert, J. (1997). *Stochastic phonological grammars and acceptability.* Paper presented at the Third meeting of the ACL special interest group in computational phonology, Somerset, NJ.

Delattre, P. (1962). Some factors of vowel duration and their cross linguistic validity. *Journal of the Acoustical Society of America, 34*, 1141-1143.

Diver, W. (1979). Phonology as Human Behavior. In R. R. (Ed.), *Psycholinguistic Research: Implications and Applications* (pp. 161-182). Hillsdale, N.Y.: Lawrence Erlbaum.

Frisch, S., Broe, M., & Pierrehumbert, J. (2004). Similarity avoidance and the OCP. *Natural Language and Linguistic Theory, 22*, 179-228.

Frisch, S., Large, N. R., & Pisoni, D. B. (2000). Perception of wordlikeness: Effects of segment probability and length on the processing of nonwords. *Journal of Memory and Language, 42*, 481-496.

Frauenfelder, U.H., Baayen, R.H., Hellwig, F.M., & Schreuder, R. (1993). Neighborhood density and frequency across languages and modalities. *Journal of Memory and Language, 32*, 781-804.

Fudge, E. C. (1969). Syllables. *Journal of Linguistics, 5*, 253-286.

Gegenfurtner, K. R. (1992). PRAXIS: Brent's algorithm for function minimization. *Behavioral Research Methods, Instruments, and Computers, 24*, 560-564.

Goldrick, M. (2004). Phonological features and phonotactic constraints in speech production. *Journal of Memory and Language, 51*, 586-603.

Hammond, M. (1999). *The phonology of English, a prosodic optimality-theoretic approach*. Oxford: Oxford University Press.

Harris, J. (1994). *English sound structure*. Oxford: Blackwell.

Kent, R. D., & Reed, C. (1992). *The Acoustical analysis of speech*. San Diego, CA: Singular Publishing.

Kessler, B., & Treiman, R. (1997). Syllable structure and the distribution of phonemes in English syllables. *Journal of Memory and Language, 37*, 295-311.

Landauer, T.K. & Streeter, L.A. (1973). Structural differences between common and reare words: Failure of equivalence assumptions for theories of word recognition. *Journal of Verbal Learning and Verbal Behavior*, *12*, 119-131.

Lindblom, B., MacNeilage, P., & Studdert-Kennedy, M. (1984). Self-organizing processes and the explanation of phonological universals. In Dahl (Ed.), *Explanations for Language Universals* (pp. 181-203). Berlin: Mouton.

Pierrehumbert, J. B., & Nair, R. (1995). Word games and syllable structure. *Language and Speech, 38*, 77-114.

Pitt, M. A., & McQueen, J. M. (1998). Is compensation for coarticulation mediated by the lexicon? *Journal of Memory and Language*, *39*, 347-370.

Prince, A., & Smolensky, P. (2004). *Optimality Theory: Constraint interaction in generative grammar*. Oxford: Blackwell.

Redford, M. A., Chen, C. C., & Miikkulainen, R. (2001). Constrained emergence of universals and the variation in syllable systems. *Language and Speech, 44*, 27-56.

Rumelhart, D. E., McClelland, J. L., & Group, t. P. R. (Eds.). (1986). *Parallel distributed processing: Explorations in the microstructure of cognition* (Vol. 1). Cambridge, MA: MIT Press.

Rumelhart, D. E., Smolensky, P., McClelland, J. L., & Hinton, G. E. (1986). Schemata and sequential thought processes in PDP models. In J. L. McClelland, D. E. Rumelhart &. the PDP Research Group (Eds.), *Parallel distributed processing: Explorations in the microstructure of cognition* (Vol. 2). Cambridge, MA: MIT Press.

Smolensky, P. (1986). Information processing in dynamical systems: Foundations of harmony theory. In D. E. Rumelhart, J. L. McClelland &. the PDP Research Group (Eds.), *Parallel distributed processing: Explorations in the microstructure of cognition* (Vol. 1). Cambridge, MA: MIT Press.

Treiman, R. (1988). Distributional constraints and syllable structure in English. *Journal of Phonetics, 16*, 221-229.

Treiman, R., & Kessler, B. (1995). In defense of an onset-rime syllable structure. *Language and Speech, 38*, 127-142.

Treiman, R., & Kessler, B. (1997). Syllable structure and the distribution of phonemes in English syllables. *Journal of Memory and Language, 37*, 295-311.

Treiman, R., Kessler, B., Knewasser, S., Tincoff, R., & Bowman, M. (2000). English speakers' sensitivity to phonotactic patterns. In J. B. Pierrehumbert (Ed.), *Papers in laboratory phonology V: Language acquisition and the lexicon* (pp. 269-282). Cambridge: Cambridge University Press.

Vitevitch, M. S., & Luce, P. A. (1998). When words compete: Levels of processing in perception of spoken words. *Psychological Science, 9*, 325-329.

Williams, R. J. (1986). The logic of activation functions. In D. E. Rumelhart J. L. McClelland, & the. PDP Research Group (Eds.), *Parallel distributed processing: Explorations in the microstructure of cognition* (Vol. 1). Cambridge, MA: MIT Press.

Table 1

Per Vowel Counts for Rhymes Containing Stop Consonants

							Voiced					
	Long						Short					
Coda	alone	n_	l_	z_	_z	_d	alone	n_	l_	z_	_z	_d
d	*12.2*	*2*	*1.9*				*14.4*	*6.8*	*1*			
g	*1.2*						*15.4*					
b	*2.4*						*13*		*0.4*			
							Unvoiced					
	Long						Short					
Coda	alone	n_	l_	s_	_s	_t	alone	n_	l_	s_	_s	_t
t	*19.2*	*2.6*	*1.1*	*3.9*	*0.2*		*22.6*	*8*	*4.6*	*7.8*	*0.2*	
k	*13.5*	*0.1*		*0.8*	*0.2*		*21.6*	*12*	*2.2*	*3.2*	*5.6*	*1.8*
p	*8.7*			*0.5*	*0.2*		*17*	*9.4*	*1.6*	*1.2*	*0.8*	*0.8*

Note: * = no coda, n_ = preceding homorganic nasal, l_ = preceding liquid, z/s_ preceding coronal fricative, _z/s subsequent coronal fricative, _d/t subsequent coronal stop.

Table 2

Per Vowel Counts for All Rhyme Types in Corpus

	Voiced											
	Long						Short					
Coda	alone	n_	l_	z_	_z	_d	alone	n_	l_	z_	_z	_d
*	36.4											
d	12.2	2	1.9				14.4	6.8	1		.2	
b	2.4						13		0.4			
g	1.2						15.4					
l	18.1						13.4					
n	16.8						15.8		0.2			
m	9.7						14.6		0.8			
N							9.8					
z	7.4						2.6	0.6				
v	5.9						2.8		1.4			
D	1.8						0.2					
Z	0.4											
dZ	3.1	0.7					6	2.8	0.4			

	Unvoiced											
	Long						Short					
Coda	alone	n_	l_	s_	_s	_t	alone	n_	l_	s_	_s	_t
t	19.2	2.6	1.1	3.9	0.2		22.6	8	4.6	7.8	0.2	
p	8.7			0.5	0.2		17	9.4	1.6	1.2	0.8	0.8
k	13.5	0.1		0.8	0.2		21.6	12	2.2	3.2	5.6	1.8
s	8.7	1.2	0.2			3.9	9.4	4.2	0.4			7.8
f	4.6					0.9	9.4	0.6	1.6			4.2
T	2.8						2.4	0.4	0.2			
S	0.8						11.6		0.4			
tS	4.6	0.8					9.6	5.4	1			

Note. * = no coda, n_ = preceding homorganic nasal, l_ = preceding liquid, z/s_ preceding coronal fricative, _z/s subsequent coronal fricative, _d/t subsequent coronal stop.

Table 3

Summary of Statistical Tests for Analysis 1

	Violation	Short Vowel	Long Vowel (Full)	Long Vow (Red.)	Test 1	Test 2	Test 1 binom	Test 1 t-test	Test 1 sig	Test 2 binom	Test 2 t-test	Test 2 sig
ix	VN-VNk	49 60			V-V		0.169	-0.81	0.462			
Stops	Vld-VVld	5	19	16	V-VF	V-VR	0.138	-0.65	0.527	0.013	-1.24	0.250
	VVlt-VVld		11 19	6 16	VF-VF	VR-VR	0.100	-1.24	0.247	0.026	-2.39	0.075
	Vnt-VNk	40 60			V-V		0.028	-1.6	0.180			
	Vnt-Vmp	40 47			V-V		0.260	-1.1	0.330			
	Vts-Vps	1 4			V-V		0.188	-0.885	0.426			
	Vts-Vks	1 28			V-V		0.000	-5.823	0.004			
	Vd-Vg	72 77			V-V		0.372	-0.30	0.781			
	Vt-VVt	113	192	126	V-VF	V-VR	0.925	1.28	0.222	0.218	-0.89	0.402
	Vd-VVd	72	122	81	V-VF	V-VR	0.883	1.18	0.380	0.259	-0.59	0.569
Fricatives	Vv-VVv	14	59	48	V-VF	V-VR	0.006	-1.271	0.226	0.000	-3.32	0.011
	Vz-VVz	13	74	58	V-VF	V-VR	0.000	-3.239	0.012	0.000	-1.69	0.115
	VD-VVD	1	18	18	V-VF		0.104	-1.657	0.121	0.000	-5.38	0.001
	VZ-VVZ	0	4	3	V-VF	V-VR	0.198	-1.255	0.231	0.125	-1.5	0.172
	VVf-VVv		46 59	30 48	VF-VF	VR-VR	0.121	-0.803	0.443	0.027	-1.25	0.279
	VT-VVT	12	28	15	V-VF	V-VR	0.397	-0.377	0.712	0.351	-0.44	0.672
	Vls-Vlf	2 8			V-V		0.055	-1.633	0.178			
	Vlz-Vlv	0 7			V-V		0.008	-1.871	0.135			
	Vz-Vv	13 14			V-V		0.500	-0.123	0.908			
	Vs-VS	47 58			V-V		0.165	-0.527	0.626			
	Vs-VVs	47	87	53	V-VF	V-VR	0.701	0.9	0.260	0.309	-0.42	0.685
Semi-Vow	Vl-VVl	67	181	117	V-VF	V-VR	0.021	-0.85	0.411	0.000	-1.69	0.130
	Vn-VVn	79	168	115	V-V	V-VR	0.352	-0.211	0.836	0.006	-1.63	0.143

Note. Columns show counts for the rhyme types with the short, long, and reduced long vowel set, the type of tests run (described below), and the results. **V-V**: short vowel to short vowel set comparison, t-test was paired with 4 degrees of freedom (df); **VF-V-F**: long vowel to long vowel comparison, full set, t-test was paired with 8 df; **VR-VR**: long vowel to long vowel comparison, reduced set, t-test was paired with 4 df; **V-VF**: short vowel to long vowel comparison, full set, t-test was unpaired with 13 df; **V-VR**: short vowel to long vowel comparison, reduced set, t-test was unpaired with 8 df.

Table 4

Goodness of Fit Measures for Several Quantitative Models

Model	-log(p)	r^2	N(sig)	hits	misses*	VVNk	FPs	CRs
Product	401.7	.826	36	40	0	+	>16	< 11
Threshold Linear	368.9	.850	23	38	1+1	+	4	23
TL+Ncx1	331.6	.931	22	35	1+4	~	3	24
TL+Ncx2	325.5	.940	19	35	1+4	-	2	25
TL+Ncx3	316.6	.946	15	35	1+4	-	2	25
TL+Ncx3+	308.2	.950	16	35	1+4	-	1	26

Note: The number of Significant Deviations in each of the last three rows does not count deviations for VVNk since the occurrence of this form in 'Oink' is considered marginal. However, this word is counted in the log(p) analysis.

*The first value listed in the miss column the number of serious misses, i.e. cases in which a form that occurs with several different vowels across many dialects of English is not predicted. The second value is the number of 'marginal' misses, i.e. cases in which the form occurs either a very small number of times (2 or less) or occurs with just a single vowel, and then only in some dialect.

Table 5

Parameters of Linear and Interaction Models

Parameter	Linear Model	Interaction Model
Baseline	20.73	22.33
↓VV	-2.91	-4.46
↓VO	-4.85	-5.75
↓LIQ	-15.51	-15.76
↓NAS	-11.55	-10.11
↓NC-lab	-3.06	-3.64
↓NC-vel	-2.75	-1.29
↓preS	-14.26	-15.10
↓postS	-14.57	-16.37
↓postT	-16.43	-17.87
↓NC x VV		-6.04
↓ NC x Nas x (Vo or VV)*		-7.00
↓NC x (l_ or s_)		-1.83
↓Vel x Vo x (VV or E)		-3.68

*The value of this parameter is indeterminate since none of the forms defined by the conjunction actually occur. The actual value shown is large enough to account for the non-occurrence of any of these forms.

Table 6

Details of the incorrectly predicted rhyme types

Rhyme	Actual Count	Sum Model Predictions	Sum + Interactions Model Predictions	Words
V zd	0	1.62*	1.47	
V Ng	0	1.59*	-1.51	
V mb	0	1.28*	-0.18	
VV mp	0	3.22*	-4.93	
VV Nk	0.1/0	3.52*	-2.57	*oink*
VV sp	0.5/0.0	0.50	-8.74	*_asp*(5)
VV sk	0.8/0.2	0.80	-6.39	*_ask* (7), *brusque*
VV ps	0.2/0.2	0.20	-2.14	*corpse*, *traipse*
V lb	0.4/0.4	-2.68*	-4.66	*alb, bulb*
VV ld	1.9/3.2	-2.53*	-3.64	*_ald*(2), *_old*(9),*_ield*(4),*_ild*(3)

.

Appendix

Phonemes and CELEX Symbols Used

Short Vowels

p**i**t	I
p**e**t	E
p**a**t	&
p**u**tt	V
p**o**t	O

Basic Long Vowels

b**ea**n	i:
b**ay**	eI
b**uy**	aI
b**oo**n	u:
n**o**	@U

Extended Long Vowels

b**ar**n	A:
b**or**n	O:
b**ur**n	3:
b**oy**	OI
br**ow**	aU

Consonants

pat	p
bad	b
tack	t
dad	d
cad	k
game	g
ba**ng**	N
mad	m
nat	n
lad	l
rat	r
fat	f
vat	v
thin	T
then	D
sap	s
zap	z
sheep	S
mea**s**ure	Z
yank	j
lo**ch**	x
had	h
why	w
cheap	tS
jeep	dZ

Figure Captions

Figure 1. Top: Probability of various observed occurrence rates for several underlying rates of a variable distributed according to the Poisson distribution. Middle: probability of different average occurrence rates for average of 5 Poisson variables. Bottom: Same as middle for average of 10 Poisson variables.

Figure 2. Partial ordering for English stop rhyme types. Dashed lines indicate violated implications, greyed out lines and forms indicate they are unattested, large arrows indicate the set of implications from each embellished rhyme to its corresponding unembellished form, and solid smaller arrows indicate normal, satisfied implications.

Figure 3. Partial ordering for English rhyme types containing an affricate, and their relationship to stop rhymes. Greyed out lines and forms indicate they are unattested, large arrows indicate the set of implications from each embellished rhyme to its corresponding unembellished form, and solid smaller arrows indicate normal, satisfied implications.

Figure 4. Partial ordering for English fricatives rhyme types. Dashed lines indicate violated implications, greyed out lines and forms indicate they are unattested, large arrows indicate the set of implications from each embellished rhyme to its corresponding unembellished form, and solid smaller arrows indicate normal, satisfied implications.

Figure 5. Partial ordering for rhymes with Null, Liquid-only, and Nasal-only codas. Dashed lines indicate violated implications, greyed out lines and forms indicate they are unattested, large arrows indicate the set of implications from each embellished rhyme to its corresponding unembellished form, and solid smaller arrows indicate normal, satisfied implications.

Figure 6. Relationship between actual and predicted average occurrence rates per vowel, based on the predictions of the multiplicative model. The logarithmic function a= log((p+1)/2), has been scaled by eye to illustrate the curvilinear relationship between the predicted (p) and actual (a) average occurrence rates.

Figure 7. Observed average occurrence rates for rhymes types containing stops.

Figure 8. Predicted and observed average occurrence rates for rhyme types containing stops. Predications are based on the linear threshold model with interactions. Error bars indicate the 90% confidence interval around the predicted value for each rhyme type.

Figure 9. Average occurrence rates for several rhyme types, for words falling at or above several frequency cutoffs. The left-most data point on each curve includes words listed in the corpus having a 0 frequency count. Other cut-offs fall at half or whole powers of 10 (e.g $10^{.5}$, 10^{1}, $10^{1.5}$, etc.). The panel at right allows comparison of the coronal variants of the rhyme types containing stops to each other and to the curve for the very common unembellished long vowel rhyme type, VV.

Figure 10. Fraction of words of the indicated classes of rhyme types that end in a coronal consonant (/t/, /d/, /n/) rather than a non-coronal (/k/ or /p/; /g/ or /b/; /N/ or /m/) within each of several frequency bands. The 'Ave' curve is the average of the four classes of rhyme types listed below Ave in the key. The error bars reflect the range of values across the four rhyme types contributing to each average. The observed coronal proportion falls within the 95% confidence interval for a sample of the given size around the average value for 18 of the 19 combinations of frequency and rhyme class, matching the expected 5% error rate; there are 19 rather than 20 combinations because there are no words in the highest frequency class for the V*v* rhyme class. Since there are two non-coronal and only one coronal consonant in each class, the expected proportion coronal in the absence of a preference for coronals would be .333; this value is indicated by the dotted line across the figure.

Figure 11. Proportion of rhymes containing long vowels within classes ending in the single coronal consonants /d/, /t/, or /n/. The proportion in a given band is calculated by summing, over the three coronals, the number of words with a short vowel in the given frequency band, and dividing by the sum over the three coronals of the number of words with either a short or a long vowel. Since there are twice as many long vowels as short vowels, the proportion that would be expected in the absence of a preference for short vowels would be .333, corresponding to the dotted line across the figure.

Figure 1

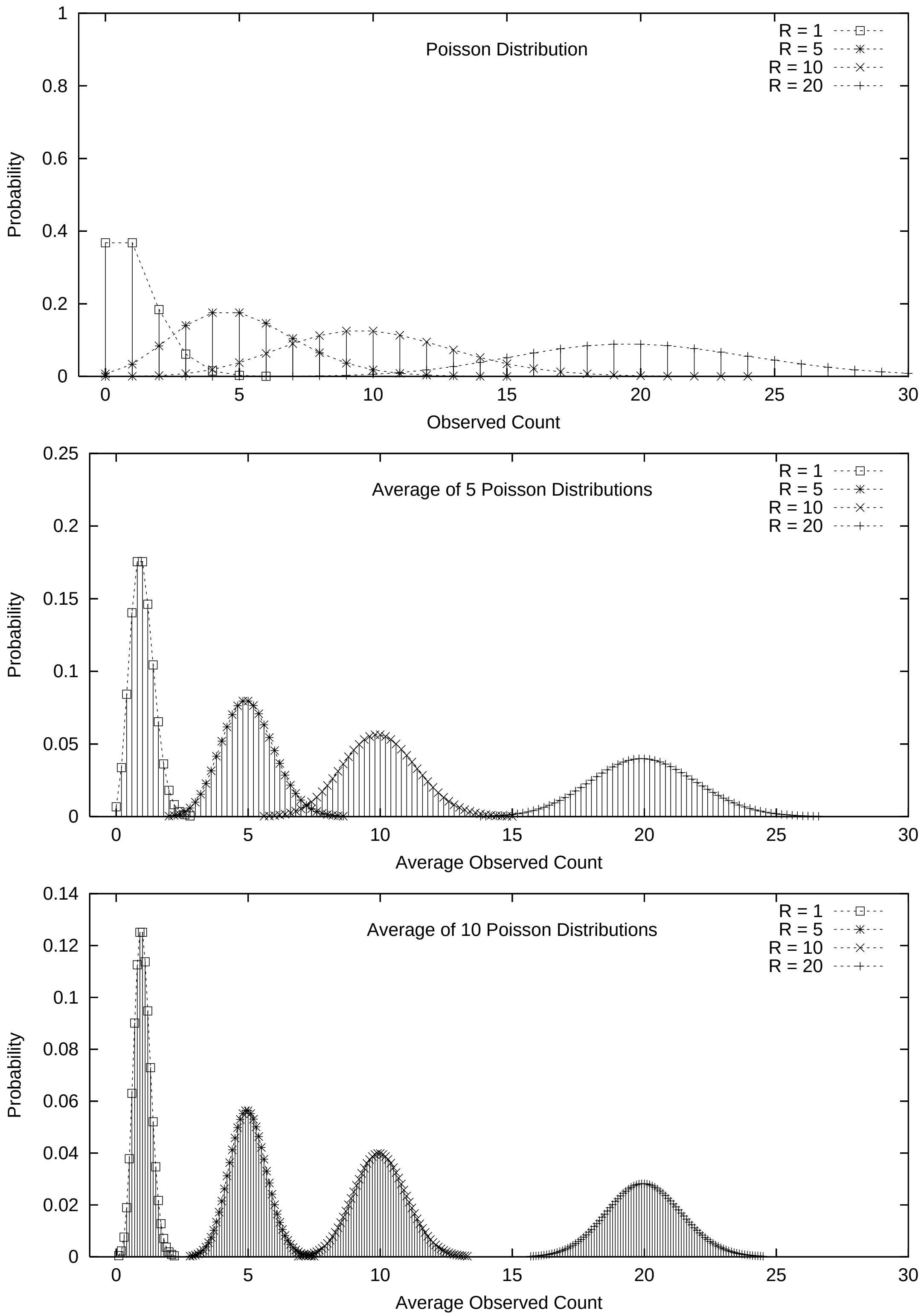

Poisson Distribution
R = 1
R = 5
R = 10
R = 20
Probability
Observed Count
0
0.2
0.4
0.6
0.8
1
0
5
10
15
20
25
30
Average of 5 Poisson Distributions
R = 1
R = 5
R = 10
R = 20
Probability
Average Observed Count
0
0.05
0.1
0.15
0.2
0.25
Average of 10 Poisson Distributions
R = 1
R = 5
R = 10
R = 20
Probability
Average Observed Count
0
0.02
0.04
0.06
0.08
0.1
0.12
0.14

Figure 2

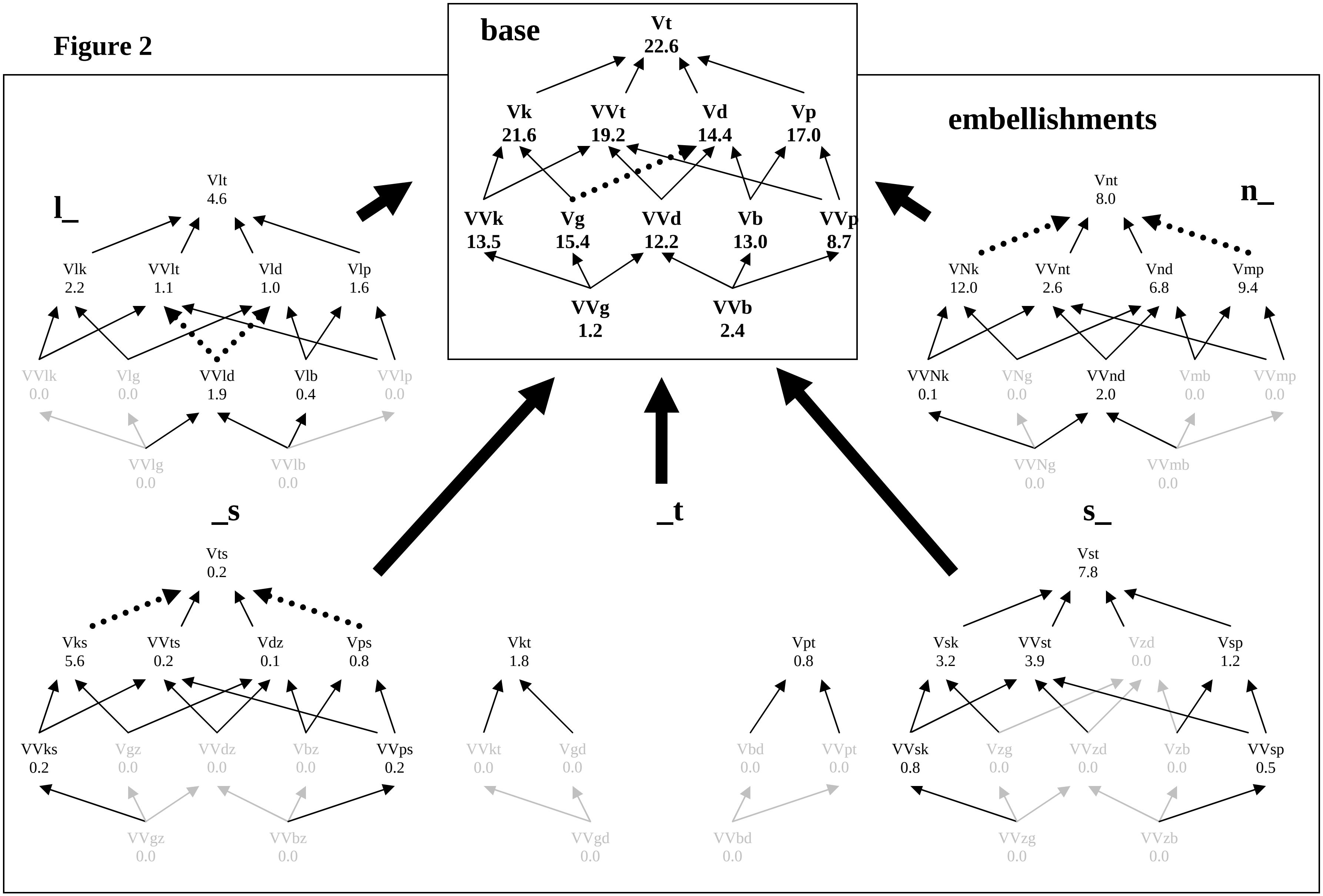

base
Vt
22.6
Vk
21.6
VVt
19.2
Vd
14.4
Vp
17.0
VVk
13.5
Vg
15.4
VVd
12.2
Vb
13.0
VVp
8.7
VVg
1.2
VVb
2.4
embellishments
l_
Vlt
4.6
Vlk
2.2
VVlt
1.1
Vld
1.0
Vlp
1.6
VVlk
0.0
Vlg
0.0
VVld
1.9
Vlb
0.4
VVlp
0.0
VVlg
0.0
VVlb
0.0
n_
Vnt
8.0
VNk
12.0
VVnt
2.6
Vnd
6.8
Vmp
9.4
VVNk
0.1
VNg
0.0
VVnd
2.0
Vmb
0.0
VVmp
0.0
VVNg
0.0
VVmb
0.0
_s
Vts
0.2
Vks
5.6
VVts
0.2
Vdz
0.1
Vps
0.8
VVks
0.2
Vgz
0.0
VVdz
0.0
Vbz
0.0
VVps
0.2
VVgz
0.0
VVbz
0.0
_t
Vkt
1.8
Vpt
0.8
VVkt
0.0
Vgd
0.0
Vbd
0.0
VVpt
0.0
VVgd
0.0
VVbd
0.0
s_
Vst
7.8
Vsk
3.2
VVst
3.9
Vzd
0.0
Vsp
1.2
VVsk
0.8
Vzg
0.0
VVzd
0.0
Vzb
0.0
VVsp
0.5
VVzg
0.0
VVzb
0.0

Figure 3

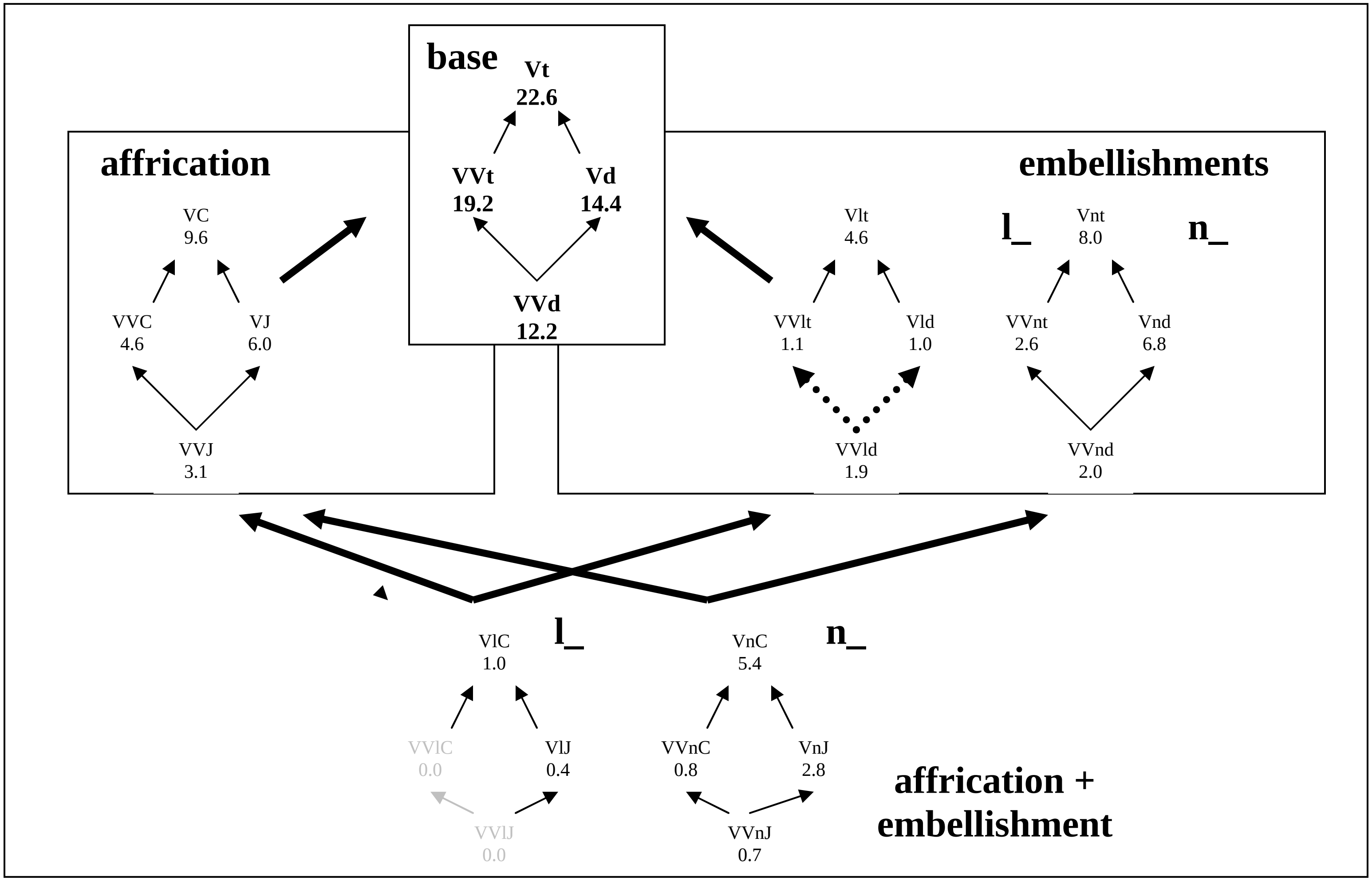

base
Vt
22.6
VVt
19.2
Vd
14.4
VVd
12.2
affrication
VC
9.6
VVC
4.6
VJ
6.0
VVJ
3.1
embellishments
l_
n_
Vlt
4.6
VVlt
1.1
Vld
1.0
VVld
1.9
Vnt
8.0
VVnt
2.6
Vnd
6.8
VVnd
2.0
l_
n_
VlC
1.0
VVlC
0.0
VlJ
0.4
VVlJ
0.0
VnC
5.4
VVnC
0.8
VnJ
2.8
VVnJ
0.7
affrication +
embellishment

Figure 4

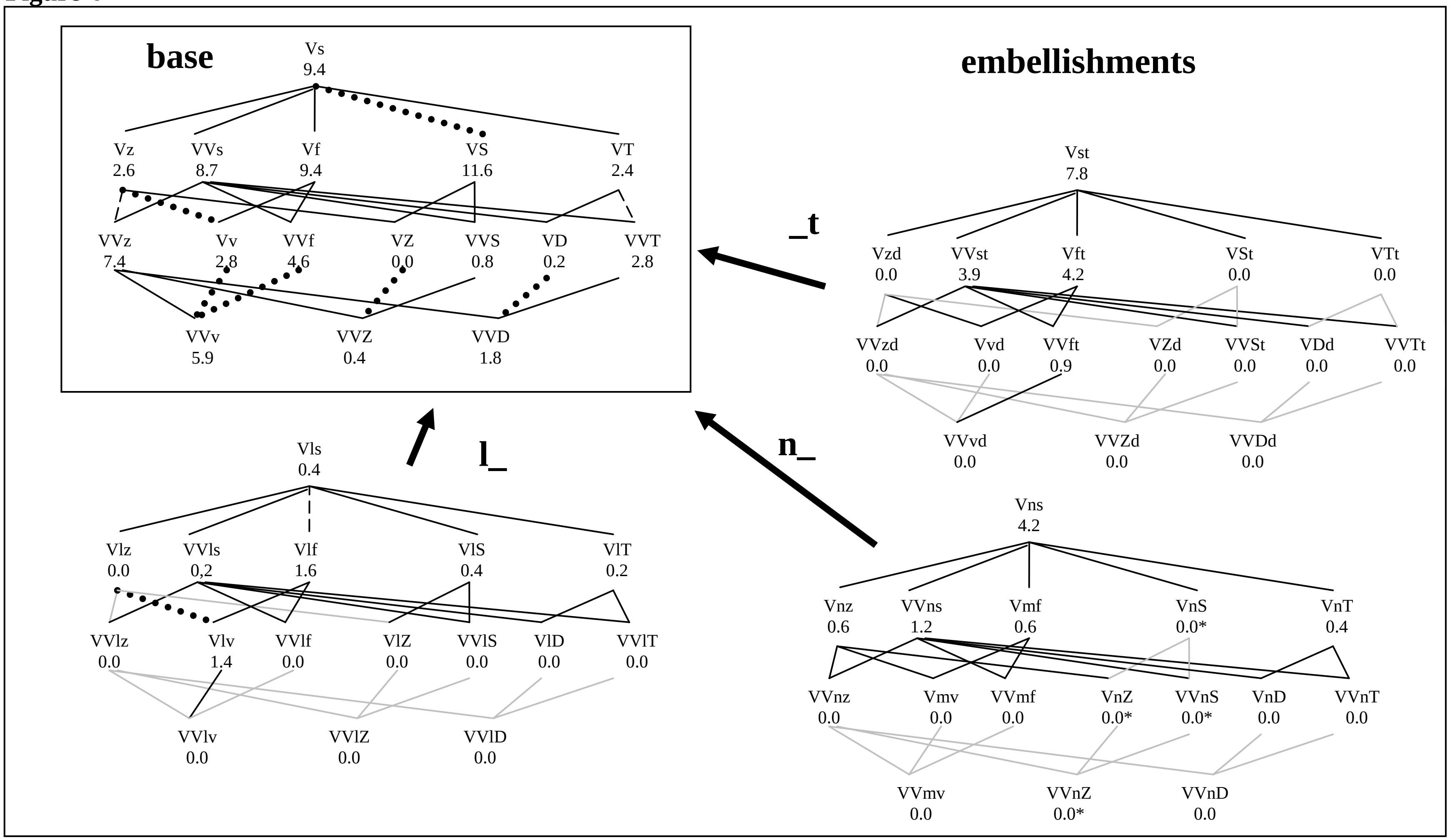

base
Vs
9.4
Vz
2.6
VVs
8.7
Vf
9.4
VS
11.6
VT
2.4
VVz
7.4
Vv
2.8
VVf
4.6
VZ
0.0
VVS
0.8
VD
0.2
VVT
2.8
VVv
5.9
VVZ
0.4
VVD
1.8
embellishments
_t
Vst
7.8
Vzd
0.0
VVst
3.9
Vft
4.2
VSt
0.0
VTt
0.0
VVzd
0.0
Vvd
0.0
VVft
0.9
VZd
0.0
VVSt
0.0
VDd
0.0
VVTt
0.0
VVvd
0.0
VVZd
0.0
VVDd
0.0
l_
n_
Vls
0.4
Vlz
0.0
VVls
0,2
Vlf
1.6
VlS
0.4
VlT
0.2
VVlz
0.0
Vlv
1.4
VVlf
0.0
VlZ
0.0
VVlS
0.0
VlD
0.0
VVlT
0.0
VVlv
0.0
VVlZ
0.0
VVlD
0.0
Vns
4.2
Vnz
0.6
VVns
1.2
Vmf
0.6
VnS
0.0*
VnT
0.4
VVnz
0.0
Vmv
0.0
VVmf
0.0
VnZ
0.0*
VVnS
0.0*
VnD
0.0
VVnT
0.0
VVmv
0.0
VVnZ
0.0*
VVnD
0.0

Figure 5

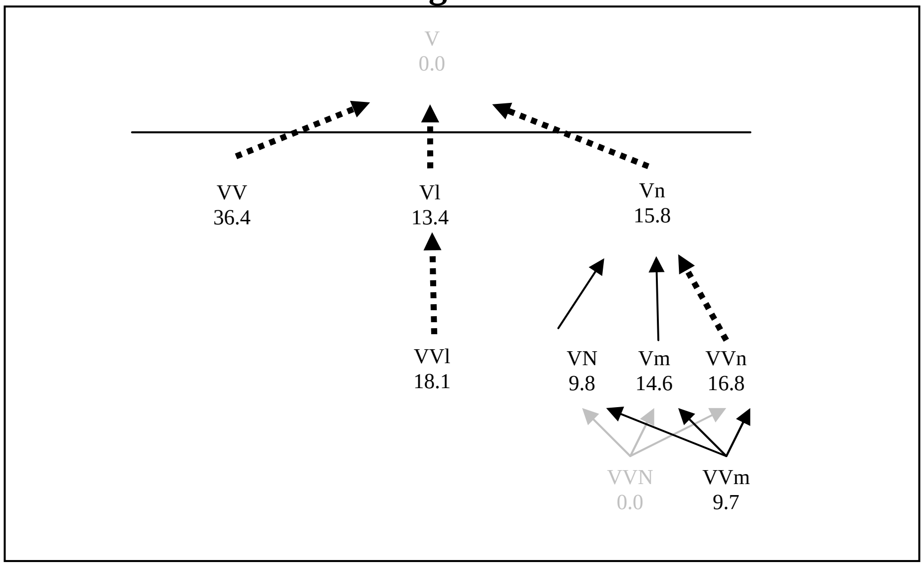


Figure 6

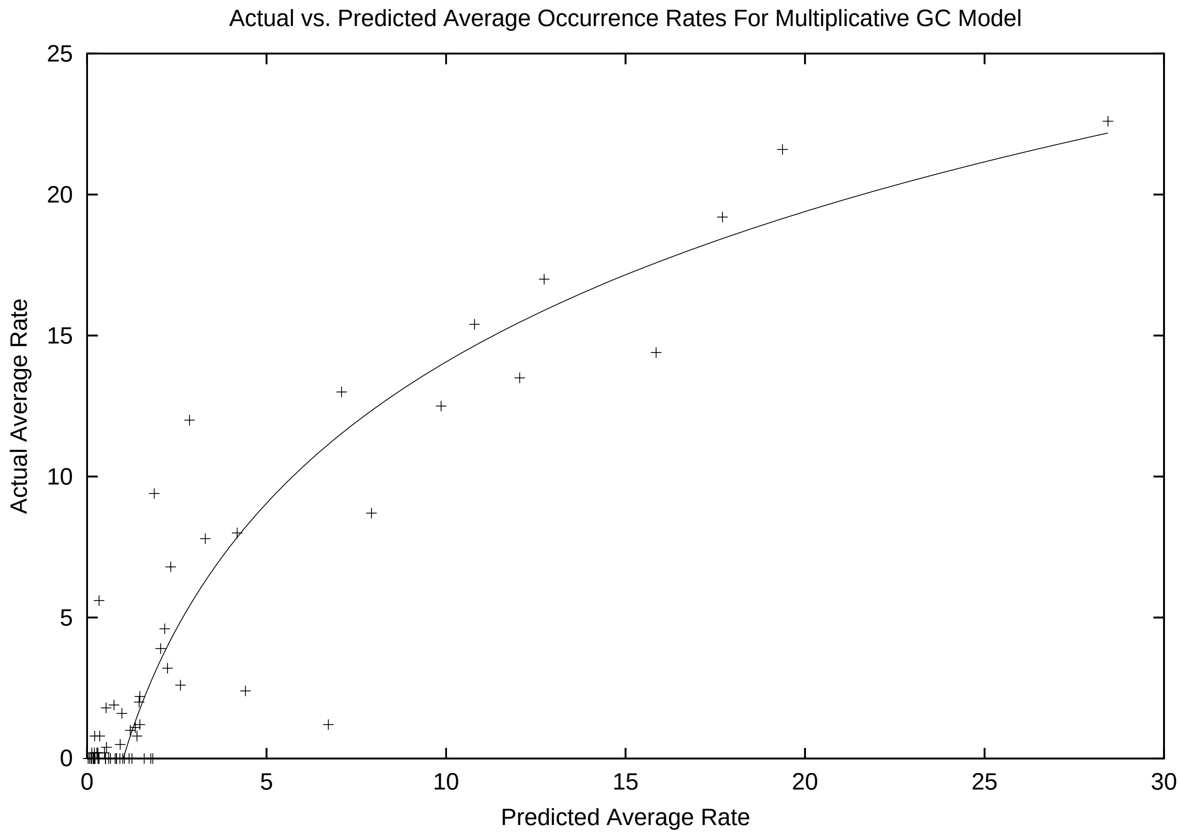

Actual vs. Predicted Average Occurrence Rates For Multiplicative GC Model
Actual Average Rate
Predicted Average Rate
0
5
10
15
20
25
30

Figure 7

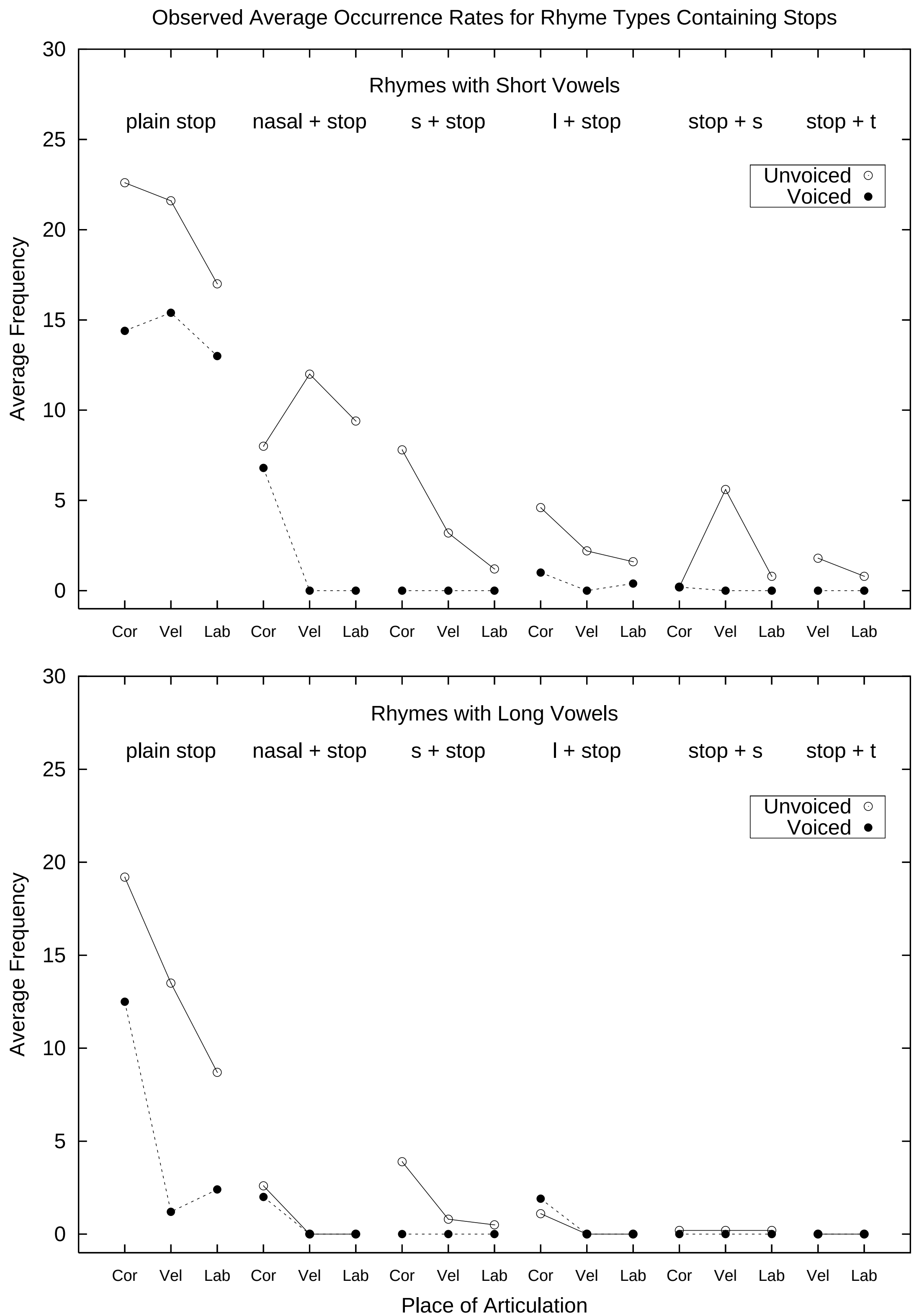

Observed Average Occurrence Rates for Rhyme Types Containing Stops
Rhymes with Short Vowels
plain stop
nasal + stop
s + stop
l + stop
stop + s
stop + t
Unvoiced
Voiced
Average Frequency
30
25
20
15
10
5
0
Cor Vel Lab Cor Vel Lab Cor Vel Lab Cor Vel Lab Cor Vel Lab Vel Lab
Rhymes with Long Vowels
plain stop
nasal + stop
s + stop
l + stop
stop + s
stop + t
Unvoiced
Voiced
Average Frequency
Cor Vel Lab Cor Vel Lab Cor Vel Lab Cor Vel Lab Cor Vel Lab Vel Lab
Place of Articulation

Figure 8

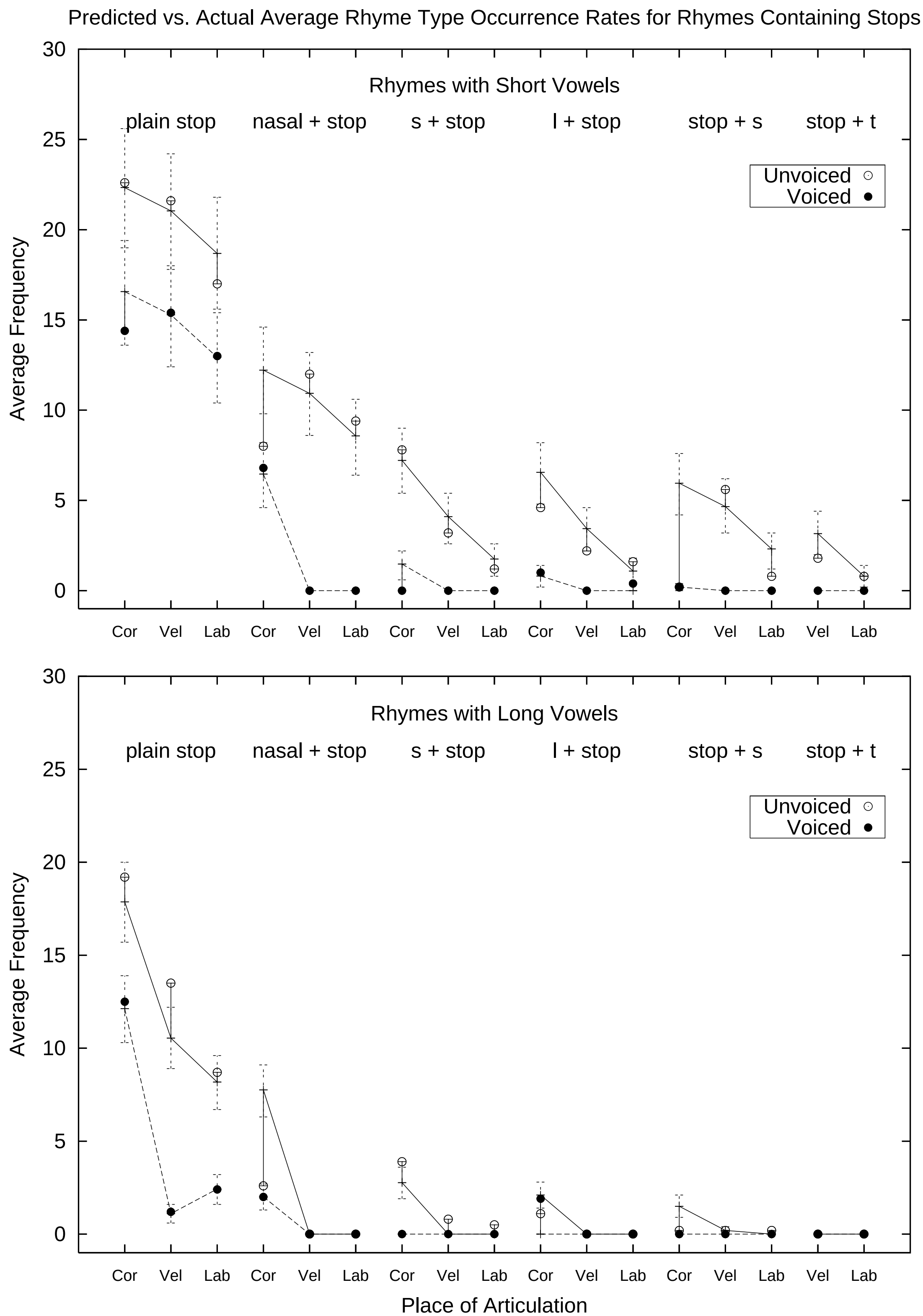

Predicted vs. Actual Average Rhyme Type Occurrence Rates for Rhymes Containing Stops
Rhymes with Short Vowels
plain stop
nasal + stop
s + stop
l + stop
stop + s
stop + t
Unvoiced
Voiced
Average Frequency
Cor Vel Lab Cor Vel Lab Cor Vel Lab Cor Vel Lab Cor Vel Lab Vel Lab
Rhymes with Long Vowels
Place of Articulation

Figure 9

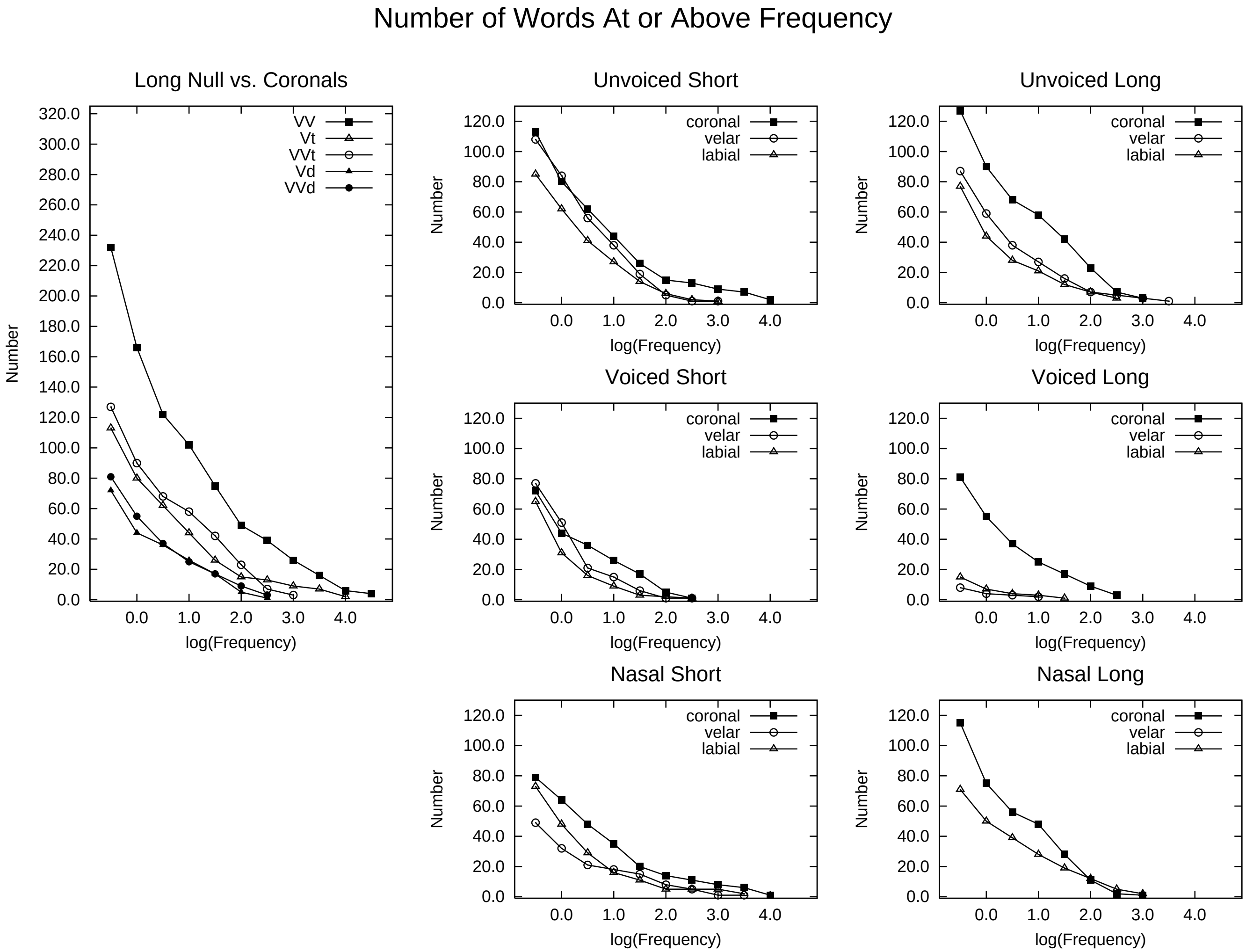

Number of Words At or Above Frequency
Long Null vs. Coronals
VV
Vt
VVt
Vd
VVd
Number
log(Frequency)
Unvoiced Short
Unvoiced Long
Voiced Short
Voiced Long
Nasal Short
Nasal Long
coronal
velar
labial

Figure 10

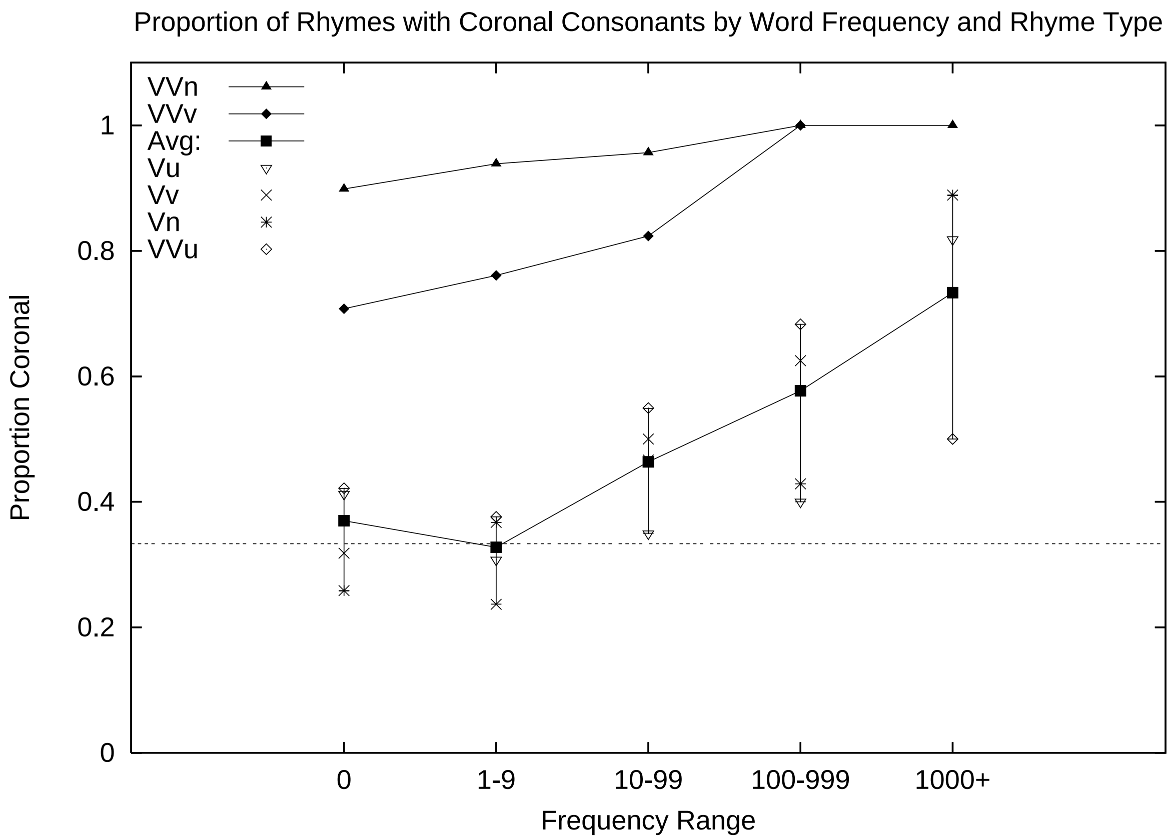
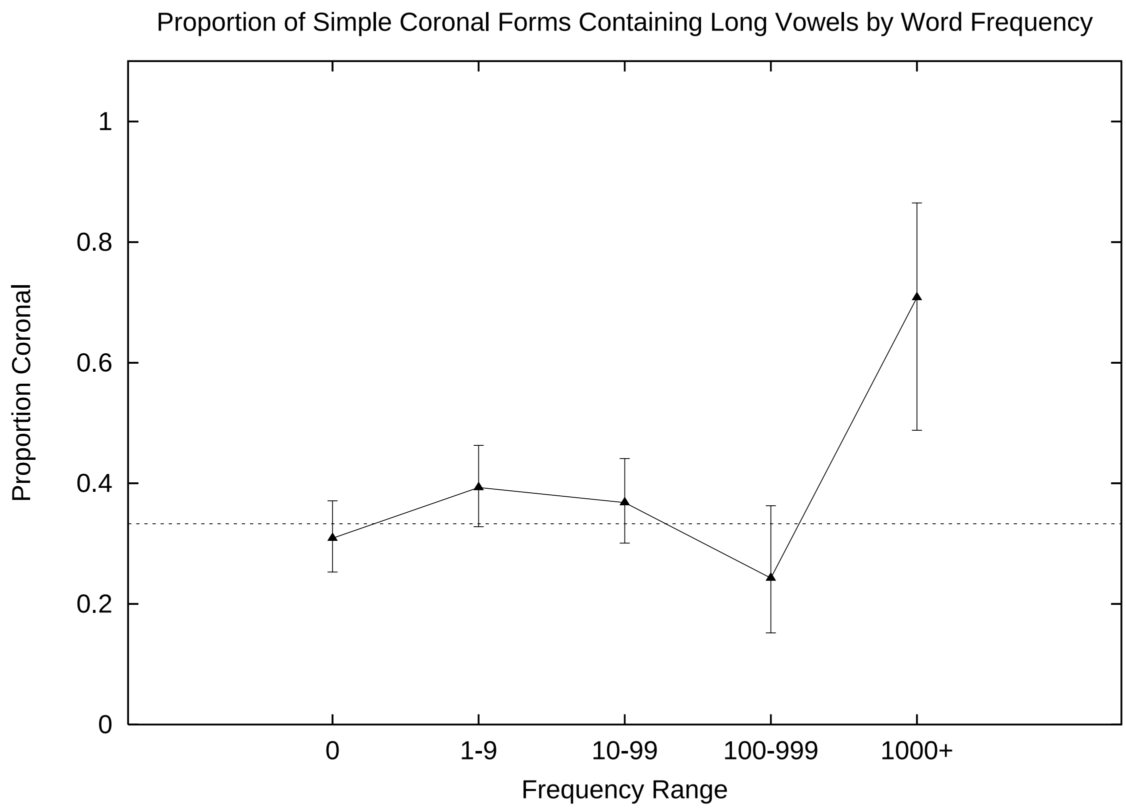

Proportion of Rhymes with Coronal Consonants by Word Frequency and Rhyme Type
VVn
VVv
Avg:
Vu
Vv
Vn
VVu
Proportion Coronal
0
0.2
0.4
0.6
0.8
1
0
1-9
10-99
100-999
1000+
Frequency Range

Figure 11


Proportion of Simple Coronal Forms Containing Long Vowels by Word Frequency
Proportion Coronal
1
0.8
0.6
0.4
0.2
0
0
1-9
10-99
100-999
1000+
Frequency Range