Language Variation and Change, 9 (1997), 149–164. Printed in the U.S.A.

Inherent variability and the obligatory contour principle

GREGORY R. GUY
York University

CHARLES BOBERG
University of Pennsylvania

ABSTRACT

English coronal stop deletion is constrained by the preceding segment, so that stops and sibilants favor deletion more than liquids and nonsibilant fricatives. Previous explanations of this constraint (e.g., the sonority hierarchy) have failed to account for the details, but we show that it can be comprehensively treated as a consequence of the obligatory contour principle (OCP). The OCP, introduced to account for a variety of categorical constraints against adjacent identical tones, segments, and so forth, can be generalized as a universal disfavoring of sequences of like features: $*[\alpha F]\,[\alpha F]$. Therefore, coronal stop deletion, which targets the set of segments /t,d/ defined by the features [−son, −cont, +cor], is favored when the preceding segment shares any of these features. But this requires adopting the assumption of inherent variability and interpreting the OCP as a probabilistic constraint with cumulative effects (the more shared features, the greater likelihood of deletion). This suggests an attractive theoretical integration of categorical and variable processes in the grammar.

The systematic patterns of variation revealed by sociolinguistic studies of language use are generally modeled in theoretical frameworks that rely on the assumption of "inherent variability" (Labov, 1969; Weinreich, Labov, & Herzog, 1968); that is, the hypothesis that the human language faculty necessarily accommodates and generates variation, and that the workings of grammar have a quantitative, noncategorical, and nondeterministic component. This assumption underlies the variable rule model (Cedergren & Sankoff, 1973; Labov, 1969), which is given a more general formulation as a probabilistic generative model by Sankoff (1978). One of the attractive features of such models is that they offer an integrated account in which social and stylistic dimensions of variation can be modeled along with linguistic dimensions. Thus, in a standard variable rule analysis, social differences between individuals in a speech community are represented by their characteristic values of an input parameter, whereas independent linguistic parameters describe favorable and unfavorable linguistic contexts for the occurrence of particular variants. Stylistic variation is another independent parameter, which

This work was supported in part by grant number 410-92-1765 from the Social Sciences and Humanities Research Council of Canada.

adjusts an individual's characteristic rate up or down according to social/psychological factors.

Tested empirically, the variable rule approach has proven quite successful, providing an accurate and quantitatively detailed model for many complex variables showing social stratification, style shifting, and linguistic conditioning. The basic assumptions of the model—such as the independence of the parameters, the constancy of linguistic parameter values across individuals within a speech community, and so forth—have been fairly well confirmed by research work in this tradition, although there are still some open questions.

But despite this record of success, the variable rule model has run into conflict with one of the theoretical traditions of linguistics, namely, the tradition that says that linguistic competence is categorical, discrete, and nonquantitative. In the latter approach, an assumption of invariance replaces the inherent variability hypothesis. This position is grounded in a long history of categorical, invariant constructs in linguistic theory: it is hypothesized that sound changes are exceptionless, distinctive features are binary, parameter settings of universal grammar are polar, speech perception is categorical, and so forth. But it presents a basic conflict with the variationist ideas that the output of a linguistic rule can be probabilistic rather than discrete, and that a linguistic constraint can have a quantitative effect on the outcome of a process—boosting it or hindering it by a certain probability, rather than categorically triggering or blocking it.

This conflict is not about the social and stylistic aspects of variation; the scholarly community in linguistics generally accepts the idea that there are quantitative differences of behavior along such dimensions. Nor can there be any argument about the existence of linguistic constraints on variability. These are undisputed empirical observations. Rather, the conflict is about how to represent these facts in our models of language. Those who accept the assumption of invariance postulate a mental grammar that is categorical and has no representation of frequency or likelihood. Probabilistic patterns of language use are treated as a product of performance, arising outside the grammar. Any explanation for them has to be sought in some (as yet undeveloped) theory of performance.

By contrast, if we postulate inherent variability, we have at least the prospect of an integrated theory of language. Instead of partitioning the evidence and attributing categorical facts to competence and variable facts to performance, we can postulate a unified probabilistic grammar that accounts for both kinds of patterns. Such a grammar is certainly capable of generating both categorical and variable patterns, because in probabilistic terms categorical outcomes are those that have a maximally high probability.

On what basis can a choice be made between these two alternatives? In this article, we suggest an empirical approach to the question. A theory that adopts the assumption of invariance makes a clear negative prediction about the relevance of grammatical principles to the patterning of variability. The principles of a categorical competence are inside the grammar and do not govern what goes on outside, in other areas of cognition and behavior. Therefore, there is no obvious reason to expect that theoretical principles designed to account for categorical phenomena should also explain variable phenomena. But in a theory that postulates inherent variability within the grammar, it is perfectly plausible that some principles could govern both variable and invariant properties. So the empirical question arises, are there such cases? If so, this would suggest that variability is located in the grammar, and that variable and invariant properties can be treated in an integrated way, rather than assigning them to separate spheres of competence and performance.

In this article, we show that one well established principle of formal phonology, the obligatory contour principle (OCP), which was originally proposed to account for invariant properties of morpheme structure and phonotactics, also provides a compelling explanation for probabilistic linguistic constraints on a variable rule, the English rule of final coronal stop deletion. This discovery validates both the integrated account of variation provided by the variable rule framework and the formal theory of phonology that led to the postulation of the OCP. However, it raises substantial questions about the assumption of invariance.

In the balance of this article, we discuss the constraints on coronal stop deletion. We show how one of them can be explained in terms of OCP and how this analysis makes novel, confirmable predictions. Finally, we suggest the implications of these results for the larger theoretical questions we have raised.

ENGLISH CORONAL STOP DELETION

The deletion of clustered final coronal stops is a well studied phenomenon of natural speech in English, whereby words like *west* and *old* lose their final /t/ or /d/, especially in phrasal contexts like *wes' side* and *ol' man*. The process is universally variable in English—that is, every speaker that has been observed deletes some, but not all, of their final stops. Previous studies of the process in a variety of dialects have shown that the rule is subject to the same kinds of conditioning that have been found for other sociolinguistic variables. There is social differentiation (e.g., the class differences reported in Wolfram, 1969), stylistic variation (see Guy, 1980), and a complex set of linguistic constraints, so that rate of deletion is affected by morphological structure, adjacent segments, syllable structure, and other factors (see Guy, 1980, 1991, for a fuller explication).

For the purpose of demonstrating the theoretical utility of an integrated approach to variable and invariant phenomena, we focus on one of the linguistic constraints on coronal stop deletion, namely, the preceding segment effect. Recent work by one of us (Guy, 1991, 1992) characterized this rule as a delinking of the final stop from its structural position, with subsequent

TABLE 1. *Preceding segment effect on coronal stop deletion (Guy, 1980 corpus: 9 Philadelphians)*

Preceding Segment	Probability of Deletion[a]
Sibilant fricatives	1.00
Stops	0.84
Nasals	0.78
Nonsibilant fricatives	0.69
Laterals	0.66
$N = 1{,}860$	

[a]Factor weights from VARBRUL 1.

removal by stray erasure, as indicated in (1). Unsurprisingly, the rule is influenced by the other elements of the coda to which the targeted stop belongs.

(1) Final stop delinking
⟨variable, unmarked domain of application⟩

$$
\begin{array}{ll}
\sigma\ ] & \\
\ldots C & C \\
\langle \alpha F1 \rangle & [-\text{son}] \\
\langle \beta F2 \rangle & [-\text{cont}] \\
 & [+\text{cor}] \\
\vdots & \vdots
\end{array}
$$

The open question on this point is, what exactly is the nature of this influence? What features of the preceding segment promote or hinder deletion? Previous studies of the process have not been definitive on this point. Quantitative analyses generally have classified the preceding segments in terms of broad manner classes, distinguishing stops, fricatives, nasals, and liquids. The liquid class is often restricted to laterals, because deletion after /r/ is rare among white speakers (although it is more common in AAVE). The fricatives are often subdivided into sibilants and nonsibilants, because it was noted in some early studies of this process that sibilants are associated with a higher deletion rate (i.e., more deletion in *wrist* than *rift*).

Some typical results for this kind of five-class analysis are given in Table 1 (from Guy, 1980). They show, for a group of nine native speakers of English from Philadelphia, the highest rate of deletion after a preceding sibilant, followed in descending order by stops, nasals, nonsibilant fricatives, and laterals. However, it should be noted that the individuals in this study showed a lot of deviations from this particular rank order, especially for the intermediate categories like nasal and stop. Some of the scatter in the individual results was shown to be due to small sample size and random error, but even so, the scatter was appreciably greater for the preceding segment effect than for the other constraints on the rule. Furthermore, the question of whether all the categories are significantly different from one another has never been adequately addressed. In Table 1, the values for laterals and nonsibilant fricatives are very close: 0.66 and 0.69; nasals and stops are also fairly close. Although the significance of this factor group as a whole is well established, it could be the case that some of the distinctions made within it are not statistically significant.

Previous studies have offered a variety of explanations for the preceding segment effect, such as ease of articulation, markedness considerations about coda clusters, and sonority hierarchy effects on syllable structure. Thus, sequences of sonorant + stop are seen as being easier to produce, less marked, or more phonotactically favored than obstruent + stop clusters. But such explanations generally fail to account for the details of the constraint. For example, stop + stop sequences should be harder to articulate, more marked, and therefore more often deleted than fricative + stop clusters. Certainly the latter are far more common in English codas than the former, but the deletion rates usually show that a preceding stop is associated with less deletion of a final /–t,d/ than a preceding sibilant fricative. Similarly, it could reasonably be argued that, based on frequency of occurrence, /st/ clusters are less marked and perhaps easier than /ft/, but it is the /st/ cases that always show higher rates of deletion. A sonority hierarchy account comes closest to predicting the observed order. If lower sonority in the preceding segment is associated with more deletion of the final /–t,d/, then the deletion rate rankings for preceding contexts should be stop > fricative > nasal > lateral. But this does not explain why sibilants and nonsibilant fricatives differ, or why nasals often promote deletion as much as stops and fricatives. It is also unclear why deletion is so rare after /r/, which is arguably equal in sonority to /l/. Clearly, an adequate account of these facts is still needed.

AN OCP ACCOUNT

Recent work has suggested a solution to the problem. We derive the preceding segment effect on coronal stop deletion from a more general constraint on phonological sequences, namely, the OCP. The OCP was originally proposed as a prohibition against adjacent identical tones in lexical representations (Leben, 1973), but in recent theoretical work it has been expanded to account for a variety of processes that involve avoidance of adjacent identical segments (McCarthy, 1986) and adjacent identical features (Yip, 1988). Yip suggested that the OCP is essentially a universal disfavoring of identical sequences on the same tier. McCarthy (1988:88) gave a general formulation of the principle: "Adjacent identical elements are prohibited." On a featured level, these observations might be formalized as in (2).

(2) The Obligatory Contour Principle
*[αF] [αF]

For our purposes, it is important that most of the formal accounts of the OCP, including all of those cited here, have treated it as a categorical principle, which absolutely prohibits violations. This follows directly from the assumption of invariance discussed earlier. But recently a few treatments of the OCP have appeared that recognize it as gradient and noncategorical. Several studies of consonant sequences in Afro-Asiatic lexical roots (Buckley, 1993; Pierrehumbert, 1993) have shown that co-occurrence of like elements is underrepresented, but not categorically suppressed, as a gradient function of similarity and adjacency.[1] Similarly, Berkley (1994) demonstrated a significant underrepresentation in English monosyllables of similar consonants. The main feature she considered was place of articulation (homorganic consonant pairs are relatively rare compared to non-homorganic), but she also found, in the coronals, an additional effect for shared sonority values (coronal obstruent-obstruent and sonorant-sonorant pairs are infrequent compared to mixed, obstruent-sonorant pairs).

This approach to the OCP, extending it to noncategorical phenomena, provides a basis for organizing the facts about the effect of preceding segment on English coronal stop deletion. Deletion rates are correlated with the featural similarity between the targeted final /−t,d/ and the preceding segment. Deletion is more likely to apply to targets with sequences of adjacent identical features. To develop this point, let us reconsider the five segment classes used in Table 1 in terms of distinctive features. The sibilants are defined by the features [−son, +cont, +cor]. The stops are defined by [−son, −cont], but since geminates are disallowed in English, the only stops that can cluster with final /t,d/ are the noncoronals /p,b,k,g/ (e.g., in words like *act, rapt, dragged*); hence, the specification [−cor] can be added to this group. Nasals are [+son, −cont]; within a morpheme they are also [+cor], but in past tense verbs, clusters occur with [−cor] nasals (e.g., rammed, banged). Nonsibilant fricatives are [−son, +cont, −cor], and laterals are [+son, +cont, +cor]. In all cases, the relevant features are [son, cont, cor].

If the target of OCP prohibition is adjacent like features, then coronal stop deletion should be promoted when the preceding segment shares features with the final /t,d/. The coronal stops for these three features are defined as [−son, −cont, +cor]. Whenever a preceding segment has one of these feature values, the OCP would favor deletion. By the variationist hypothesis that such effects are probabilistic, this favoring should be cumulative: the more features shared, the more deletion. Viewed in this light, we have a straightforward account for previous findings. Sibilants share with the targeted /t,d/ the features [−son, +cor], and stops share the features [−son, −cont]; these are the categories that most favor deletion. The preceding nasals include /n/, which shares two targeted features [−cont, +cor], and /m,ŋ/, which share only one feature [−cont]. Since this category mixes segments sharing two fea-

TABLE 2. *Preceding segment effect: An OCP analysis (Guy & Boberg, 1994 corpus: 3 Philadelphians)*

Preceding Segment		Deletion		
		n	%	Factor Weight[a]
/t,d/	[+cor, −son, −cont]	–	(categorical absence, i.e., 1)	
/s,z,ʃ,ʒ/	[+cor, −son]	276	49	0.69
/p,b,k,g/	[−son, −cont]	136	37	0.69
/n/	[+cor, −cont]	337	46	0.73
/f,v/	[−son]	45	29	0.55
/l/	[+cor]	182	32	0.45
/m,ŋ/	[−cont]	9	11	0.33
/r/	?	86	7	0.13
vowels	–	–	(nearly categorical retention, i.e., ≈0)	
N = 1,071, log likelihood = −533.173				

[a]From VARBRUL 2.

tures and one feature, it shows a somewhat lower deletion rate than stops and sibilants. Next in rate of deletion come nonsibilant fricatives, which are really just the noncoronal fricatives, having only one deletion-favoring feature [−son]. Finally, there are the laterals, which share only the place feature [+cor] with the coronal stops.

A better test of this analysis would be to divide the nasal category in Table 1 into coronals and noncoronals, since the model predicts more deletion with preceding /n/ (with the targeted features [−cont, +cor]) than with preceding /m, ŋ/ (with just [−cont]). In Table 2, we show the results of an analysis of a new corpus of Philadelphia speech compiled by Boberg.[2] As predicted by the OCP, the three categories that have two of the targeted features cluster together at the top: stops, sibilants, and /n/. Their factor weights are all high and very closely grouped at 0.69 to 0.73. Lower and more loosely clustered are the three categories with only one of the favoring features, laterals and noncoronal fricatives and nasals, with factor weights ranging from 0.33 to 0.55. As our model predicts, there is a big difference between /n/ and the noncoronal nasals. This is a previously unsuspected result.

Let us now consider differences within the two clusters of categories. Are stops, sibilants, and /n/ significantly different from one another in deletion rates, or are they substantially the same? What about the differences among laterals and noncoronal nasals and fricatives? As we have observed, such questions have not been adequately addressed in previous studies. The theory we have presented leaves the question open. If each favoring feature has exactly the same effect, then any combination of two favoring features should promote deletion at the same rate. In that case, stops, sibilants, and /n/ should have essentially equal values. Similarly, the bottom three categories, each with only one favoring feature, should be nearly equal. However, it is also possible that each favoring feature could have a different effect, in which case significant differences between categories within a cluster would be observed.

In our data, no significant differences emerge within the clusters. In other words, although having two favoring features produces significantly more deletion than having only one, it does not matter which two features are present. Stops, sibilants, and /n/s are all substantially equal. Also, the laterals and noncoronal nasals and fricatives do not differ significantly from each other. These results suggest that each of the features [son], [cont], and [cor] has about the same effect on the rule.

This analysis can be extended to encompass the cases at the extremes, which are also cited in Table 2. As we have noted, the English prohibition on geminates means that the final /t,d/ can never be preceded by another /t/ or /d/. This is an absolute constraint on underlying forms, and devices are available in the morphophonology to avoid creating derived geminates. The best known of these is the syllabic allomorph of the regular past tense suffix which occurs just after a root-final /t/ or /d/ (cf. *wanted, sanded* vs. the nonsyllabic suffix in *walked, canned, missed*). This epenthesis rule (or nondeletion of an underlying suffixal vowel) is, in itself, an OCP effect (see Borowsky 1986:138–153), and it occurs just where all three features are shared between target and context. In terms of the present discussion, it can be expressed as a categorical absence of final /t,d/ after a preceding /t,d/, which is equivalent to a rule probability of 1. In this way, the single explanatory principle of OCP accounts for both categorical and variable conditions on the occurrence and absence of final coronal stops.

At the other extreme of the continuum, the OCP also explains the effect of preceding vocalic elements. In most English dialects, coronal stop deletion occurs rarely or not at all after vowels (e.g., *hid, head*) and glides (e.g., *hide, loud*). The same is true, as we noted earlier, of /r/. As formulated in (1), the rule requires a filled C slot before the targeted stop, and so preceding glides and vowels do not trigger the rule. But if the facts are recast in terms of preferred contours on the feature tiers, it becomes possible to subsume this condition under the OCP as well. Glides and vowels share none of the deletion-promoting features: they are [+son, +cont, −cor]. Therefore, no undesirable sequences of like features occur when they are followed by a coronal stop coda. Hence, they should have the lowest deletion rate, which they do.

Under this analysis, the separate structural requirement for a preceding C on the skeletal tier can be abandoned. The rule can then be restated as in (3), which variably delinks coda segments when sequences of like features occur. In this way, all the facts about possible coda sequences are accounted for in one statement, phrased in terms of the featural tiers and the operation of the OCP.

(3) Coda cluster reduction by OCP
⟨variable, unmarked domain of application⟩

$$\begin{array}{cc} \sigma\ ] & \\ | & \\ \ldots x & x \\ \langle \alpha F1 \rangle & \alpha F1 \\ \langle \beta F2 \rangle & \beta F2 \\ \vdots & \vdots \end{array}$$

However, there still remains the matter of preceding /r/. It is traditionally treated as [+cor] and so should trigger deletion at a rate equal to /l/. But, as we see in Table 2, it actually shows a much lower rate: 7% deletion, and a factor weight of 0.13. Why is it different? A possible answer to this question begins with the observation that /r/ in this position is also postvocalic. This means that, in the r-less dialects of English, these instances of /r/ are fully vocalized or deleted. Even in the r-pronouncing dialects, they are ordinarily realized phonetically as the vocalic segment /ɚ/. We can formally represent these facts by assigning postvocalic /r/ to a glide position in syllable structure. Veatch (1991:41–45) argued for such an analysis in his phonology of English vowels; for him, postvocalic /r/ functions structurally as a glide. Then, if coronal stop deletion is defined as it was in (1), preceding /r/ would automatically be excluded as a triggering context, because it would not fill the obligatory C slot preceding the target. But if we want to formalize the rule in terms of features and the OCP, as in (3), a further step in the argument is required. If postvocalic /r/ is structurally a glide and is phonetically the central vowel /ɚ/, then it should be classified as [−cor]. In that case, it, too, lacks all the features that favor coronal stop deletion and is predictably associated with a low deletion rate.

The OCP account also promises to explain another finding. It has been reported (e.g., Nesbitt, 1984; Wolfram, 1969) that clusters with contrasting voicing values show less deletion than corresponding ones with identical voicing values. Of course, given English phonotactics, a voicing contrast can only occur with a preceding sonorant, because double obstruent clusters obligatorily have the same value for voice. But in the contrasting cases, the cited studies show that words like *cold* and *tend* delete more than words like *colt* and *tent*. This, too, can be interpreted in OCP terms: having identical values for the voicing feature also promotes deletion.

Would this make a false prediction regarding the double obstruent clusters? If the OCP avoids like sequences, why are obstruent + stop clusters required to agree in voicing in English (and many other languages)? Our OCP treatment seems to make the empirically incorrect prediction that preceding obstruents would systematically favor coronal stop deletion over preceding sonorants because of obligatory voicing agreement.

An answer to this objection is readily available. Precisely because they show obligatory voicing agreement, obstruent clusters must have a single voicing specification which is doubly linked to the skeletal slots. Underspecification requires the noncontrasting voicing value to be filled in by feature spreading from a single underlying entry, giving a representation as in (4). The sonorant + stop clusters are represented with separate [voice] specifications, as in (5), because they involve contrastive voicing for the stops. Consequently, there is a potential OCP clash on the featural tier for [voice] in the sonorant case, but not in the obstruent case. Whenever a structure like (5) occurs, deletion should be promoted.

(4) Voicing in obstruent + stop clusters

```
 x   x
  \ /
[+voice]
```

(5) Voicing in sonorant + stop clusters

```
    x           x
    |           |
[+voice]    [+voice]
```

This prediction is confirmed in the present corpus. In these data, a final /t,d/ after a sonorant is deleted at a rate of 41% when the adjacent segments agree in voicing, but at a rate of only 25% when they disagree, VARBRUL analysis bears out our OCP analysis as well, yielding a deletion probability of 0.64 with voice agreement versus 0.36 with contrasting voice.

FEATURE EFFECTS

Finally, one further step in the quantitative analysis is needed. The results in Table 2 are expressed in terms of the segmental classes defined by feature combinations rather than in terms of the features themselves. Since our model assumes that each feature has an individual and independent effect, it is desirable to run a VARBRUL analysis in which each feature constitutes a separate factor group. In this approach, liquids, sibilants, and /n/ are all coded [+cor], in opposition to stops and other fricatives and nasals, which are coded [−cor]. At the same time, in another factor group, nasals and liquids are all coded [+son], whereas the other classes are coded [−son]. In a third group, fricatives and liquids are coded [+cont], whereas stops and nasals are coded [−cont]. Finally, a fourth factor group expresses the cluster voicing constraint. Sequences of nasal or liquid /+/ /d/ are coded as agreeing in voice, whereas nasal or liquid /+/ /t/ are coded as nonagreeing.

If this featural analysis fits the data adequately, it will provide the best confirmation of the hypothesis that a probabilized OCP constraint operates in-

TABLE 3. *Feature analysis of preceding segment effect on coronal stop deletion (Guy & Boberg, 1994 corpus)*

Features of Preceding Segment	Factor Weight (for Deletion)[a]
Sonority	
[−son]	0.58
[+son]	0.42
Continuancy	
[−cont]	0.65
[+cont]	0.35
Coronal Place	
[+cor]	0.65
[−cor]	0.35
Voice (preceding sonorants only)	
[αvoice]	0.64
[−αvoice]	0.36
Log likelihood = 535.033; $\chi^2 = 3.72(3)$, $p > .25$[b]	

[a]Targeted features: /t,d/ = [+cor, −son, −cont, αvoice].
[b]Comparison with Table 2.

dependently on each defining feature. The results of a VARBRUL analysis appear in Table 3.

The results conform neatly to the predictions of the model. The two favoring features [−cont] and [+cor] are equal at 0.65, whereas [−son] appears slightly weaker at 0.58. Agreement in the [voice] feature is associated with a favoring weight of 0.64. The difference between this analysis and the previous one is not significant, and so the hypothesis that each feature has an independent effect is confirmed.

The similarity of the factor values in Table 3 is quite interesting. Three of the four favoring factors are valued at 0.64/0.65; only the 0.58 for [−son] is appreciably different. We might hypothesize, given this evidence, that all features have an equal effect for OCP purposes. This would be the simplest quantitative interpretation of the statement of the OCP as *[αF] [αF]. The modest divergence of the values in the [±son] factor group might then be explained by the mathematics of VARBRUL analysis. Since voicing agreement is in effect a subcategorization of the [+son] cases, the algorithm can produce a trading relation between the two, attenuating one while augmenting the other. We found evidence of this in our data set. In several analyses where the voicing factor group was dropped, the [±son] effect appeared appreciably stronger.

This feature-by-feature quantification of the OCP effect has interesting implications for the extreme cases: it does not predict either categorical deletion or retention. Combining all the deletion-favoring features does not predict 100% deletion after a preceding coronal stop, and combining all disfavoring values does not predict 0% deletion after vowels, glides, and /r/. For stops, this is not a problem, because the absolute prohibition on geminate sequences in English holds on underlying representations; so there will never be any /tt/ or /dd/ sequences for the rule to (fail to) apply to. But underlying forms with vowels, glides, or /r/ preceding the final stop do exist; thus, the model would lead us to expect at least occasional deletion in these forms, although the rate should be very low. This prediction is confirmed in the present corpus, where we have found a low but nonzero deletion rate after /r/. Deletion after vowels and glides is known to occur in some English dialects, especially AAVE, and elsewhere we have found it to occur at a low rate in white speakers of North American English. (The absence of such reports in previous studies of this dialect is evidently a consequence of the investigators' assumptions rather than an empirical finding.)

The occurrence of deletion in the postvocalic environment causes us to prefer the probabilistic OCP constraint in (3) over the absolute structural constraint in (1), where a preceding C slot is a necessary condition for the rule. Note that the variable rule model can accommodate very low deletion rates by the addition of more disfavoring features. This is possible because vowels and glides are distinguished from coronal stops by more features than the ones studied here. But in the standard logistic model of constraint interaction, the rate should only reach zero when there is a specifically categorical constraint on the rule.

In view of the voicing effect, this model would further predict that, after vowels and glides, there should be more deletion of /d/ than /t/, because the vowel + /d/ sequences will have the clashing structure in (5). Although we have no quantitative data on this point, it conforms to our intuitions. Deletion seems more likely in a phrase like *made me* (cf. *He ma' me do it*) than one like *hate me* (cf. *Don't hate me*).

Finally, these results offer an explanation for the scattering of values found by previous studies for this factor group. First, the previous practice of conflating /n/ with /m, ŋ/ meant that the value obtained for the conflated class fluctuated from sample to sample, depending on the relative proportion of coronal and noncoronal nasals. Second, since some of the classes may not have been significantly different from each other (like stops and sibilants, and fricatives and laterals), their rank order would be expected to fluctuate due to random sampling error. And finally, there is the possibility that different speakers or dialects could have small differences in the values associated with the features, which would affect the relative positions of the segmental classes. Thus, a dialect with a higher value for [−cont] than for [+cor] would show higher deletion rates in stops and nasals relative to fricatives and liquids, whereas a dialect with a higher value for [+cor] would enhance deletion rates in /n/, /l/, and the sibilants relative to other segments.

Some aspects of this featural analysis receive independent confirmation in the work by Santa Ana (1996) on Chicano English. He broke out [cor] as a separate factor group, with results similar to ours (preceding coronals favor deletion over preceding noncoronals by about a 2:1 ratio). But his analysis relied on a continuous sonority scale and did not provide a comparable analysis treating the manner features separately. The rank order of his sonority classes suggests that the effect of continuancy may outweigh sonorancy in Chicano English.

THEORETICAL IMPLICATIONS

For present purposes, the most significant consequence of the approach that we have taken here is that it states the OCP as a quantitative preference rather than as a categorical constraint. It disfavors same-tier identical adjacent sequences, but does not absolutely rule them out. It has a gradient or probabilistic effect, so that the more similar adjacent segments are, the less likely that both will surface. Similarity in this treatment is a function of shared distinctive features, particularly the features [cor], [son], [cont], and [voice].

Additional support for various elements of this approach can be found in several of the studies we have cited. The fact that the OCP does not require absolute identity of the adjacent segments has been noted even by linguists who treat it as a categorical principle. Thus, Borowsky (1986) noted that the surface vowel in English *-ed* and *-es* suffixes occurs not just after /d,z/, where the vowel prevents an identical sequence, but also after /t,s/, where there is a voicing distinction, and (for *-es*) after all the other sibilant fricatives and affricates, which differ from the target by minor place and manner features. This is the kind of evidence that has led to the hypothesis that the OCP is not restricted to fully identical segments or autosegments, but that it can target sequences that are identical only in part, as for example with respect to a sequence of features on some particular tier.

The position that the OCP is gradient—in that it is sensitive to the degree of similarity between targeted elements—has been argued by linguists such as Pierrehumbert (1993), Buckley (1993), and Berkley (1993). Pierrehumbert (1994) argued, as we do, that this gradiency is evidence for inherent variability. The specific features implicated in this calculation of similarity are principally place features, but there is some evidence that manner features also play a role. Thus, Berkley found evidence that the sonorant/obstruent distinction was a significant factor in her data, and Santa Ana's results for sonority effects also implicated continuant/noncontinuant.

The OCP is therefore no longer an "obligatory" contour principle, although the acronym could be preserved by renaming in the optional contour preference. This gives it the character of what Goldsmith (1990) termed a soft universal, defining "a marked configuration that grammars tend to avoid though they do not necessarily always succeed in doing so" (Kenstowicz, 1994:534). Such phenomena have received increasing recognition in theoretical work in phonology in recent years. Perhaps the best known work in this vein is the optimality theory (OT) of Prince and Smolensky (1993) and others. In this theory constraints on phonological form are explicitly treated as violable and rankable. This makes them extremely similar to constraints in the variable rule model. Because optimality constraints can be violated, they are by definition variable, not categorical, and their rankability means that some are stronger than others, which implies a quantification of magnitude analogous to the probabilities used in VARBRUL analysis. A number of recent works have taken advantage of this similarity to interpret variable results in an OT framework. Although we are encouraged by this convergence of approaches, we note that, in works such as Berkley (1994) and Guy (1994), the OT model becomes problematic for a gradient OCP constraint sensitive to partial featural similarities, such as the one proposed here.

Finally, it is important to note that the approach taken here to the preceding segment effect on coronal stop deletion has implications for the analysis of this process within the framework of lexical phonology that is presented in Guy (1991, 1992). It is argued there that the rule in (1) operates at all levels of the phonology (both lexical and postlexical), and that underived, monomorphemic words are exposed to the rule more times than derived forms (e.g., past tense verbs), which only satisfy the structural description of the rule late in their derivational history. This explains the higher deletion rate in the former category. As a consequence of these repeated applications of the rule, the conditioning effects of the preceding segment are shown in Guy (1992) to be magnified, whereas constraints that apply only once at the postlexical level are relatively attenuated. Since the present work argues that the preceding segment effect is really a consequence of the OCP, we must conclude that the OCP also operates throughout the phonology, without restriction as to lexical level. This is consistent with the conclusions of Borowsky (1986).

CONCLUSIONS

The OCP, treated as a variable constraint, offers a cogent and integrated account of the preceding segment effect on English coronal stop deletion. This account is superior to any previous explanation, subsumes a broader range of facts, and makes novel predictions that have been confirmed in the present study. Our finding that the quantitative effects of the four defining features are very close, perhaps identical in value, has a theoretically attractive interpretation in terms of the uniformity of the OCP effect across different features. All of these results suggest that this analysis is entitled to a strong supposition of truth at the present stage of our understanding. Therefore, one conclusion of this work is that a genuine explanation has been found for the preceding segment effect on coronal stop deletion. But what does this imply for the larger issue we raised at the outset? Does it say anything about whether variation is inside or outside of the grammar?

The most important feature of this analysis is that it relies on a single theoretical principle—the OCP—to account for both obligatory and variable constraints on English syllable structure. The OCP was originally postulated by theoretical phonologists to account for categorical phenomena. In English, it explains the categorical absence of geminates, the categorical distribution of the syllabic allomorphs of the *-ed* and *-es* suffixes, and various other invariant properties (cf. Borowsky, 1986; Lamontagne, 1993). These are not optional or variable statements; they are absolutely true of all relevant English lexical items. But, as we have seen, the OCP also provides the best explanation of a constraint on the variable process of /t,d/ deletion.

This implies that variability is governed by the grammar, and that the assumption of invariance is incorrect. If variability is exclusively the product of extragrammatical factors of performance, why and how would it be governed by a grammatical principle like the OCP? To account for the performance facts, we would be forced to postulate a second version of the OCP outside the grammar, one which would be separate but equal to the categorical OCP. In such a model, it would be a theoretical surprise, a random coincidence, that the two OCPs existed and were so parallel in operation and effect. Normal scientific practice requires us to view such a coincidence with extreme suspicion. Far more attractive would be a model that has only one OCP governing both categorical and variable phenomena. Such a model incorporates Weinreich, Labov, and Herzog's hypothesis of inherent variability: a unified probabilistic grammar that sees orderly heterogeneity in language use as an extension of the same processes that generate categorical outputs. To maintain the alternative assumption, that competence and grammar are invariant, requires an unmotivated fragmentation of what is manifestly a unified phenomenon and leaves unexplained the data considered here.

NOTES

1. Similar results, but different theoretical conclusions, are presented in McCarthy (1993).
2. Older corpora available to us were not coded in such a way as to make the required coding distinctions recoverable. Hence, a fresh database was necessary.

REFERENCES

Berkley, Deborah M. (1994). Variability in Obligatory Contour Principle effects. In *Proceedings of CLS 30*. Chicago: Chicago Linguistic Society.

Borowsky, Toni. (1986). *Topics in the lexical phonology of English*. Doctoral dissertation, University of Massachusetts, Amherst.

Buckley, Eugene. (1993). Tigrinya root consonants and the OCP. Manuscript, University of Pennsylvania.

Cedergren, Henrietta, & Sankoff, David. (1974). Variable rules: Performance as a statistical reflection of competence. *Language* 50:333–355.

Goldsmith, John. (1990). *Autosegmental and metrical phonology*. Oxford: Blackwell.

Guy, Gregory R. (1980). Variation in the group and the individual. In W. Labov (Ed.), *Locating language in time and space*. New York: Academic. 1–36.

______ (1991). Explanation in variable phonology: An exponential model of morphological constraints. *Language Variation and Change* 3:1–22.

______ (1992). Contextual conditioning in variable lexical phonology. *Language Variation and Change* 3:223–239.

______ (1994, October). *Violable is variable: Principles, constraints, and linguistic variation.* Paper presented at NWAVE-XXIII, Stanford University.

Kenstowicz, Michael. (1994). *Phonology in generative grammar.* Oxford: Blackwell.

Labov, William. (1969). Contraction, deletion, and inherent variability of the English copula. *Language* 45:715–762.

Lamontagne, Gregory. (1993). *Syllabification and consonant co-occurrence conditions.* Doctoral dissertation, University of Massachusetts, Amherst.

Leben, Will. (1973). *Suprasegmental phonology.* Doctoral dissertation, MIT.

McCarthy, John. (1986). OCP effects: Gemination and antigemination. *Linguistic Inquiry* 17:207–263.

______ (1988). Feature geometry and dependency: A review. *Phonetica* 43:84–108.

______ (1993). The phonetics and phonology of Semitic pharyngeals. In P. Keating & B. Hayes (Eds.), *Papers in laboratory phonology III.* Cambridge: Cambridge University Press.

Nesbitt, Chris. (1984). *The linguistic constraints on a variable process: /t,d/ deletion in Sydney speech.* BA Honors thesis, University of Sydney.

Pierrehumbert, Janet. (1993). Dissimilarity in the Arabic verbal roots. In *Proceedings of NELS 23.* Amherst, MA: University of Massachusetts.

______ (1994). Knowledge of variation. In *Proceedings of CLS 30.* Chicago: Chicago Linguistic Society.

Prince, Alan, & Smolensky, Paul. (1993). Optimality theory: Constraint interaction in generative grammar. Manuscript, Rutgers University and University of Colorado at Boulder.

Santa Ana, Otto. (1996). Sonority and syllable structure in Chicano English. *Language Variation and Change* 8:63–89.

Sankoff, David. (1978). Probability and linguistic variation. *Synthèse* 37:217–238.

Veatch, Thomas. (1991). *English vowels: Their surface phonology and phonetic implementation in vernacular dialects.* Doctoral dissertation, University of Pennsylvania.

Weinreich, Uriel, Labov, William, & Herzog, Marvin. (1968). Empirical foundations for a theory of language change. In W. Lehmann & Y. Malkiel (Eds.), *Directions for historical linguistics.* Austin: University of Texas Press. 95–195.

Wolfram, Walt. (1969). *A sociolinguistic description of Detroit Negro speech.* Washington, DC: Center for Applied Linguistics.

Yip, Moira. (1988). The Obligatory Contour Principle and phonological rules: A loss of identity. *Linguistic Inquiry* 19:65–100.