[This is a longer version of a paper to appear in *Language* 78.2, March 2002.]

The Unabridged **Feature Classes in Phonology**[*]

Jaye Padgett - University of California, Santa Cruz - May 2001

Abstract

This paper argues for a particular understanding of feature class behavior - the recurrent patterning together of certain phonological features, such as place of articulation and laryngeal features. The proposals build on the well known work of Feature Geometry in assuming the importance of feature classes in phonology, but differ in that features of a class are targeted directly and invidually by constraints (or rules), even when a feature class such as 'Place' is mentioned. Further, constraints mentioning feature classes are gradiently violable. Evidence for this view of feature classes comes from two sources. First, assimilation involving feature classes is sometimes only partially successful; an adequate understanding of such cases requires the proposed view of feature classes. Second, there are broad categories of feature class generalization that require it, including dissimilatory effects usually handled by the Obligatory Contour Principle. Overall, the proposals broaden the explanatory potential of the feature class idea due to Feature Geometry. At a more general level, the results here suggest that linguistic representations sometimes need to be reconsidered in the context of Optimality Theory (Prince and Smolensky 1993), since they can effectively function as inviolable constraints and so hinder our understanding of the more subtle kind of phenomena revealed by analyses employing gradiently violable constraints.

1. Feature classes. It is a familiar observation that certain phonological features pattern together recurrently across languages in phenomena such as assimilation, dissimilation, and reduction. Perhaps the phenomenon most frequently cited as an example of this is nasal place assimilation. In Kpelle (a Mande language of West Africa), for example, nasal place assimilation occurs as in (1).

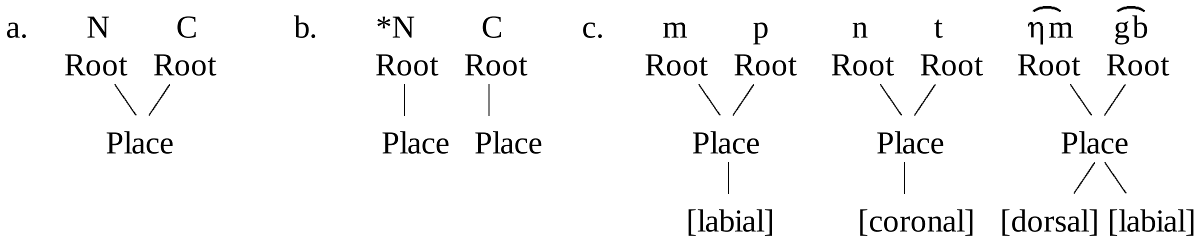
(1) Kpelle nasal place assimilation

[m]bolu	'my back'	[ŋ]gɔɔ	'my foot'
[ɱ]vela	'my wages'	[ŋ͡m]g͡bìŋ	'myself'
[n]duɛ	'my front'		

Nasal place assimilation as in Kpelle is very common, though languages differ in the precise conditions they place on it. Of interest here is the fact that it is frequently enforced over all places of articulation, as in the Kpelle case, and not over just one place of articulation; further, the rule affects only place of articulation features, rather than any random subset of phonological features. As the familiarity of the term 'place assimilation' implies, there seems to be a recurrent, unitary generalization covering all and only the places features.

How should this generalization about place features be understood? A well-known answer to this question is provided by Feature Geometry (Mohanan 1983, Mascaró 1983, Clements 1985, Sagey 1986, McCarthy 1988, and much subsequent work). Feature Geometry posits that the relevant features are grouped together under a node Place in a universal autosegmental representation; rules or constraints target this node directly, and thereby affect all of the place features indirectly. For the case at hand, suppose a constraint requires that a nasal consonant share Place with a following obstruent, making the partial representation in (2)a well formed and that in (2)b ill formed. The indirect result of this will be that the two consonants share whatever features Place dominates, as in the representative examples of (2)c. In this way Feature Geometry provides a simple, unitary characterization of nasal place assimilation.

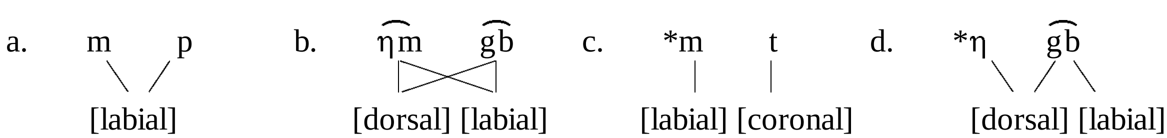
(2)



Given some further assumptions (see below), Feature Geometry more generally accounts for the fact that only a few privileged sets of features are distinguished by this recurrent behavior (e.g. place features, laryngeal features, vowel height features), rather than any arbitrary subset of phonological features. Let us call these distinguished sets feature classes.

This paper argues for a different understanding of feature class generalizations, called Feature Class Theory. As Feature Geometry does, Feature Class Theory assumes a limited inventory of feature classes such as Place, Laryngeal, Vowel Height, etc., and it allows rules or constraints to refer directly to these classes (as well as individual features), thus carrying over from Feature Geometry the advantage of a unitary characterization of nasal place assimilation and similar facts. Feature Class Theory is different, however, in its central claim that features are always affected directly and individually. Even when a rule or constraint targets a class such as Place or Laryngeal, it is the relevant place or laryngeal features themselves that are affected, rather than any higher order class node. Continuing with the example above, what constraints driving Kpelle nasal place assimilation require is that every feature that is a Place feature be shared in a nasal-obstruent cluster. Given this, forms such as (3)a-b are licit, while (3)c-d are not.[1]

(3)



If features are affected directly and individually, how is the unitary characterization of a class preserved under this view? Here I suggest an approach in which feature classes are simply defined according to properties that features have, understood as set-theoretic postulates in a way familiar from work in semantics: Place simply stands for the set {[labial], [coronal], [dorsal],...}; Laryngeal is {[voiced], [aspirated],...}; and so on. Equivalently, these features have the respective properties of placehood and laryngeality. These sets, rather than the class nodes, provide the essential unity of 'place' and 'laryngeal'.

There are two arguments for the claim that features are always targeted directly and individually. First, within the category of assimilations there are feature class generalizations missed otherwise. These derive from the possibility of partial class behavior: where constraints target a feature class such as Place, some members of the class conform to the constraint while others do not. In the language of Optimality Theory (Prince and Smolensky 1993), the relevant feature class constraints are gradiently violable. Such partial class behavior is surprising from the perspective of Feature Geometry. Given the assumptions of that theory, class nodes have always been interpreted to function in an all-or-nothing fashion, essentially as 'hard' constraints. The unfortunate result of this property is that evidence for feature class behavior can be missed.

Second, though Feature Geometry nicely handles many facts involving assimilation and reduction, it fails altogether to extend to certain other broad categories of feature class generalization. For example, though the recurring existence of dissimilatory constraints involving feature classes has been cited as an argument for feature geometry, the logic of Feature Geometry does not in fact extend to them. Rather, capturing these constraints requires essentially the Feature Class Theory understanding of feature classes.

The results here bear on very general questions concerning the role of constraints versus representations in linguistic theory. Linguistic representations are effectively 'hard' or inviolable constraints, or at least they are often effectively interpreted as such. In the context of Optimality Theory, in which constraints are violable, the question arises whether this property of representations is desirable. The findings here imply that it sometimes is not.

Section 2 reviews earlier work on feature classes and Feature Geometry. Sections 3 and 4 present two systematic kinds of partial class behavior that motivate Feature Class Theory, the first involving the feature class Color (the vocalic features [back] and [round]), the other Place. The first is particularly interesting, because it shows that without the Feature Class Theory point of view, evidence for a feature class can be missed altogether. Section 5 outlines a means of formalizing Feature Class Theory. Section 6 demonstrates that the theory straightforwardly handles certain processes apart from assimilation that are difficult for Feature Geometry. Section 7 provides general discussion, and a comparison of the two approaches to feature classes. Section 8 is the conclusion.

2. Feature classes and Feature Geometry. Cross-linguistic phonological patterns provide evidence for privileged subsets of phonological features that correspond roughly to traditional phonetic dimensions such as 'place of articulation' and 'laryngeal'. Though features were often classified according to these dimensions for expositional reasons, such classifications played no formal role in generative phonology before the mid-seventies, an inadequacy pointed out on occasion. In a discussion of Spanish nasal place assimilation, for example, Cressey (1974) observed of the rule in (4)a (from Harris 1969) that it seemed to miss the very generalization that nasals are homorganic with a following obstruent. He made this point more explicit by offering for comparison the rule in (4)b. This second rule requires agreement in voicing, continuancy, and only a subset of the place features. According to the feature-counting evaluation metric of Chomsky and Halle (1968), rules (4)a and (4)b are equally valued and hence should be equally likely, an incorrect result.

(4)

$$\text{a.}\quad [+\text{nasal}] \rightarrow \begin{matrix}\alpha\text{cor}\\ \beta\text{ant}\\ \gamma\text{back}\\ \delta\text{distr}\end{matrix} \;/\; \_\_\_\; (\#) \begin{matrix}+\text{obstr}\\ \alpha\text{cor}\\ \beta\text{ant}\\ \gamma\text{back}\\ \delta\text{distr}\end{matrix}$$

$$\text{b.}\quad [+\text{nasal}] \rightarrow \begin{bmatrix} \alpha\text{ant} \\ \beta\text{cor} \\ \gamma\text{voice} \\ \delta\text{cont} \end{bmatrix} \;/\; \_\_\_ \;(\#) \begin{bmatrix} +\text{obstr} \\ \alpha\text{ant} \\ \beta\text{cor} \\ \gamma\text{voice} \\ \delta\text{cont} \end{bmatrix}$$

To remedy the problem, Cressey proposed a 'point of articulation' abbreviatory device for use in rules, and wrote the rule for Spanish shown in (5). [αPA] was to be expanded as [αanterior, βcoronal, γdistributed,...], essentially standing in for all of the place of articulation features. (5) rather than (4)a was the representation of nasal place assimilation to be submitted to the evaluation metric (this was the point of abbreviatory devices); the lack of any similar abbreviatory device covering the features grouped in (4)b entailed that the latter was formally more complex. Furthermore, this result implied that nasal place assimilation should occur more often, or be more 'natural', than the process indicated by (4)b. Very similar proposals were made independently by Lass and Anderson (1975:151-6) and Lass (1976:260-3), who proposed that segmental matrices be divided into submatrices corresponding to 'articulation' and 'phonation' (Lass and Anderson) or 'oral' and 'laryngeal' (Lass). (See Anderson and Ewen 1987 for a review and further proposals.) These latter works motivated the idea with reference not only to place assimilations, but to other feature class generalizations such as the reduction /p t k/ → [ʔ], argued to be a formal reduction of the form [oral] → 0, an idea foreshadowing Clements (1985) and McCarthy (1988).

(5)

$$[+\text{nasal}] \rightarrow [\alpha\text{PA}] \;/\; \_\_\_ \;(\#) \begin{bmatrix} +\text{obstr} \\ \alpha\text{PA} \end{bmatrix}$$

These arguments demonstrate some strategic assumptions guiding Feature Geometry and most work leading up to it, and it is worth making them explicit.

(6) a. Assumption of Privilege

Within the full set of features, a limited number of universal and (according to many) phonetically motivated subsets are distinguished. Reference by rules to just these subsets is made simple (with the help of implementational assumptions of some kind)

b. Assumption of Simplicity and Naturalness

Simpler rule formulations correspond to more natural (or more common) phenomena

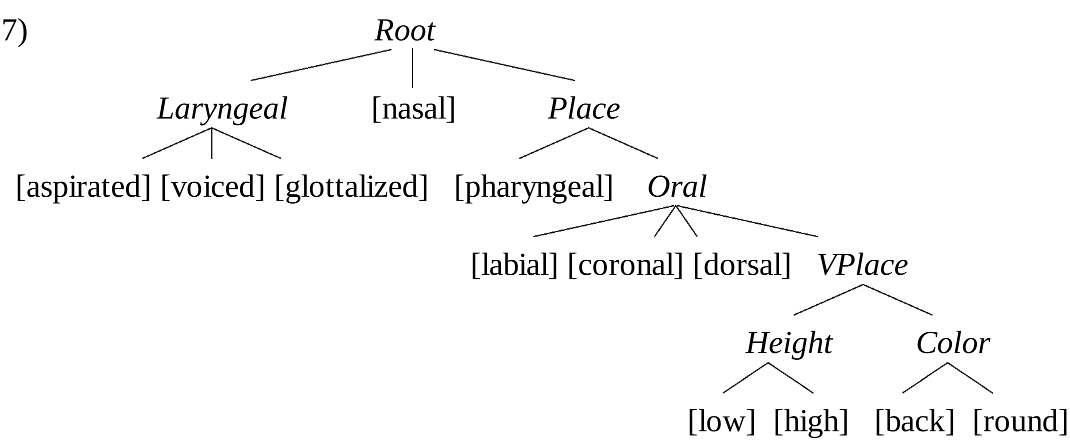
It should be clear that both assumptions are crucial to the intended arguments. It is the assumption of privilege that prevents us from stipulating opportunistically for one case that the features mentioned in (4)b exhaustively form a class called (say) Hodgepodge that can be referred to by a rule, rendering this rule as simple as that in (5)a. The assumption of simplicity and naturalness (an example of which is the evaluation metric of Chomsky and Halle 1968) is what in turn makes the connection to the typological facts.

Feature Geometry developed as a particular implementation of these ideas in the framework of autosegmental phonology (on the latter Goldsmith 1976). Considering again the

Spanish nasal place assimilation facts, Goldsmith (1981) proposed to understand the recurrent unity of place features by means of an autosegmentalized point of articulation feature bundle (or in his terms, 'segment'), also given a unitary characterization ('PA') in rule formulations. At about the same time Thráinsson (1978) proposed a similar understanding of laryngeal and supralaryngeal classes. However, the beginning of Feature Geometry is most clearly evident in Mascaró (1983) and Mohanan (1983), both of which posit universal hierarchically structured feature class representations encompassing all features (at least in principle). In Mohanan (1983) autosegmental operations can target one of three 'levels' of the segment: a terminal feature, a privileged bundle of features, or the entire segment at the superordinate 'Root' tier. These ideas are notably developed and elaborated into Feature Geometry by Clements (1985), and further by Sagey (1986) and McCarthy (1988). Though later works make proposals concerning new or revised feature classes, the basic premises of Feature Geometry are largely laid out in the cited works. For overviews of Feature Geometry, see Broe (1992) and Clements and Hume (1995).

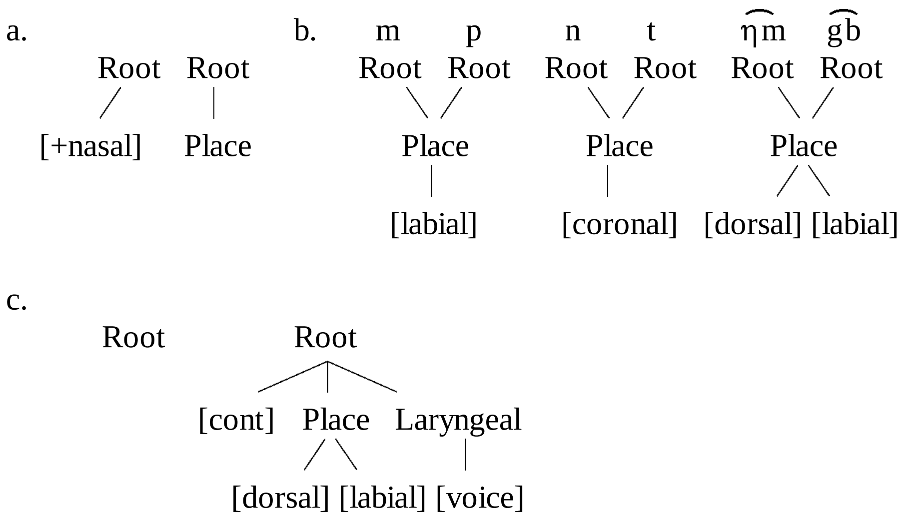
A representative Feature Geometry representation is shown in (7). This representation builds on the cited references in its incorporation of Place and Laryngeal classes, with the former dominating privative articulator features (on the laryngeal features shown, see Lombardi 1991 and Lombardi 1995). Since the focus here is largely on classes involving place features, this is the area of the geometry with the most detail provided. (7) represents a fairly standard synthesis of proposals of Steriade (1987), Hyman (1988), Clements (1991), McCarthy (1991, 1994), Odden (1991), Ní Chiosáin (1994a), among others. This geometry remains underdetermined in various respects. My concern here is not with the question what precise inventory of feature classes or hierarchical arrangement among classes is right (assuming there is one universal answer), nor with issues such as what features the theory should assume and whether they are privative, binary, etc. The focus is instead on the manner in which feature classes are implemented in the theory.

(7)



In Feature Geometry, the assumption of simplicity and naturalness relies on a conception of simplicity slightly elaborated to embrace autosegmentalism. In particular, the simplest rules are those that insert or delink one association line (Clements 1985, McCarthy 1988, building on the references cited above). The assumption of privilege is redeemed through the central notion of class nodes such as Place and Color in (7), which serve as key mediators of feature class behavior. Thus a rule of nasal place assimilation has the form shown in (8)a and produces outputs such as those shown in (8)b. (Alternatively, an output constraint requires that representations be as in (8)b.) In contrast, a rule would have to be much more complex to derive linking of exactly [voiced], [continuant], and two place features, since this is not a privileged class (i.e. in Feature Geometry there is no such class node). Assuming [cont] is immediately dominated by Root, as in McCarthy (1988), four separate association lines are needed, as shown in (8)c. (If [cont] is instead dependent on the place of articulation features, as in Padgett 1991/95a, then three lines are needed.)

(8)



Arguments for feature classes in Feature Geometry have been based largely on evidence that certain subsets of features recurrently fall under one generalization of assimilation, neutralization, or dissimilation. In terms of autosegmental constraints or operations, we find evidence that features spread, delink, or are restricted by the Obligatory Contour Principle together (McCarthy 1988).[2] In this context, class nodes are mediators of feature class behavior in the following sense: rules or constraints that are intended to target a feature class in fact refer to and affect directly the relevant class node, as shown above; the features making up that class are affected only indirectly, by virtue of being dominated by the class node.

Feature Class Theory, the view argued for here, assumes a different means of cashing in on the assumption of privelege. As we will see, a good deal of evidence suggests that features of a class are targeted directly and individually, rather than via intermediate nodes in a representation. (This fact has implications for our interpretation of the assumption of simplicity and naturalness too.) The next two sections present the evidence for this point of view. I return to a general comparison of the two theories afterward.

3. Feature class theory and color harmony.

3.1. Turkish. In many Turkic languages we find patterns of harmony involving both roundness and backness in vowels. The most familiar example is that of Turkish. The analysis of Turkish here draws especially on the work of Lees (1961), Haiman (1972), Clements and Sezer (1982), Hulst and Weijer (1991), Kirchner (1993), and references therein. The Turkish vowel inventory is given below.[3]

(9) Turkish vowel inventory

	[-back]		[+back]	
[+high]	i	ü	ɨ	u
[-high]	e	ö	a	o
	[-rnd]	[+rnd]	[-rnd]	[+rnd]

The data in (10), taken from Clements and Sezer (1982) and expanded, show suffix alternations due to vowel harmony. When a suffix vowel is high as in the accusative and genitive cases, it agrees with the preceding vowel in both backness and roundness. Otherwise, it agrees only in backness, as shown with the plural marker. These alternations are quite regular and productive across Turkish suffixes, with the exception of a few 'disharmonic' suffixes, e.g., the progressive suffix -iyor.

(10) Turkish vowel harmony

stem	acc.sg.	gen.sg.	nom.pl.	
ip	ip-i	ip-in	ip-ler	'rope'
kɨz	kɨz-ɨ	kɨz-ɨn	kɨz-lar	'girl'
yüz	yüz-ü	yüz-ün	yüz-ler	'face'
pul	pul-u	pul-un	pul-lar	'stamp'
el	el-i	el-in	el-ler	'hand'
sap	sap-ɨ	sap-ɨn	sap-lar	'stalk'
köy	köy-ü	köy-ün	köy-ler	'village'
son	son-u	son-un	son-lar	'end'

When suffixes are strung together, backness harmony applies throughout, as in ip-ler-in 'ropes (gen.)' and kɨz-lar-ɨn 'girls (gen.)'. The same is true of roundness harmony where high vowels are concerned. However, non-high vowels not only fail to become [+round] by roundness harmony, they also block the spreading of this feature to any following vowels, even if the latter are high, e.g., pul-lar-ɨn 'stamps (gen.)', yüz-ler-in 'faces (gen.)', cf. *pul-lar-un, *yüz-ler-ün. Suffix vowels following non-high vowels are systematically [-round].

There is some controversy over the issue of whether vowel harmony can be said to hold within roots, given the existence of many disharmonic roots largely (but not entirely) due to borrowings, e.g. politika 'politics', peron 'railway platform', mezat 'auction'. Clements and Sezer (1982), for example, conclude based on the large number of disharmonic forms that vowel harmony does not operate within roots. On the other hand Hulst and van de Weijer (1991), among others, assume that harmony applies within roots that appear harmonic, e.g. somun 'loaf', adɨm 'step', öküz 'ox', and treat disharmonic forms as exceptionally specified for backness and roundness features. The issues involving disharmony, patterns of disharmony, underlying forms, and the status of harmony within roots are complex. (For proposals and interesting discussion see also Kirchner 1993, Polgárdi 1994, and Inkelas, Orgun, and Zoll 1997.) In this paper, I take the conservative route of assuming suffix harmony only, as Clements and Sezer (1982) do, since the points to be made can be made based on suffix harmony, and hold regardless of the truth about root harmony.

Of interest here is the existence within a single language of both [back] and [round] harmony. Though Turkish provides the best known case of these harmonies in coexistence, they occur together in many other languages, as we will see. Indeed, while [back] and [round] harmonies coexist often, languages having harmonies involving [back] and [high], for example, or [round] and [ATR], do not. Why should this be so? I will show that this fact can be accounted for given the assumption reflected in (7) that [back] and [round] form a feature class. However, in order for this to be so, we must revise our notion of what a feature class is, and how it operates. Let us first consider the independent evidence for [back] and [round] as a feature class.

3.2. The feature class Color. There is a long history in phonology of classifying vowels according to a 'vertical' axis of vowel height on the one hand, as opposed to a 'horizontal' dimension of vowel quality—backness and roundness—on the other. Trubetzkoy (1939) and Jakobson and Halle (1956), for example, refer to this latter dimension as that of 'timbre' or 'tonality' (respectively). They motivate this distinction based on facts of acoustics, inventory markedness, and alternation.

The acoustic facts are well known: the acoustic correlates of articulatory backness and roundness involve essentially the second and higher vowel formant frequencies, or their frequencies in relation to the first formant, while those of vowel height involve the first formant frequency (see Stevens 1998:283 for a recent summary). Perhaps more telling, the perceptual effect of the second and higher formants can be modeled well by means of a single formant frequency, called 'F2 Prime' by Carlson et al. (1970), which is mathematically derived based on values of the first three formants, and which mimics aspects of the acoustic to auditory transformation done by the human ear to the higher formants. In other words, [back] and [round] share a unitary perceptual correlate, distinct from that for height. Articulatory backness can therefore enhance the acoustic effect of roundness, and vice versa (Stevens, Keyser, and Kawasaki 1986, Stevens and Keyser 1989), and in fact these articulatory dimensions are traded off in the pronunciation of vowels across dialects, speakers, and phonetic contexts (de Jong 1995). This acoustic and perceptual unity of backness and roundness was expressed in the feature system of Jakobson, Fant and Halle (1952) through the feature [grave]. (It should be noted, however, that a theory having [grave] and no [back] and [round] could not account for the many facts where these features crucially diverge, as in the case of Turkish vowel harmony.)

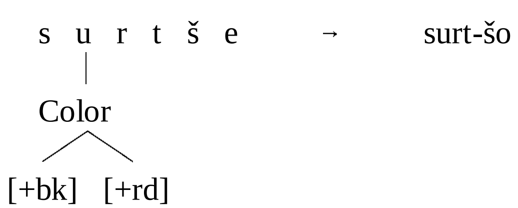
Many phonologists following Clements (1985) and Sagey (1986) assume that feature classes have a phonetic basis, or at least should be broadly compatible with what we know of phonetic processes. However, work on feature classes within phonology has largely been based on phonological patterns. Moving to vowel inventory patterns, for example, a well known generalization is that vowels tend to contrast more in the backness/roundness dimension the higher they are, so that low vowels are least likely to contrast in this way, and high vowels most likely, giving the familiar triangularity of vowel systems. This is a phonological generalization that treats [back] and [round] as a class. Schane (1972, 1984) notes the phonetic and phonological relationship between backness and roundness, while Ewen and Hulst (1988) and Hulst (1988) posit a structural grouping of the elements I and U in a variant of Dependency Phonology theory (see Anderson and Ewen 1987 and references therein), based on inventory considerations.

In the realm of alternations, many researchers have noted that [back] and [round] tend to assimilate or harmonize within the same language, and have taken this to be evidence that these features are in some kind of close structural relationship. Archangeli (1985) makes the important observation that, while there are languages having [back] harmony and [round] harmony, there are few if any languages having [round] harmony and [high] (or [low]) harmony, or [back] harmony and [high] (or [low]) harmony. She accounts for this in part by grouping [back] and [round] together in an autosegmental representation in which they are dependent on height features. Mester (1986) analyzes Kirghiz [back] and [round] harmony in terms of a dependency relation between these features themselves (an idea Archangeli also explores), thus making of them a phonological constituent. Similarly, Hulst and Smith (1987) and Hulst (1988) argue for such a grouping in accounting for [back]/[round] harmonies (see also Selkirk 1991a). Odden (1991), working within Feature Geometry, argues for a node [back]/[round], based on a range of cases in which these features spread together in what appears to be a single phonological rule. Examples come from diverse languages, including Eastern Cheremis (Uralic), Tunica (Gulf, North America), Fe?-Fe? Bamileke (Niger-Congo), and Wikchamni Yokuts (California). Data from Eastern Cheremis are given below, all involving the 3rd singular possessive suffix -že.

(11) Eastern Cheremis [back]/[round] assimilation

a.	surt-šo	'his house'	b.	üp-šö	'his hair'
	boz-šo	'his wagon'		šör-žö	'its milk'
	kornə-žo	'his way'		pörtəštə-žö	'in his house'
c.	kit-še	'his hand'			
	bokten-že	'beside it'			
	šužar-že	'his sister'			

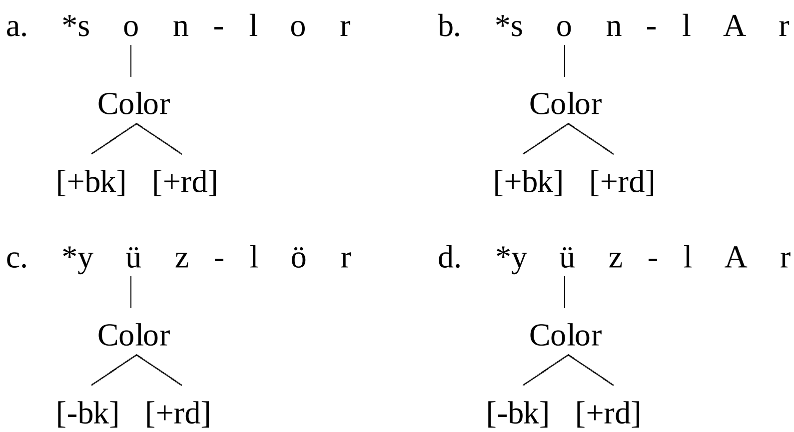
Odden assumes that the suffix is underlyingly -že, since this alternant occurs following [i,e,a,ə]. ([ə] can also occur between a trigger and target vowel, as seen above, and is therefore treated as transparent to spreading by virtue of lacking feature specifications.) We therefore require a means of spreading both [back], to account for the forms in (11)a, and [round], to explain the forms in (11)b. All of Odden's examples are like Cheremis in having inventories, and alternations, that require both of these features to achieve the account: one cannot say, for example, that front vowels are redundantly unrounded. Odden also argues here (and for other cases) that this is one rule, and not separate rules of [back] and [round] spreading: both spreadings have the condition that only non-low vowels are targeted, and words that are an exception to [back] spreading are also an exception to [round] spreading. Odden therefore posits a rule spreading a constituent grouping [back] and [round], as in (12). Following Donegan (1978) and others, I use the term Color to designate this feature class.

(12)


To summarize, there is a good deal of independent evidence converging from both phonetic and phonological considerations that [back] and [round] form a feature class. In this light, languages such as Turkish raise an intriguing question: can they be understood as instances of Color spreading too? As noted above, some researchers have taken it to be significant that [back] and [round] harmonize within a single language recurrently. However, no one has suggested that [back] and [round] harmony as in Turkish could be genuinely united in any way, for example in a single process of Color harmony, just as in Eastern Cheremis. Rather, research on Turkish invariably assumes two separate processes of harmony, implicitly or explicitly: one of [back], and one of [round]. It is particularly striking that no one working within Feature Geometry has taken the facts of Turkish as evidence for a Color node, given the precedent for this feature class.

It is not difficult to see why, however. As we have seen, Turkish is unlike Cheremis, and other cases examined by Odden (1991), in that [back] harmony and [round] harmony do not apply identically within the language; while [back] harmony targets all vowels, [round] harmony targets only non-high vowels. In this sense we have 'two separate rules'. Consider an example such as son-lar 'ends': taking harmony to involve Color node spreading, we would wrongly entail assimilation of [round] by the non-high vowel as well, as shown in (13)a. If Color spreading failed due to an independent constraint prohibiting this [round] spreading (see below), then we would wrongly fail to spread [back] as well, leaving the target vowel underdetermined for this feature, (13)b. We cannot avoid this outcome by assuming a default [back] value for non-high vowels, since the same problem arises whatever the value of [back]: compare yüz-ler 'faces' in (13)c-d. That is, [back] values are determined by harmony. The essential point is that Feature Geometry does not lead us to expect that Color harmony might produce an intermediate outcome for certain forms, the spreading of [back] but not [round]. Therefore, two rules are written.

(13)



This is a problem, because to assume separate rules of [back] and [round] harmony implies incorrectly that their cooccurrence within Turkish is no more likely than a cooccurrence of [back] and [high] harmony, or [round] and [nasal] harmony. Elaborating on Archangeli (1985)'s observation above, the crucial fact is that [back] harmony and [round] harmony occur together in many languages—even when subject to different conditions as in Turkish—to a degree that other logical possibilities do not.[4] How can this fact be explained?

The problem that Turkish presents is not at all isolated. Backness and roundness harmony occur together throughout the Turkic language group. For example, Korn (1969) surveys approximately 15 Turkish languages having both harmonies, including Yakut, Shor, Kazakh, Khakass, Tartar, and Uzbek. Backness and roundness harmony together are a feature of Altaic and Uralic languages more broadly. Hungarian is one well-known example from the latter group (see for example Vago 1980). Mongolian provides an example from Altaic outside of Turkic. In Classical Mongolian there was harmony of both [back] and [round], though [round] harmony had effects limited to roots. Khalka Mongolian has [round] harmony along with [ATR] harmony; however, the [ATR] alternation resulted from a reanalysis of a former system of [back] harmony. (See Svantesson 1985 on Mongolian.) Outside of these language groups, the two features both harmonize in Djingili and Nyangumarda, languages of Australia, (Hulst and Smith 1985). In all of the cases just mentioned, [back] harmony and [round] harmony apply under different conditions, just as in Turkish.

We are therefore presented with a problem. The cross-linguistic evidence suggests that there is a privileged status to the coexistence of [back] and [round] harmonies. The logic of feature classes therefore suggests that we capitalize on an independently motivated feature class Color in order to explain this fact. Yet how can [back] and [round] harmonies reflect a unitary generalization in languages such as Turkish where they apply under different conditions? In what follows I will show how this problem can be resolved.

3.3. Turkish as Color harmony. In Turkish words such as the accusative ip-i 'rope', köy-ü 'village', kɨz-ɨ 'girl', and son-u 'end', the root and suffix vowels agree in both [back] and [round] values. Such forms would be straightforwardly handled by a Color spreading rule such as that assumed for Eastern Cheremis. The same is true of words where both vowels are unround, even if the suffix vowel is non-high, such as kɨz-lar 'girls', and sap-lar 'stalks'. As we have seen, the problem arises with forms such as son-lar 'ends', and yüz-ler 'faces'. Even in these forms, however, there is a sense in which Color harmony does partially succeed: though [round] does not harmonize, [back] still does. If Turkish harmony is indeed Color harmony, then it has succeeded only partially in these cases; we could call this partial class behavior.

Partial class behavior finds a natural home within the framework of Optimality Theory (Prince and Smolensky 1993). A claim to be pursued here is that constraints mentioning feature classes like Place and Color are gradiently violable in the sense established by Optimality Theory, and partial class spreading is one result of this gradient violability. That is, partial class spreading is simply the (minimal) violation of the harmony requirement. As we will see, the other key idea is that features are always affected directly and individually by constraints requiring spreading, deleting, and so on. This is so even though constraints themselves can mention classes of features such as Color. How precisely this idea should be formally implemented is a question taken up in section 5. The first step is to show its relevance to the kind of typological problem just laid out. The following account assumes an understanding of Optimality Theory. Readers wanting more background and comprehensive discussion are referred to Prince and Smolensky (1993), Kager (1999), and McCarthy (to appear).

To begin we require a constraint enforcing assimilation, or feature spreading. Padgett (1995b) assumes a very general spreading constraint requiring that a feature be linked to (or be a property of) every segment in some domain, such as the prosodic word. This idea is made precise in a formulation by Walker (1998), on which the following informal statement is based. (14) targets the feature class Color, the central point of interest here.

(14) SPREAD(COLOR, PWD): For all color features *f* in a prosodic word, if *f* is linked to any segment, it is linked to all segments

The formulation of Spread assumed is obviously very general. I assume that potentially unwanted effects of this general constraint are reined in by constraints on locality, segment wellformedness, faithfulness (see below), and so on.[5]

One of the constraints reining in Spread is the faithfulness constraint Ident, defined below, where F is any feature (see McCarthy and Prince 1995). For our purposes, it will always be clear that identity of color features is at stake, hence I will simply refer to Ident.

(15) IDENT(F): Let S_I and S_O be corresponding segments of the input and output. Then if S_I is specified [αF], S_O is specified [αF], where α is +,-, or 0.

Spread(Color) and Ident can conflict: the former constraint requires harmony in forms such as ip-i and son-u, entailing that the accusative suffix surface as one of [i,ü,ɨ,u]. No matter what we assume about this suffix's underlying representation, there will be a change of features in many cases during the derivation. Suppose we assume (non-crucially) that suffix color features are unspecified.[6] Then input forms such as /son-I/ and /köy-lAr/, for example, where /I/ and /A/ are high and non-high vowels respectively, violate Ident in surfacing as son-u and köy-ler: the suffix in son-u is [+back, +round], and that in köy-ler [-back, -round], while their respective inputs were not specified for color at all. Since spreading of color features occurs in Turkish, leading to alternations in suffixes, Spread(Color) must outrank Ident. All of this is shown in the tableau in (16). To simplify notation, feature specifications are indicated not by association lines but by bracketing. The domain of [back] (B) is indicated by square brackets, that of [round] (R) by curly brackets. The notation used implies that consonants bear the color features also. (See Ní Chiosáin and Padgett to appear and references therein.) However, the points here do not change whether this is true or not, and for simplicity Ident and Spread violations are counted only with respect to vowels.

Candidate (16)a is identical to the input, and so satisfies Ident. However, it violates Spread, since harmony has not occured. There are two violations, because both [+back] and [+round] have failed to link to the suffix vowel. This candidate's suffix vowel lacks color specifications in the output, a state of affairs we can assume to be ruled out independently by high ranking constraints; from here on I consider only surface forms that are fully specified. Candidate (16)b satisfies Spread completely, and therefore violates Ident twice (the suffix vowel differs from its input in both [back] and [round] specifications). Candidates (16)c-d each harmonize one feature only ([back] and [round] respectively). Assuming full specification, values for [back] and [round] that are not determined by harmony must be inserted separately. The values chosen here are arbitrary. Finally, (16)e is like (16)a in failing to harmonize altogether; however, the suffix vowel is fully specified. Spread is violated four times, because there are four feature specifications in this form that are linked to one vowel rather than both. (16)b is chosen as optimal.

(16)

UR: /son-I/	Spread(Color)	Ident
a. $\{[\text{son-}]_{+B}\}_{+R}$-I	*!*	
b. ☞ $\{[\text{son-u}]_{+B}\}_{+R}$		**
c. $[\{\text{son-}\}_{+R}\{\text{ɨ}\}_{-R}]_{+B}$	*!*	**
d. $\{[\text{son-}]_{+B}[\text{ü}]_{-B}\}_{+R}$	*!*	**
e. $[\{\text{son-}\}_{+R}]_{+B}\ [\{\text{ɨ}\}_{-R}]_{+B}$	*!***	**

With this basic understanding of harmony in hand, we can return to the problem of partial class behavior. Important to the account is a well-known markedness generalization noted already (Trubetzkoy 1939): color feature contrasts are most favored in high vowels, less so in mid vowels, and least of all in low vowels. Further, environments of vowel reduction tend to display less marked inventories. Haiman (1972) and Hulst and Weijer (1991) note that the reduced vowel inventory derived by Turkish vowel harmony, shown in (17), should be viewed as less marked in this sense.

(17) Vowels produced by Turkish harmony

i	ü	ɨ	u
e		a	

In practice, this reduction has sometimes been enforced through a stipulation on roundness harmony itself (e.g., Lees 1961, Clements and Sezer 1982). Yet there is an independent need in the theory for an account of the relative markedness of non-high round vowels, as evidenced by patterns in languages involving both underlying and reduction inventories, and it is therefore greatly preferable to attribute this property of roundness harmony to the interaction of an independent constraint, factoring the height stipulation out of the harmony generalization altogether, as advocated by Haiman (1972). Putting aside further questions concerning the best account of the cross-linguistic height-color interaction, for convenience I will refer to a constraint *[+round, -high] (Kirchner 1993, Kaun 1995, abbreviated *[+rnd, -hi]).

Applied to the analysis at hand, the idea is that the ranking between the two constraints *[+rnd, -hi] >> Spread(Color) is the source of partial class spreading in Turkish. Consider the next constraint tableau, whose candidates are parallel in structure to those of (16), except that the candidate with no suffix vowel features is omitted. Candidate (18)a, with perfect harmony, is no longer optimal, because by spreading [+round] to a non-high vowel it violates *[+rnd, -hi] twice. Candidates (18)b-c spread only one feature, and (18)d none at all. Because (18)c spreads [+round], it again violates *[+rnd, -hi] twice. Of (18)b and d, both violate Spread, but (18)b violates it less and is therefore the output.

(18)

UR: /son-lAr/	*[+rnd, -hi]	Spread(Color)	Ident
a. {[son-lor]$_{+B}$}$_{+R}$	**!		**
b. ☞ [{son-}$_{+R}${lar}$_{-R}$]$_{+B}$	*	**	**
c. {[son-]$_{+B}$[lör]$_{-B}$}$_{+R}$	**!	**	**
d. [{son-}$_{+R}$]$_{+B}$ [{lar}$_{-R}$]$_{+B}$	*	***!*	**

It is important to note again the manner in which the constraint Spread(Color) is assumed to work. (19) shows the output forms of (16) and (18) in a more conventional autosegmental notation. As the form of the candidate outputs makes clear, there is no feature geometric Color node at work, such as Odden (1991) assumes for cases such as Eastern Cheremis. Instead, the features [back] and [round] are manipulated directly and individually. Yet they are manipulated by force of Spread(Color), which must therefore have some means of 'recognizing' these features as color features, a point I return to in section 5.

(19)

In (18) even the optimal candidate violates *[+rnd, -hi], due to the root vowel [o]. It is important therefore to consider candidates that lack this initial [+round] specification. As the Turkish facts make clear, the restriction against non-high round vowels is not respected in initial syllables, or arguably, in roots, given many historical borrowings such as melankoli. Why should this be? Following Beckman (1997, 1998), I assume a differentiation of faithfulness constraints into two classes: those restricted to certain phonological or morphological categories on the one hand, and those that are not (such as Ident assumed here). The former are argued to involve a limited universal inventory of positions that are of particular phonetic prominence or psycholinguistic salience. As noted earlier (section 3.1), I assume that harmony does not obtain within roots, given the larger range of vowel contrasts found there (due to many disharmonic forms). Therefore the priveleged position I have in mind is the morphological root, and the relevant positional Ident constraint is given below. The reader is referred to Beckman and references therein for extensive motivation of the root as a cross-linguistically priveleged position.

(20) IDENT$_{RT}$(F): Let S_I and S_O be corresponding segments of the input and output, where S_O is in the Root. Then if S_I is specified [αF], S_O is specified [αF], where α is +,-, or 0.

Putting together the rankings justified so far, we have *[+rnd, -hi] >> Spread(Color) >> Ident. Since the vowels [o,ö] are preserved in roots by Ident$_{Rt}$, this constraint must be highest ranking. This overall ranking is demonstrated in (21). Candidates (21)a,b repeat (18)a,b. These should be compared to (21)c, which satisfies Spread and *[+rnd, -hi] completely. This is achieved by failing to preserve the [+round] sprecification of the root vowel, however, violating Ident_{Rt}. It should be noted that Ident_{Rt} accounts not only for the occurrence of [o,ö] in roots, but for the fact that harmony is root controlled: spreading from suffix to root will always violate this constraint. (See note ?.)

(21)

UR: /son-lAr/	Ident_{Rt}	*[+rnd, -hi]	Spread(Color)	Ident
a. $\{[\text{son-lor}]_{+B}\}_{+R}$		**!		**
b. ☞ $[\{\text{son-}\}_{+R}\{\text{lar}\}_{-R}]_{+B}$		*	**	**
c. $\{[\text{san-lar}]_{+B}\}_{-R}$	*!			***

Compare (21) to (22), which applies the entire analysis so far to the form son-u once again. Of particular interest here are candidates (22)a,b: obviously, the constraint *[+rnd, -hi] plays no role in governing harmony when a suffix vowel is [+high]. The analysis therefore predicts correctly that Color harmony will succeed entirely in such cases.

(22)

UR: /son-I/	Ident_{Rt}	*[+rnd, -hi]	Spread(Color)	Ident
a. ☞ $\{[\text{son-u}]_{+B}\}_{+R}$		*		**
b. $[\{\text{son-}\}_{+R}\{\text{ɨ}\}_{-R}]_{+B}$		*	*!*	**
c. $\{[\text{san-ɨ}]_{+B}\}_{-R}$	*!			***

As noted earlier, Color harmony completely succeeds also in forms like kɨz-lar 'girls', in which both vowels are unround. The constraint *[+rnd, -hi] will have no effect in these words either. This is shown in tableau (23).

(23)

UR: /kɨz-lAr/	Ident_{Rt}	*[+rnd, -hi]	Spread(Color)	Ident
a. ☞ $\{[\text{kɨz-lar}]_{+B}\}_{-R}$				**
b. $[\{\text{kɨz-}\}_{-R}\{\text{lor}\}_{+R}]_{+B}$		*!	**	**
c. $\{[\text{kuz-lor}]_{+B}\}_{+R}$	*!	*		***

For completeness, (24) shows how the analysis handles forms like pul-lar-ɨn 'stamps (gen.)'. In such forms, all vowels following the non-high suffix vowel are predictably unround. This follows from harmony: since the first suffix vowel is forced to be [-round] by *[+rnd, -hi], the best option is to spread this feature rightward as well, as in candidate (24)b. Compare this to the case of 'transparency' in (24)c, in which the final vowel is [+round] just as the initial vowel is. In this form, it bears its own [+round] specification, however, and this leads to more violations of Spread. An alternative true transparency scenario, in which the [+round] value of the root actually spreads to this final vowel while skipping the medial vowel, is assumed to be ruled out by whatever constraints prohibit such vowel skipping, if such a representation is allowed in the theory at all (see Walker 1998 for recent discussion).

(24)

UR: /pul-lAr-In/	$Ident_{Rt}$	*[+rnd, -hi]	Spread(Color)	Ident
a. {[pul-lor-un]$_{+B}$}$_{+R}$		*!		****
b. ☞ [{pul-}$_{+R}${lar-ɨn}$_{-R}$]$_{+B}$			***	****
c. [{pul-}$_{+R}${lar-}$_{-R}${un}$_{+R}$]$_{+B}$			****!**	
d. {[pɨl-lar-ɨn]$_{+B}$}$_{-R}$	*!			*****

Summing up, partial class behavior in Turkish, and the concomitant partial violation of Spread(Color), is due to the force of a higher ranked and conflicting constraint, *[+rnd, -hi]. Where this constraint plays no role Color harmony is complete. It should be clear from all of this that a grammar having the reverse ranking Spread(Color) >> *[+rnd, -hi] derives Color harmony in all suffixes, including non-high ones, giving forms such as son-lor, yüz-lör. This is the more familiar case exemplified by Odden (1991) of total class spreading. Such spreading occurs in the Turkic language Kirghiz as well (Korn 1969, Kaun 1995). It has no special status in my terms next to partial class spreading, but simply represents the case where no higher ranking constraint interferes with Spread.

In forms such as son-lar, [back] and [round] harmony have a particular way of behaving differently: both vowels originate in the root vowel, but they spread differently. Turkish also presents cases having just the opposite arrangement: [back] and [round] differing not in what they target, but where they originate. They involve disharmonic stems ending in palatalized [l^y], as in (27)a (data from Clements and Sezer 1982). There is a contrast in Turkish stem-finally between the plain and palatalized lateral; compare (27)b. The palatalized lateral not only bears the feature [-back], but triggers [back] harmony; this is why the accusative suffix is [-back] in the forms of (27)a. What is interesting is that the suffix is targeted as expected by the [+round] specification of the root-final vowel in (27)a, even though it is obviously not affected by the [+back] sprecification of that vowel. Compare these forms to those in (27)b, which behave as expected according to my earlier analysis.

(25) Turkish harmony triggered by l^y

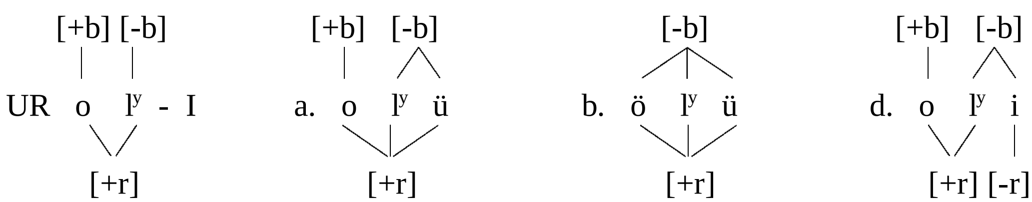
a.	usul^y	usul^y-ü	'system'
	petrol^y	petrol^y-ü	'petrol'
	sual^y	sual^y-i	'question'
b.	okul	okul-u	'school'
	karakol	karakol-u	'police station'
	tʃatal	tʃatal-ɨ	'fork'

The treatment of these forms here translates ideas of Clements and Sezer (1982) in its essential respects, except for my reliance on a unitary generalization of color spreading. The account is shown in the tableau below. In order to simplify the discussion only the last three segments of the form petrol^y-ü 'petrol (acc.)' are considered. In order to aid the reader in interpreting the representations, the underlying form and candidates are also given in autosegmental notation below the tableau (except for (26)c, which is much like (26)b). Since now we are explicitly considering the role of a consonant in harmony, the lateral counts for violations of Spread and Ident.[7] Candidate (26)a, the grammatical form, violates the harmony requirement three times. (The [+back] specification fails to spread to two segments, [l^y] and [ü]; the [-back] specification fails to spread to [o].) In spite of faring much better on this score, candidates (26)b,c are worse overall, because they sacrifice underlying color specifications of the root, in violation of Ident_{Rt}. Candidate (26)d preserves root specifications, as does (26)a. Yet it represents the failure to spread as fully as possible (subject to higher-ranking constraints); given the minimal gradient violability of Spread, this is not a possible outcome, and (26)a wins.

(26)

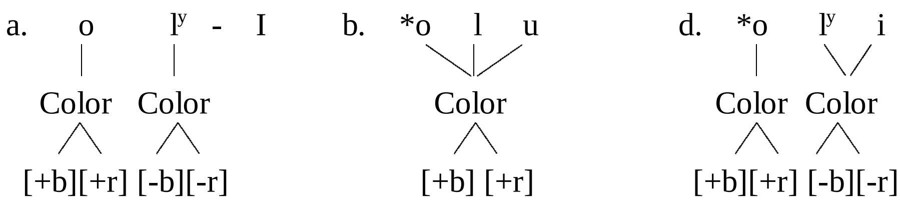
UR: /(petr)ol^y-I/	Ident_{Rt}	*[+rnd, -hi]	Spread(Color)	Ident
a. ☞ $\{[o]_{+B}[l^y\text{-ü}]_{-B}\}_{+R}$		*	***	**
b. $\{[\text{öl}^y\text{-ü}]_{-B}\}_{+R}$	*!	*		****
c. $\{[\text{ol-u}]_{+B}\}_{+R}$	*!	*		****
d. $\{[o]_{+B}\ [l^y\}_{+R}\ \{\text{-i}]_{-B}\}_{-R}$		*	****!**	**

Underlying representation and candidates in autosegmental notation:



This kind of example causes the same sort of problem for Feature Geometry, in which Color harmony is seen as the linkage of a superordinate feature geometric node Color, as do the earlier cases of partial class spreading. Consider the underlying form in (27)a. Spreading of the first Color node incorrectly distributes the [+back] and [+round] values of the stem vowel across all three final segments, as in (27)b. On the other hand, spreading of the second Color node forces both the [-back] and [-round] values of [l^y] onto the suffix, (27)c. There is no way that spreading of the node Color will distribute the features as needed (barring complicated and unnecessary repair mechanisms generally argued against by Optimality Theory, see section 7.2).

(27)



To conclude this section, let us consider one final point about the typology of color harmonies. In languages having both [back] and [round] harmony, [round] harmony is often restricted in some way, while [back] harmony rarely is. Korn (1969)'s survey of Turkic languages classifies them according to the degree to which roundness harmony is restricted. What emerges from his survey is that the patterns of restrictions found are far from idiosyncratic or random, but rather can be understood in terms of a few principles of phonological wellformedness, likely grounded in phonetic constraints. Indeed, the fact that [round] is singled out over [back] in itself suggests that systematic forces are at work, rather than randomly working conditions. Kaun (1995), building on Korn (1969) and Steriade (1981), among others, notes the following generalizations, and proposes a set of constraints intended to capture them. Turkish is a case of (28)a.

(28) Roundness harmony can fail when

a. it would create a non-high round vowel
b. it would involve linking between vowels of different height
c. the trigger is high
d. the trigger is back

For our purposes, the important point is that the cross-linguistic patterns of harmony are subject to systematic conditions, and that the latter are plausibly viewed as reflecting general, independently motivated constraints on wellformedness. Pursuing the logic seen above, this suggests that these other languages should indeed be handled analogously to Turkish. First, the relevant restrictions should be factored out of any harmony constraints (or rules) and attributed to independently working constraints. This factoring out of conditions is a familiar strategy in phonology. The novel point here is the bearing this has on our understanding of feature classes: once conditions on constraints of [back] and [round] spreading are removed from the formulation of the spreading constraints themselves, we have no basis on which to distinguish [back] and [round] spreading constraints at all. Independently of this, the cross-linguistic co-patterning of these features we considered at the outset suggests that our goal must be to unite them.. In the context of Optimality Theory, and assuming features are targeted directly and individually, we can do this, formulating a single Color harmony constraint.

This thinking obviously challenges a strategy whereby rules are judged to be different due to different conditions placed on them. According to this strategy, [back] harmony and [round] harmony in Turkish are clearly different rules (and accordingly have been treated as such). Odden (1991) is a particularly relevant work in this regard: in order to establish the unity of [back] and [round] spreading in Eastern Cheremis, as we saw, Odden is careful to argue that the two feature spreadings are subject to identical conditions. Odden's case for the class Color is made rather challenging by this strategy, given that color harmonies very often display more restricted behavior by the feature [round]. To the extent that conditions on rules are truly idiosyncratic (perhaps in the sense either of phonetically unnatural, morphologically restricted, or both), this strategy for equating or distinguishing rules may be quite valuable. However, to the extent that they manifest independently motivated dimensions of wellformedness, as they often do, our real imperative is surely to factor the conditions away from the rules altogether.

4. Partial class spreading: place features. Recall that nasal place assimilation occurs in Kpelle, across all places, giving forms such as those in (29). The data are from Welmers (1962, 1969). The nasal involved is tone bearing and syllabic, and indicates first person possession. Tones are not shown.

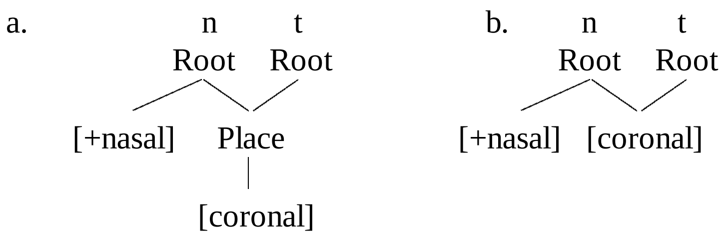
(29) Kpelle nasal place assimilation

[m]bolu	'my back'	[ŋ]gɔɔ	'my foot'
[ɱ]vela	'my wages'	[ŋm]gbiŋ	'myself'
[n]duɛ	'my front'		

Nasal place assimilation such as this is extremely common across languages, and offers perhaps the best known evidence for a feature class. As is typically the case, assimilation here involves at least all of the major articulator features (for Kpelle, [labial], [coronal] and [dorsal]), and so the generalization must be over the class Place. (There is no smaller class including these features, assuming the feature classes presented in section 2.)

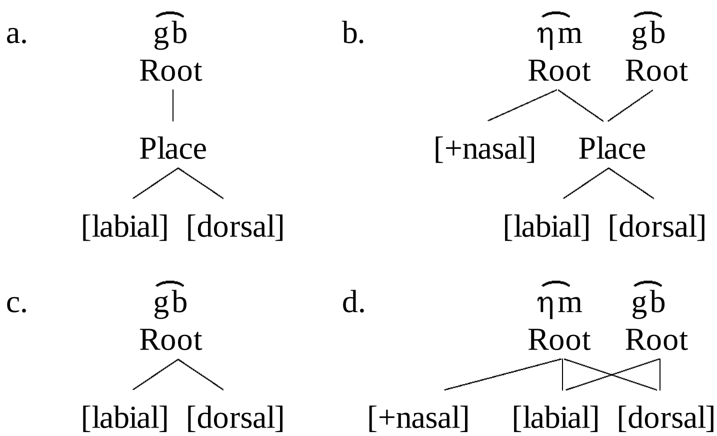
Within Feature Geometry, a rule or constraint of Place assimilation targets a node Place in the representation, thereby indirectly leading to the sharing of place features such as [coronal], as shown in (30)a. In contrast, Feature Class Theory assumes that features such as [coronal] are targeted directly and individually, leading to the direct linking of such features, as in (30)b, even when a constraint mentions Place.

(30)



Consider now nasal place assimilation to the labiovelar [g͡b], also seen in (29). The labiovelars [g͡b] and [k͡p] are a well-known kind of complex segment, that is, single segment having more than one place of articulation. Certain instances of labiovelars have sometimes been treated as simply [labial] or [dorsal] according to a language's phonology, as suggested by gaps in the phoneme inventory (see especially Anderson 1976). However, most are clearly phonologically complex, contrasting with both plain labials and plain velars within a language; this is the case in Kpelle. As is true in many West African languages, assimilation to [g͡b] in Kpelle yields [η͡mg͡b], a homorganic sequence of a doubly articulated nasal followed by a doubly articulated stop. (The labial and velar articulations of a labiovelar largely overlap; [η͡mg͡b] should be understood as a single constriction of overlapping labial and velar gestures, nasalized at first and then released orally.) An important result of Feature Geometry noted by Sagey (1986) involves its straightforward handling of such assimilation: if [g͡b] is a single segment represented as in (31)a, then linking the node Place will entail linking of both [labial] and [dorsal] without further ado, (31)b. Feature Class Theory retains this result: nasal place assimilation requires that all Place features be linked, and the place features found in (31)c are targeted directly, by virtue of their status as place features, without the mediation of a mother node, as in (31)d. (On the crossed lines in (31)d see note 1.)

(31)



However, another outcome of nasal place assimilation before labiovelars is observed in many languages. Assimilation can give simple [η], partial assimilation, i.e., ηg͡b and ηk͡p. This kind of partial assimilation has been reported to occur in Gonja (Painter 1970:74), Nkonya, Efik (Ohala and Lorentz 1977, citing Reinecke 1972 and Cook 1969 respectively), Birifor, Konkomba, Anufo, Kɔnni, and other Gur and Kwa languages cited by Cahill (1995) (see references there), to name just some. Working from his own field notes, Cahill notes the occurrence of both types of assimilation in Kɔnni: nasal place assimilation is partial within words (including compounds) and total across words. Padgett (1995b) reports something close to the reverse in Gã, a Kwa language—assimilation is total within a morpheme but partial across a morpheme boundary, a difference that leads to surface minimal pairs such as η͡mk͡pai 'libation' versus η-k͡pai 'my cheeks'.[8] The facts of Kɔɔni within words are illustrated below. There is a general process of nasal place assimilation in the language, giving results across all places of articulation as in (32)a. Partial assimilation before labiovelars is shown in (32)b.

(32) Kɔɔni nasal place assimilation (Cahill 1995)

a.	démbíŋ	'man'	b.	tɪŋg͡báŋ	'floor'
	dantɪ-má	'greet!'		bɪŋk͡pɪáŋ	'shoulder'
	koŋkógiŋ	'mountain'			

More generally, instances of partial nasal place assimilation can be found involving the full range of different complex segment types across languages. Assimilation to clicks, for example, is typically partial (Maddieson and Ladefoged 1989). The relevant examples involve the prefixation of a nasal-final class marker in Bantu languages having clicks, as in the Zulu examples in (33). The final nasal of the class 10 prefix <u>iziN</u>- assimilates across all places of articulation, as shown in (33)a. Before the clicks in (33)b (respectively dental, palato-alveolar and lateral), assimilation produces simple velar [ŋ], as before [k] . (Doke 1926:78, Doke 1931:14, 52-3, data from the latter.) Since these clicks are corono-velars (Sagey 1986, building on Chomsky and Halle 1968 and Halle 1983), this assimilation is partial. (The clicks are also voiced following the nasal.)

(33) Zulu nasal place assimilation

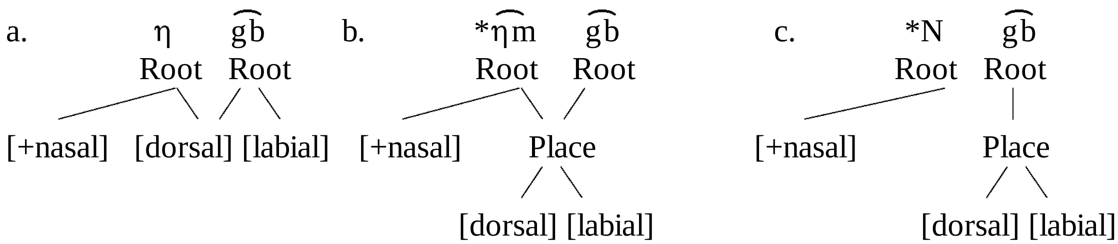
a.	izim-paphɛ	'feathers'	b.	iziŋ-|ezu	'slices'
	izin-ti	'sticks'		iziŋ-ǂuŋǂulu	'species of bird (pl.)'
	iziŋ-kɛzɔ	'spoons'		iziŋ-‖aŋ‖a	'green frogs'

Assimilation to the glide [w] (another kind of labiovelar) is also typically partial and occurs widely, with results including [ŋgw] (e.g. Kihungan, Clements 1987), [ŋw] (e.g. Bakweri, John Kingston, p.c.), and (less commonly) [mw] (e.g. Chukchee, Skorik 1961). Partial assimilation to the glide [j] occurs also, as in Kihungan /N + j/ → [ŋgj] (Clements 1987), where the output nasal is purely [dorsal].[9] In Central Catalan, assimilation of /n/ to palatal [j] and alveopalatal [ʎ] gives not (alveo)palatal [ɲ], though this is a phoneme of the language, but postalveolar [n̠], the same nasal that occurs before postalveolar [ʃ] and [ʒ], e.g. <u>so[n̠ʒ]ermans</u> 'they are brothers', <u>so[n̠ʎ]iures</u> 'they are free', <u>u[n̠j]ogurt</u> 'a yogurt' (Mascaró 1976, Recasens 1991, Recasens 1993). In order to capture this last partial assimilation, Mascaró (1976) omitted [high] from a rule of nasal place assimilation employing greek variable notation, taking the postalveolars to be different from palatals and alveopalatals in lacking the specification [+high]. This is no longer a possible solution, since the variable notation has been eliminated from the theory. The featural make-up of palatal and alveopalatal segments is not a matter of consistent agreement today, but a widely-held view due to Keating (1988) treats them as complex segments, both [coronal] and [dorsal]. We might hypothesize that nasal place assimilation succeeds in causing only [coronal] (and dependent features) to link. In any case, all of the cases just mentioned involve partial assimilation to a complex segment as part of a more general place assimilation process affecting all places of articulation.

There is an alternative to viewing sequences like [ŋg͡b] as instances of partial assimilation: suppose place assimilation to a complex segment is simply blocked altogether, and there are actually two independent velar gestures. The nasal is velar by default. (There are many examples of velar behaving as default place for coda nasals, see Trigo 1988.) The cross-linguistic facts do not support such a picture in general. Default place for nasals is often coronal rather than velar. If assimilation could fail before complex segments, leading to default place, we would equally predict [ng͡b] to occur sometimes in a language otherwise showing nasal place assimilation, and yet apparently no cases are attested.

The issues raised by these cases are just those raised earlier by color harmonies. In Kɔɔni and Zulu, there is a general nasal place assimilation process (targeting all places of articulation), just as in Turkish there is a general color harmony. In spite of this fact, spreading is only partial in certain forms. Once again partial assimilation is likely due to the pressures of an independently motivated wellformedness constraint—in this case a ban on complex segments (see below). Partial assimilation is illustrated in (34)a below: some, but not all, of the Place features of [g͡b] spread from a single segment. Partial class behavior in such cases is unexpected in Feature Geometry, for reasons by now familiar. The targeting of a node Place predicts that either all place features will assimilate, as in (34)b, or none will, (34)c.

(34)



Let us assume that spreading in nasal place assimilation occurs due to a constraint Spread(Place), shown below; this constraint is identical to Spread(Color), except for the feature class affected. (Again I assume that other constraints not considered here limit assimilation to within a consonant cluster as required. This is not crucial to the point.)

(35) SPREAD(PLACE, PWD): For all place features *f* in a prosodic word, if *f* is linked to any segment, it is linked to all segments

The tableau in (36) presents an analysis of Kɔɔni partial nasal place assimilation that is entirely parallel to that of Turkish partial color harmony. Since nasal place assimilation generally occurs, e.g. démbíŋ 'man', we can infer Spread(Place) >> Ident. On the other hand, since it does not succeed completely before labiovelars, Spread(Place) must in turn be dominated by a constraint the disfavors the expected outcome [ŋ͡m]. Complex segments are generally marked in comparison to simple segments, motivating a universal constraint *CompSeg ('no complex segments') that prohibits consonants having more than one place of articulation. This constraint will ensure that assimilation is only partial in the relevant forms. Therefore, even though candidate (36)a satisfies Spread(Place) completely, it loses in comparison to (36)b-c. Violation of Spread(Place) is not total, however, as in candidate (36)f. Rather, the optimal candidates (36)b-c violate Spread(Place) minimally, that is, only to the extent required by*CompSeg. Finally, just as non-high round vowels are allowed in Turkish roots, complex segments do occur in Kɔɔni syllable onsets. The syllable onset is a priveleged phonological position in the sense discussed earlier, allowing more contrasts, and more marked segments, than the syllable coda. The highest ranked constraint posited here, accordingly, is Indent$_{Onset}$, a constraint favoring faithfulness in a perceptually prominent position, analagously to Ident$_{Root}$. (See Beckman 1997, 1998 and references therein on this constraint, and on its connection to earlier 'onset licensing' approaches to the onset-coda asymmetry.) It is this constraint that explains why complex segments occur at all in Kɔɔni words. Indent$_{Onset}$ also explains why it is the nasal that assimilates to the place of the onset consonant, rather than the reverse; the latter spreading would violate this constraint. (The assumption that the underlying nasal is unspecified for place features is therefore not needed to explain this last fact, but is simply a matter of convenience; see note ?.) Indent$_{Onset}$ also explains why candidates (36)d-e are not optimal, in spite of satisfying both Spread(Place) and *CompSeg perfectly.

(36)

UR: /N + g͡b/	Ident$_{Onset}$	*CompSeg	Spread(Place)	Ident
a. ŋ͡m g͡b		**!		**
b. ☞ ŋ g͡b		*	*	*
c. ☞ m g͡b		*	*	*
d. ŋ g	*!			**
e. m b	*!			**
f. N g͡b		*	**!	

So far the analysis chooses two candidates as optimal, [ŋg͡b] and [mg͡b]. In fact, it is the former that occurs in all known cases of partial assimilation to a labiovelar stop. Why is this the case? Labiovelar stops are produced with labial and velar constrictions that are nearly, but not completely, overlapped in time. Generally the velar closure is achieved slightly earlier than the labial closure, and likewise the velar release precedes the labial release. The perceptual cues to the presence of the velar constriction are therefore more robust at the onset of the consonant closure, and those to the labial constriction at the release of the closure (Ladefoged and Maddieson 1996). The cluster [ŋg͡b] is very much like simple [g͡b] in this respect: the velar gesture is cued first, this time by the velar nasal that is homorganic to the velar portion of [g͡b], and the labial gesture last. The cluster [mg͡b] is different: it begins with a labial nasal, and ends with the homorganic labial portion of [g͡b]. In other words, this transcription implies that the labial gesture begins before the velar gesture, and ends after it, while being held continuously. In this scenario there would be no significant acoustic cue to the presence of the velar gesture at all, and the cluster would be virtually indistinguishable from [mb]. These considerations suggest one possible reason for the favoring of [ŋg͡b] over [mg͡b] when assimilation is partial, namely a requirement that a feature or gesture be perceptually recoverable. It would be possible to posit a new constraint or family of constraints with this intent, but Steriade (1995b) suggests an alternative that is preferable: since faithfulness constraints in Optimality Theory already exist in order to preserve features in the input to output mapping, perhaps they should be minimally reinterpreted to require specifically that input features be perceptually realized in the output. A statement of Ident$_{Onset}$ along these lines is offered below, with the novel portion emphasized. The same amendment is assumed to hold for all Ident constraints.

(37) IDENT$_{ONSET}$(F): Let S_I and S_O be corresponding segments of the input and output, where S_O is in the Onset. Then if S_I is specified [αF], S_O is *recoverably* specified [αF], where α is +,-, or 0.

With this new understanding of faithfulness, the needed distinction is made between our candidates, as shown in (38). Candidates (38)b-c correspond to (36)b-c above. The reformulation of Ident has no consequences for the violations incurred by the other candidates of (36).[10]

(38)

UR: /N + g͡b/	Ident$_{Onset}$	*CompSeg	Spread(Place)	Ident
b. ☞ ŋ g͡b		*	*	*
c. m g͡b	*!	*	*	**

This explanation for the favoring of velar assimilation over labial assimilation cannot obviously extend to the case of clicks, where the same fact holds, or to labiovelar glides (but note that [mw] is attested, as mentioned above); for an alternative explanation for this asymmetry in general, see Ohala and Lorentz (1977) and Ohala and Ohala (1993). In any case, the means by which [ŋg͡b] is favored over [mg͡b] is not crucial to our concerns. The important point here is that assimilation before complex segments is partial, and some constraint that outranks Spread(Place) must account for that.

It is easy to see that with the ranking of Spread(Place) and *CompSeg reversed, the fully assimilated candidate (36)a would emerge as optimal. This is the ranking that derives Kpelle and other cases in which [ŋ͡mg͡b] occurs. This ranking may also be responsible for total assimilation across words in Kɔɔni, as opposed to within words, assuming that a distinction between lexical and postlexical phonology is maintained in Optimality Theory (see McCarthy and Prince 1993b, Kiparsky 1999). (Alternatively, total assimilation between words could be due to gestural overlap, a properly phonetic cause; phonetic investigation is necessary in order to distinguish these possibilities.) The general account therefore correctly predicts the existence of both full and partial assimilation, and does so with a minimum of assumptions: a constraint against complex segments and a basic account of place assimilation. Partial assimilation falls out immediately, given the existence of constraint (re)ranking, the quite general means of language variation in Optimality Theory.

5. Formalizing Feature Class Theory. The central proposal of Feature Class Theory is that rules or constraints mentioning feature classes thereby target (or hold true of) the relevant features directly and individually. In particular, no class nodes mediate as in Feature Geometry. How precisely does this work? Here I suggest a very direct approach, in which classes such as Place and Laryngeal are understood as sets of features. (For a precedent involving syntactic feature classes, see Gazdar et al. 1985.) These sets are not arbitrary lists, if we carry over from Feature Geometry the assumption that the classes have a phonetic basis; this means they can be defined equivalently by the respective properties that their members share: placehood, laryngeality, and so on. The simplest sets we might consider to start with are those of (39), sets that consist of a single feature. It is useful to define these sets, since we can then define larger feature classes as unions over them, as I do momentarily; the subset relations that obtain are made clearer this way.[11]

(39)	Nasal	$=_{def}$ {[nasal]}		Coronal	$=_{def}$	{[coronal]}
	Voiced	$=_{def}$ {[voiced]}		Dorsal	$=_{def}$	{[dorsal]}
	Aspirated	$=_{def}$ {[aspirated]}		Low	$=_{def}$	{[low]}
	Glottal	$=_{def}$ {[glottalized]}		High	$=_{def}$	{[high]}
	Pharyng	$=_{def}$ {[pharyngeal]}		Back	$=_{def}$	{[back]}
	Labial	$=_{def}$ {[labial]}		Round	$=_{def}$	{[round]}

Assuming the validity of the classes of the Feature Geometry representation in section 2 (see (7)), I carry them over in the definitions of (40). Let us assume the classes to be fixed and universal as well, in the usual way. (On the status of the segment, see below.)

(40)	Laryngeal	$=_{def}$	Aspirated ∪ Voiced ∪ Glottalized
	Place	$=_{def}$	Pharyngeal ∪ Oral
	Oral	$=_{def}$	Labial ∪ Coronal ∪ Dorsal ∪ VPlace
	VPlace	$=_{def}$	Height ∪ Color
	Height	$=_{def}$	Low ∪ High
	Color	$=_{def}$	Back ∪ Round

Terms like Place, and Color in rule or constraint formulations make reference to these postulated sets, rather than to the nodes of Feature Geometry. Consider once again the constraint Spread(Place), for example. The formulation below is precisely analogous to that of section 4 (see (35)), except for making explicit that 'place' features are features of the set Place. Since Place is defined as Pharyngeal ∪ Oral, the latter as Labial ∪ Coronal ∪ Dorsal ∪ VPlace, and so on, this constraint will hold over all instances of [pharyngeal], [labial], [coronal], [back], etc., in the relevant representation.

(41) SPREAD(PLACE, PWD): For all features $f \in$ *Place* in a prosodic word, if *f* is linked to any segment, it is linked to all segments

This constraint quantifies universally over the set Place; that is, it refers to all features of that set. So long as this constraint is interpreted as gradiently violable in the appropriate way, it will have just the effect required in section 4. Other constraints will naturally quantify differently, depending on their intent. Consider for example the constraint HavePlace (Itô and Mester 1993, among others), given below. This constraint is intended to rule out segments such as [h], [ʔ], and other placeless segments such as the 'mora nasal' of Japanese. This constraint quantifies existentially; that is, any segment having at least one of the features [labial], [coronal], [back], etc., will satisfy this constraint, as intended.

(42) HAVEPLACE: Every segment must have some feature $f \in$ *Place*

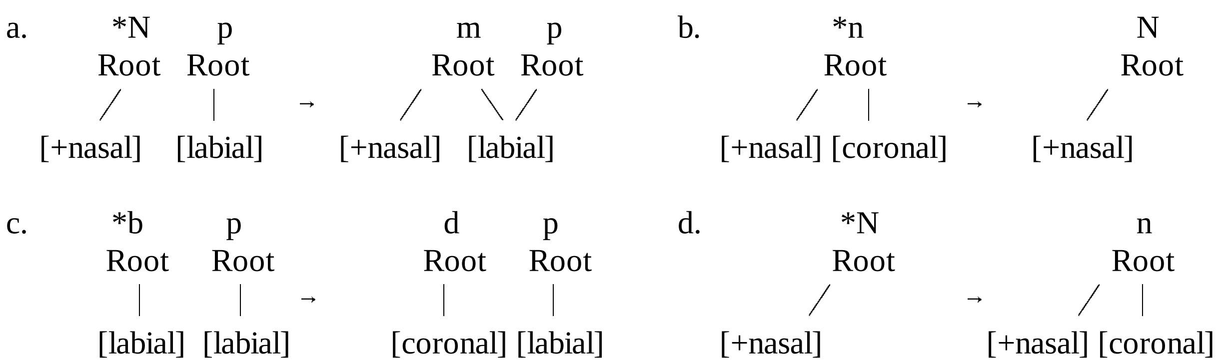
Two more constraints are worth considering for a comparison to Feature Geometry next section. The Obligatory Contour Principle relativized to Place will be discussed then. *Place requires the opposite of HavePlace, and can force the absence of place features, as with the mora nasal of Japanese. (Other constraints such as Ident$_{Onset}$ will prevent the loss of place in all enviornments.)

(43) OCP(PLACE): Adjacent identical features *f*, *g* ∈ *Place* are prohibited

(44) *PLACE: No segment may have any feature *f* ∈ *Place*

The figures in (45) illustrate the effect of all of these constraints on representations. (45)a shows the effect of Spread(Place) on a consonant cluster as seen in section 4. On the left of the arrow is an ill formed representation according to this constraint; a well formed cluster appears on the right. (Alternatively, (45)a can be understood as a rule that inserts an association line.) The individual feature [labial] is targeted directly, by virtue of being a member of the set Place. (45)b similarly shows ill- and well-formed representations according to *Place. Finally, (45)c-d do the same for OCP(Place) and HavePlace. The straightforward extension of the theory to (45)c-d will be contrasted in the next section with the difficulties encountered by Feature Geometry.

(45)



Feature Class Theory maintains the assumptions central to any account of feature class generalizations, the assumption of privilege, and the assumption of simplicity and naturalness (section 2). The assumption of privilege is handled by means of the feature sets seen above, rather than by targeting a class node in a representation as in Feature Geometry. In order for this strategy to make sense, we must also interpret the assumption of simplicity and naturalness somewhat differently. Recall autosegmental phonology's measure of simplicity for rule statements (section 2): rules may insert or remove at most one association line (or more weakly, the fewer the lines, the better). This assumption is carried over to work in Feature Geometry. It cannot be our assumption, though, precisely because of the claim that features are targeted directly and individually. This claim entails that several association lines might be inserted or removed due to the requirements of one constraint such as Spread(Color), and indeed this was a desired feature of the analyses of sections 3 and 4. If a process spreading both [back] and [round] is to be as 'low-cost' as one spreading only [back], then it cannot be the number of association lines we are counting. The obvious alternative is to concentrate on the constraint formulations themselves. The idea behind Spread(Color), after all, is that it would not be desirable to posit a constraint such as Spread(back) and Spread(round). The latter formulation is more complex, but even more important, it fails to account for the privileged recurrence of just the features [back] and [round] in assimilation, since it is no more difficult to formulate a constraint Spread(back) and Spread(high), for example. Seen in this way, the assumption of simplicity and naturalness is not something special to autosegmental theory or phonology, but reduces to the requirement of any theory to capture generalizations in the most explanatory manner possible.

What is a segmental representation in light of Feature Class Theory? Since class nodes no longer mediate feature class behavior by functioning as targets of rules or constraints, and since this was their original purpose, they have at best a questionable status in Feature Class Theory. Class nodes are not inconsistent with Feature Class Theory, however. It is entirely possible to hold that features are affected directly and individually while assuming a representation with class nodes. Indeed, some have advocated individual feature behavior while retaining class nodes (see section 7.1). It makes sense to retain class nodes if they serve a purpose even without mediating feature class behavior. Minimally, they might serve as labels through which the features of a class—those dominated by the relevant node—are identified for targeting. This is hardly necessary to the theory, however convenient it is for those viewing representations; the sets posited here do this work in a way that is formally well understood. Class nodes have a more serious potential role in explaining facts of locality (transparency versus blocking) in phonology. This has in fact been an important secondary job for class nodes in Feature Geometry. (See Clements 1985, Steriade 1987, and Clements and Hume 1995, for example.) However, other recent work advocates a different approach to locality problems, rooting such effects instead in matters of segment wellformedness, while assuming that segments are not skipped. (See Padgett 1995c, Gafos 1996, Walker 1998, Ní Chiosáin and Padgett to appear, and references therein.)

The assumption here is that there are no class nodes. This means that a segmental representation is as in (46) (with only some features shown), much like the 'bottlebrush' representation of Hayes (1990). Feature order in the diagram is randomized in order to emphasize the point that feature classes are wholly understood as sets of features that share the relevant property. The features are assumed to be associated to a Root node that serves as the representational encoding of the segment, as in Feature Geometry. Perhaps equivalently, this node might be an 'X' or 'C/V' skeletal unit as in Steriade (1982), Clements and Keyser (1983), and Levin (1985). As Broselow (1995) notes, the Root node and skeleton overlap in function, and the theory might not require both; this is all the more true if the mora is included in the theory.

(46)

Root

[coronal] [voice] [labial] [high] [nasal] [back]

The assumption of a node corresponding to the segment requires comment, since it implies that the segment is different from a feature class. In principle we could define a class Laryngeal ∪ Place ∪ Nasal corresponding to the Root node of our Feature Geometry representation (see (7)). This would correspond to a segment in the sense of including all feature content; however, the segment is not merely a collection of feature content, but has the well known additional status in phonology of a timing unit. (See the references above on the segmental skeleton; also Hayes 1990 and Scobbie 1991 on this role for the Root node of Feature Geometry.) When a segmental metathesis $S_1S_2 \rightarrow S_2S_1$ occurs, for example, we require some means of ensuring that all the features of S_1 now follow all those of S_2 in time. In Feature Geometry the Root node has the conflated roles of capturing a feature class on the one hand (feature content), and timing relations on the other (feature coordination). As Hayes (1990) demonstrates, these two roles need not be conflated. In addition, evidence suggests that only the Root node is necessary for the purposes of coordinating features in time; that is, there is no reason to assume that nodes such as Place and Laryngeal perform any such role (Hayes 1990, Scobbie 1991, Padgett 1991/95a). This fact lends further support to the formal separation of these functions. In Feature Class Theory, feature classes are understood as feature sets; this leaves only the role of timing or feature coordination for the Root node or skeletal unit. Of course, it is possible nevertheless to define a feature class Segment. Support for such a feature class would come from evidence that processes demonstrably affecting entire segments nevertheless target features directly and individually, along the lines seen in this paper. This is a matter worth further investigation, and the question is left open here.

On a separate note, the question might arise why in Feature Class Theory feature classes do not overlap (ignoring subset relations), since we might easily posit that [aspirated], for example, is a member of both Laryngeal and Place. Actually, this question arises equally in Feature Geometry, though here the analogous assumption would be that [aspirated] is linked to both the Laryngeal and Place class nodes. Both theories must either stipulate or derive a prohibition on feature class overlap, assuming it is empirically required. The appropriate constraint for Feature Geometry, sometimes known as the 'single mother node condition' in work on syntax, is not usually made explicit. Yet whether trees obey such a condition or not is something that must be formally declared. The relevant question in Feature Class Theory is, can subsets of a set ever intersect? Here I carry over Feature Geometry's implicit assumption that such overlap cannot occur. A more interesting question is why this should be the case; most likely the answer is to be found in the phonetic underpinnings of feature classes: the larynx, oral gestures, the velum, and so on, are phonetically relatively independent.

6. Dissimilation and other feature class generalizations. Sections 3 and 4 argued for Feature Class Theory based on the existence of partial class behavior in assimilations. This section makes an independent argument: there are entire classes of generalization apart from assimilation, which clearly involve feature classes, but which Feature Geometry does not account for at all. The main example considered here involves dissimilation across place of articulation.

A fair sample of languages have been found to exhibit cooccurrence restrictions over all places of articulation within roots. These include Classical Arabic (Greenberg 1950, McCarthy 1988), Russian (Padgett 1991/95a), Yucatec Maya (Straight 1976, Yip 1989, Lombardi 1990), Javanese (Mester 1986, Yip 1989), and Pomo (Yip 1989). In much work since McCarthy (1985) and McCarthy (1988), such cooccurrence restrictions have been seen as reflexes of the Obligatory Contour Principle (henceforth OCP, Leben 1973, McCarthy 1981, McCarthy 1986) applying over place features. The OCP is usually stated as in (47)a. In order to capture the recurrent privileged status of the place of articulation features under dissimilation, however, a particular instantiation of this general formulation is required, as in (47)b. As the simpler formulation does, this one abstracts away from questions such as what counts as adjacent, why the cooccurrence restrictions apply only within roots, and other matters; these issues are independent of our main concern.

(47) a. OCP: Adjacent identical elements are prohibited

b. OCP(PLACE): Adjacent identical features *f*, *g* ∈ *Place*, are prohibited

McCarthy (1988) notes that dissimilatory effects are a major diagnostic of feature classes, along with assimilations and neutralizations. An OCP relativized to place features is implicit in his argument for a feature geometric class node Place, based on root cooccurrence effects over just these features in Classical Arabic. Here I take up this idea, while arguing that it cannot in fact be pursued within Feature Geometry, but instead requires Feature Class Theory.

The facts of cooccurrence in Russian roots are considered here. They are remarkably similar to the better-known Arabic facts. Russian consonants can be grouped into the following 'identity classes' (to use a term of Yip 1989). Consonants within any group fail to cooccur within Russian roots, to a statistically significant degree. The results are based on a chi-square analysis performed by Padgett (1991/95a) on the distribution of non-adjacent consonants in a list of about 500 roots. C' is palatalized.[12]

(48) Russian identity classes

Labials	p, p', b, b', m, m'
Coronal sonorants	l, l', r, r', n, n'
Coronal stops	t, t', d, d'
Coronal fricatives	s, s', z, z'
Dorsals	k, g, x, č, š, ž [13]

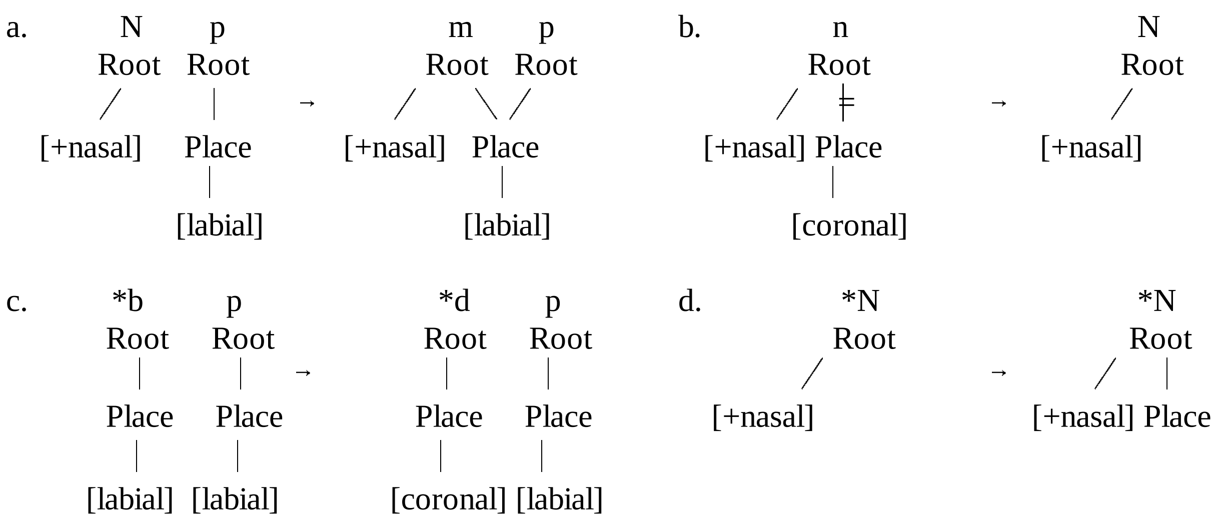
Roots can consist of consonants drawn from separate identity classes, for example brat- 'brother', bod- 'awake', sad- 'sit', greb- 'dig', koz- 'goat', tolk- 'explain'; roots with two or more consonants from the same identity class are significantly underattested: *soz-, *dat-, *pib-, *kag-, etc.

What is interesting for our purposes is that the failure to cooccur is contingent on identity of place. Russian and the other languages cited above exhibit the generalization that restrictions are imposed only on consonants of the same place of articulation—[labial], [coronal], and [dorsal]. (On the further manner requirements evident for the coronals see below.) Here we have essentially a dissimilatory phenomenon, rather than an assimilation as seen in earlier sections; yet the class recurrently targeted is a familiar one, place of articulation, and the logic that motivates feature classes holds here no less than before. Yet the formalism of Feature Geometry does not extend directly to such facts. In order to see this clearly, we must consider the precise assumptions behind that theory. The following discussion presupposes the formal interpretation of Feature Geometry provided by Bird (1995).

First, consider assimilation. Insertion of an association line as in (49)a establishes an immediate dominance relation between the first Root node and the Place node. (The reader should compare (49) to (45).) The result, shown on the right of the arrow, is that by virtue of this association, the first Root node dominates whatever Place dominates. This is so because 'dominance' is understood as the transitive closure of the immediate dominance relation (Bird 1995). That is, suppose node a_1 immediately dominates a_2, a_2 immediately dominates a_3., ... , a_{n-1} immediately dominates a_n. Then a_1 'dominates' a_n, for any n. So 'dominance' follows from a chain of immediate dominance relations, rather than being another relation separately stated between all relevant node pairs such as a_1 and a_n. If dominance implies temporal overlap, it follows that both Root nodes overlap the [labial] feature in (49)a, giving [mp].

Consider now cases of neutralization, characterized in autosegmental terms by the delinking, or removal, of an association line. For example, Spanish nasals do not contrast in place when they occupy a syllable rhyme. In most dialects, when they cannot assimilate to a following consonant, they must be [coronal] (in other dialects they must be velar), e.g. corazo[n] 'heart', pa[n] 'bread', canta[n] 'they sing', vs. *corazo[m], *pa[ɲ], *canta[ŋ], etc., (Harris 1983, Harris 1984). Neutralizations to a uniform place have been understood as a kind of reduction in autosegmental phonology, formulated as the delinking of the relevant material. Harris (1984) formulates the rule shown in (49)b (slightly recast into Feature Geometry notation); it is restricted to apply in the syllable rhyme. The desired output, lacking a Place node, is shown. Understood in the formal terms outlined above, this rule severs the immediate dominance relation between Place and the Root node, and thereby all dominance relations between Place and higher prosodic structure. Place will therefore undergo stray erasure (or fail to be phonetically realized), if this is the fate of 'stray' material (McCarthy 1979, Steriade 1982, Itô 1986, Itô 1989, Goldsmith 1990, Itô and Mester 1993). More to the point, its dependents, [labial], [dorsal], or (in this case) [coronal], and so on, are now also not dominated by any prosodic structure, for analogous reasons, and so must also stray erase. This is the desired result. A default rule filling in [coronal] is assumed to apply next, accounting for the [n] of corazo[n] etc.[14]

(49)



Feature Geometry generally works well for assimilations and reductions. In particular, the unity of the accounts has much to do with the central role of dominance, and the status of dominance as the transitive closure of immediate dominance. As we have seen, however, this approach fails for cases of partial class behavior under assimilation. The problem is equally apparent in the case of dissimilations such as that of Russian: if constraints mentioning feature classes held of the relevant class nodes themselves, then OCP(Place) would mean 'Adjacent identical Place nodes are prohibited'. This would have the undesirable effect of ruling out both the input and output representations in (49)c, and similarly nearly all sequences of segments—those segments having Place nodes. The understanding of dominance in itself contributes nothing to the desired outcome. As Yip (1989:363) observes, it is not the Place nodes but the place features themselves that matter. Yip therefore assumes that the OCP, though stated over Place, involves 'checking of identity' at the featural level. That is, the OCP mentions the class Place, but thereby holds of the relevant place features directly and individually. This is the main claim of Feature Class Theory, and within that theory the feature class formalism extends straightforwardly to OCP effects, as we have seen, involving nothing new beyond the simple assumptions needed to accommodate assimilations and reductions.

It is not hard to find other categories of constraint, outside the realm of assimilations and reductions, that raise similar issues for Feature Geometry. The constraint HavePlace seen in the last section is an example. This constraint cannot serve its purpose if understood to require merely that the class node Place appear, as seen in (49)d; what matters, again, is that an actual place feature be present. Of course, the right result can be made to follow in Feature Geometry if we make the additional assumption that every class node must dominate some terminal feature. In itself this presents no difficulty. The problem is that some different extra assumption is required to make the OCP effects work, and something differerent yet again to make partial assimilations work (see section 7.2). Feature Geometry's premise that constraints hold of class nodes is actually inadequate in diverse ways, depending on the kind of process or constraint we have in mind, and for each way a new implementational assumption is required to ensure that things work. The upshot is an account for feature class generalizations that is complicated and lacking in unity.

Recent work by Pierrehumbert (1993), Frisch, Broe, and Pierrehumbert (1995), and Frisch (1996) argues against the familiar OCP account of place cooccurrence restrictions, and in favor of an account using a gradient similarity constraint. In this view, the constraint is gradient in two ways: it becomes weaker as consonants are separated by more segments, and also to the extent that consonants are different in properties other than place of articulation, e.g. sonorancy, stricture, voicing, etc. (See also Berkley 1994 on similar but weaker effects in English.) It is the weakening of the constraint due to differences in features other than place that causes coronals to divide into separate identity classes in Russian, Arabic, and other languages. It remains true in this conception of cooccurrence restrictions, however, that the restrictions hold of place of articulation: similarity in other ways, and distance effects, matter only if place features are identical to begin with; hence Frisch (1996) continues to refer to 'OCP-Place'. It also remains true that the restrictions affect all places of articulation rather than one, across various languages. The conclusions above therefore extend to this conception of cooccurrence restrictions: a general constraint holds over place of articulation, and it necessarily targets directly the place features themselves. The points hold as well under a conception of the OCP as the local conjunction of markedness constraints, as in Itô and Mester (to appear), Alderete (1997).

7. Discussion. The goal of this section is to place Feature Class Theory within a larger context, including other work on feature classes, work in phonology, and linguistic theory more generally. Section 7.1 briefly notes precedents to Feature Class Theory. Section 7.2 considers alternatives and possible objections, and further compares the approach to Feature Geometry. Finally, 7.3 considers the status within linguistics of constraints versus representations.

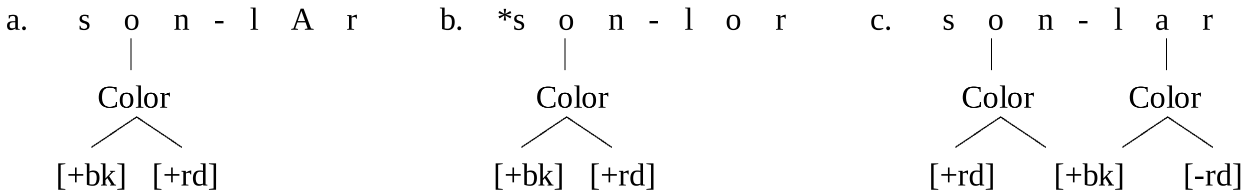
7.1 Precedents. The notion that features are affected directly and individually, even when a rule or constraint holds of a class of features, is a possibility raised by Sagey (1987) for cases of spreading she dubs 'non-constituent spreading'. She focuses most notably on a case involving

Barra Gaelic. It is unclear whether the analysis of Barra is correct (see Ní Chiosáin 1994b), but the argument is very much like that presented here based on Turkish forms ending in a palatalized lateral (section 4). However, Sagey regards 'non-constituent' spreading as a marked option, a deviation from the usual Feature Geometry scenario of class node spreading. Halle (1995) revives Sagey's argument based on Barra Gaelic. (See also Halle, Vaux, and Wolfe 2000.) Though he assumes that individual feature targeting is the rule for cases of assimilation, the idea is not meant to extend to other processes. Selkirk (1991a,b) and Padgett (1991) assume individual feature targeting as the only means of class behavior, abandoning the class node. However, there is relatively little attempt in these works to motivate this idea empirically; Padgett's (1991) main argument (elaborated above) rests on the OCP. Though superficially distanced by a difference of implementation, Hayes' (1990) understanding of 'diphthongization paradoxes' is essentially individual feature behavior as well. (See also the discussion of this issue in Goldsmith 1990:295-8, where the alternative view of class membership as properties of features is considered.) All of the works just mentioned are cast within a framework involving serially ordered rules.

Here it has been argued that partial class behavior in spreading is much more pervasive than has been assumed (sections 3 and 4); the explanatory domain of the feature class idea has in this way been extended. Further, other compelling arguments for the direct targeting of individual features have been raised (section 6). Individual feature targeting is not viewed as an option, or 'marked' in any sense, nor is it restricted to occur only with certain processes. The formalization of this idea is made explicit (section 5). Finally, the theory is couched within Optimality Theory, with its central notion of (gradient) constraint violability, a notion from which partial class behavior follows very naturally.

7.2 Further discussion and comparison. Consider again the problem for Feature Geometry of partial class behavior, for example as in Turkish forms such as son-lar 'ends' (section 3), illustrated in (50)a. If Color harmony involves spreading of a class node Color, then how are we to avoid deriving *son-lor incorrectly? There are possible responses to this problem that I did not consider earlier. For example, suppose we were to retain the assumption that class nodes mediate feature class behavior, but view individual feature targeting as some kind of available 'resort' or 'repair' mechanism? In particular: in order to target a feature class such as Color, target the appropriate class node (here, Color), unless this move is blocked for some reason; otherwise, target individually as many features dominated by that node as possible. This would have the desired effect: spreading of the node Color would fail, due to the prohibition on non-high rounded vowels. Yet any resort strategy of this kind has the 'do this unless that in which case...' quality that Optimality Theory has sought to redress, and seems unappealing on the face of it. In fact, once we allow for targeting as many features as possible directly and individually, we no longer require a separate mechanism of targeting class nodes, and the main proposal of Feature Class Theory has been adopted.

(50)



A similar objection can be raised against approaches that avoid individual feature targeting by appealing to repairs. Suppose that we allow our theory to create the incorrect intermediate representation *son-lor, as in (50)b. This form is then repaired to eliminate the link between [round] and the target vowel, giving (50)c. This second step requires the delinking of [round] only, and it is not immediately obvious how this should be accomplished, given the representation. (This constitutes a 'diphthongization paradox' as in Hayes 1990.) We might nevertheless pursue this tack and carry through the more drastic re-linkings, and Color node mitosis, required to get the result (50)c. Again, however, this approach by its very nature introduces unnecessary complexity into the analysis: it is the incorrect assumption of uniform class behavior in (50)b itself that leads to the need for a repair.

The problem that partial class spreading presents arises only assuming that spreading as in Turkish actually involves a unitary process of color harmony, as I have argued. Therefore, another means of avoiding the problem involves denying this premise. Discussing Feature Class Theory, Cahill and Parkinson (1997a,b) in fact reject a unifying approach to the facts of both Turkish and Kɔɔni. For Turkish two rules are suggested, one spreading a node Color to high vowels, the other spreading only [back] to non-high vowels. A similar approach is taken to Kɔɔni, with separate rules spreading Place and [dorsal]. Two reasons to reject a unification of the Turkish [back] and [round] harmonies are offered. First is the fact that [round] harmony applies under separate conditions from [back] harmony, affecting only non-high vowels. As I have argued, however, this reasoning runs counter to the goal of explanation in phonology; when stipulations on constraint or rule formulations can be factored out by means of independently motivated constraints, it is hard to justify not doing so. (The same reasoning applies for Kɔɔni.) The second reason involves psycholinguistic evidence due to Zimmer (1969) suggesting that Turkish speakers tolerate some violations of [back] harmony, but none of [round] harmony, in nonsense words. There are two major problems with this argument. First, Zimmer's data involve only roots. The generalizations discussed in section 3 involve suffixes, in contrast, and the suffix generalizations are uncontroversially robust and productive. Second, even assuming that Zimmer's results did diagnose a grammatical distinction between [back] and [round] harmony, they do not tell us where in the grammar this distinction is made. In any account of the Turkish facts, including the one here, the output facts of [back] and [round] harmony differ, and this distinction can presumably become psycholinguistically relevant. We have no reason to conclude that Zimmer's results reflect the grammar's formulations of spreading rules or constraints per se. Apart from these points, it is not clear how Cahill and Parkinson's arguments support their own claims, since they posit not separate [back] and [round] rules, but separate Color and [back] rules.[15]

The inference of a unitary color harmony generalization for Turkish is no trivial move, however. Here a comparison might be useful. In his work widely introducing Feature Geometry, Clements (1985) considered a similar question in relation to some facts of Kikuyu. In this language, when a nasal consonant precedes a fricative, the fricative is hardened to a stop, as in (51)a. When a nasal precedes a voiceless obstruent, the obstruent is voiced, (51)b.

(51) Fricative hardening and obstruent voicing in Kikuyu

a.	mbureetɛ	'lop off'	cf. βura	b.	ndomeetɛ	'send'	cf. toma
	ndeheetɛ	'pay'	cf. reha		nd͡ʒineetɛ	'burn'	cf. t͡ʃina
	ŋgoreetɛ	'buy'	cf. ɣora		ŋgereetɛ	'cross'	cf. kera

These facts on the surface resemble evidence that [voice] and [continuant] together form a class: the appearance is of a uniform process of obstruent hardening/voicing. As Clements notes, however, we should not rush to posit a grouping of [voice] and [continuant]. First, these features represent independent phonetic dimensions. Second, post-nasal hardening and post-nasal voicing occur often and quite independently across languages. Their chance cooccurrence in some languages is fully expected; indeed, it would be odd if no language had both. More generally, hypotheses about feature classes cannot be based on observations within a single language. Rather, they must be informed by the larger context both of the theory, and of cross-linguistic patterning. It is just such considerations that motivate color harmony in cases like Turkish. The class Color is independently well motivated, deriving support from phonetic considerations, patterns of vowel inventories, and the privileged cross-linguistic co-patterning of [back] and [round] in many vowel harmony systems. Even when we restrict our attention to languages such as Turkish where two (or more) features spread under (partially) different conditions, the privileged co-patterning of [back] and [round] is clear, and in need of explanation.

7.3 Constraints and representations. The problem partial class assimilation raises for Feature Geometry derives from the manner in which class nodes have been understood to function. When a class node is targeted for spreading, and there are two or more features under it, then spreading is essentially a hard requirement, rather than a gradient one: in effect, either assimilate all features (of that category), or assimilate none. It is worth emphasizing that it is this functioning of class nodes that is at the heart of the matter, rather than the simple presence of the class nodes in the representation. The argument from partial class behavior is against hard feature class generalizations, in whatever guise, in favor of their soft (gradiently violable) understanding.

Feature Geometry indeed has been such a 'hard' theory. Presenting the facts of Kpelle, in which assimilation to labiovelars is total (see section 4), Sagey (1986:101) provides an illustrative quote: 'Given the representation of /kp/...there would be no way for place assimilation spreading the place node to spread just the dorsal, or just the labial, articulation.' Clearly Sagey was not alone in this understanding of Feature Geometry. In the wake of her important work, the relatively widespread phenomenon of just such partial nasal place assimilations went largely unnoticed. A similar point can be made with respect to color harmonies. In spite of a significant precedent within Feature Geometry for a feature class Color, well known color harmonies such as that of Turkish, where [back] and [round] do not spread identically, were never cited as potential evidence for that class. There may be an issue here concerning the role of theories in helping to uncover the facts they are meant to explain. It is not simply that Feature Geometry did not capture these feature class generalizations. Rather, the generalizations were largely not noticed, and it is reasonable to wonder whether the theory itself contributed to this problem. The success of

Feature Class Theory in bringing partial class behavior to light, and therefore extending the explanatory range of the feature class idea, follows directly from its assumptions of gradient violability and individual feature behavior.

The results of this paper in addition have implications for our understanding of 'constraints' and 'representations' in linguistic theory. In practice, most theories distinguish between the two. For example, within Optimality Theory, constraints play a role in the evaluation of candidates, while candidates themselves are representations. However we view this distinction between constraint and representation in practice, it is not at all clear that it is more than a matter of convenience or tradition. Some linguistic frameworks in fact explicitly identify the two. In phonology, for example, Scobbie (1991) assumes that 'representations' such as /dɔg/ 'dog' are simply very specific constraints. (Compare the lexical 'identity rules' of Kiparsky 1982, e.g., /dɔg/ → /dɔg/, which were assumed to prevail over more general rules by the Elsewhere Condition.) In any case, in practice again, representations often function as 'hard' or inviolable constraints in analyses, as opposed to 'soft' or violable constraints (for theories that allow the latter). This has been the effect of class nodes in Feature Geometry (as they were always interpreted), as noted above.

This is not to say that representations cannot be constructed so as to have gradient properties; they can be. The metrical grid of Prince (1983) is one clear example. Yet representations are obviously a convenient means of conceptualizing invariant properties of language, and this is what they frequently do. To the extent that the facts to be explained do not reflect these theoretical invariants, however, such representations will only hinder our understanding of them. This point is obviously not particular to phonology.

8. Conclusion. Feature Class Theory predicts that there should be other kinds of partial class behavior—involving laryngeal features, vowel height features, and so on. Only two kinds have been offered in this paper, involving vowel color features and consonant place of articulation. With the theoretical context in place to analyze partial class behavior (gradient violability and individual feature targeting), it is to be expected that future analyses will reveal other kinds.

By way of conclusion, let us consider once again the Assumption of Privilege, an assumption common to all work on feature classes. Why should it hold? Why do features pattern into classes at all, and why the particular classes found? Though Feature Geometry and Feature Class Theory are noteworthy in capturing feature class patterning, their formal mechanisms do not provide any answer to these deeper questions. Instead, the answers have been attributed to phonetic underpinnings: feature classes have a basis in the phonetic parameters of place of articulation, laryngeal state, and so on. Yet this assumption deserves further scrutiny, not because it is likely to be wrong, but rather because more attention to the phonetic bases would probably bring a new depth of explanation to the research program. For example, Ohala (1990) argues that nasals are prone to assimilation of place (in part) because the acoustic cues to place are not robust in nasals in comparison to other consonants. On the other hand, the nasalization and voicing of nasals are perceptually robust, explaining why it is typically only place that assimilates. Because of its promise of illuminating such questions, the interplay of phonetics and phonology is currently a burgeoning area of research.

References

Akinlabi, Akinbiyi. 1997. Kalabari vowel harmony. Linguistic Review 14.2:97-138.

Alderete, John. 1997. Dissimilation as local conjunction. In Proceedings of the North East Linguistics Society 27, edited by K. Kusumoto. Amherst, MA: GLSA, Dept. of Linguistics, University of Massachusetts.

Anderson, John; Colin Ewen. 1987. Principles of Dependency Phonology. Cambridge: Cambridge University Press.

Anderson, Stephen R. 1976. On the description of multiply-articulated consonants. Journal of Phonetics 4.17-27.

Archangeli, Diana. 1985. Yokuts harmony: Evidence for coplanar representation in nonlinear phonology. Linguistic Inquiry 16:335-372.

Beckman, Jill. 1997. Positional faithfulness, positional neutralization, and Shona vowel harmony. Phonology 14.1:1-46.

Beckman, Jill. 1998. Positional faithfulness. Amherst, MA: University of Massachusetts dissertation.

Berkley, Deborah. 1994. The OCP and gradient data. Studies in the linguistic sciences 24.2.

Bird, Steven. 1995. Computational phonology. Cambridge: Cambridge University Press.

Broe, Michael. 1992. Introduction to feature geometry. Papers in laboratory phonology II: gesture, segment, prosody, ed. by G. Docherty and D. Ladd. Cambridge: Cambridge University Press.

Broselow, Ellen. 1995. Skeletal positions and moras. The handbook of phonological theory, edited by J. A. Goldsmith. Cambridge, MA: Blackwell, pp. 175-205.

Cahill, Mike. 1995. Nasal assimilation and labiovelar geometry. Paper presented at ACAL 26 [Also ms., SIL and Ohio State University, 1994].

Cahill, Mike; Frederick Parkinson. 1997a. Overgeneration in Feature Class Theory. Paper presented at the annual meeting of the LSA.

Cahill, Mike; Frederick Parkinson. 1997b. Partial class behavior and Feature Geometry: remarks on Feature Class Theory. Proceedings of the Norh East Linguistics Society 27. Amherst, MA: GLSA, Department of Linguistics, University of Massachusetts.

Carlson, R., B. Granström, and Gunnar Fant. 1970. Some studies concerning perception of isolated vowels. Speech transmission laboratory - quarterly progress and status report. Royal Institute of Technology, Stockholm 2-3:19-35.

Chomsky, Noam; Morris Halle. 1968. The sound pattern of English. New York: Harper & Row.

Clements, G. N. 1985. The geometry of phonological features. Phonology Yearbook 2. 225-52.

Clements, G. N. 1987. Phonological feature representation and the description of intrusive stops. Proceedings of the Chicago Linguistic Society 23: Parasession on Autosegmental and Metrical Phonology. Chicago: Chicago Linguistic Society. 29-50

Clements, G. N. 1991. Place of articulation in consonants and vowels: a unified theory. Working Papers of the Cornell Phonetics Laboratory 5. Ithaca: Cornell University. 77-123.

Clements, G.N.; Elizabeth V. Hume. 1995. The internal organization of speech sounds. The handbook of phonological theory, ed. by J. A. Goldsmith. Cambridge, MA: Blackwell. 245-306.

Clements, G. N.; Samuel Jay Keyser. 1983. CV Phonology: A Generative Theory of the Syllable. Cambridge, MA: MIT Press.

Clements, G. N.; Engin Sezer. 1982. Vowel and consonant disharmony in Turkish. The Structure of Phonological Representations 2, ed. by H. v. d. Hulst and N. Smith. Dordrecht: Foris. 213-55.
Cole, Jennifer; Charles W. Kisseberth. 1994. An Optimal Domains theory of harmony. Urbana-Champaign: Cognitive Science Technical Report UIUC-BI-CS-94-02, University of Illinois.
Cook, T.L. 1969. The pronunciation of Efik for speakers of English. Bloomington, IN: African Studies Program.
Cressey, William W. 1974. Homorganic in generative phonology. Papers in linguistics 7. 69-82.
de Jong, Kenneth. 1995. On the status of redundant features: the case of backing and rounding in American English. In Phonology and phonetic evidence: papers in laboratory phonology iv, edited by B. Connell and A. Arvaniti. Cambridge: Cambridge University Press, 68-86.
Doke, Clement M. 1926. The phonetics of the Zulu language. Johannesburg: University of Witwatersrand Press.
Doke, Clement M. 1931. Text-book of Zulu grammar. London: Longmans, Green and Co.
Donegan, Patricia. 1978. The natural phonology of vowels. Columbus, OH: Ohio State University dissertation.
Ewen, Colin J.; Harry van der Hulst. 1988. [high], [low] and [back], or [I], [A],and [U]. Linguistics in the Netherlands 1988, ed. by P. Coopman and A. Hulk. Dordrecht: Foris.
Frisch, Stefan. 1996. Similarity and frequency in phonology. Evanston, IL: Northwestern University dissertation.
Frisch, Stefan; Michael Broe; Janet Pierrehumbert. 1995. The role of similarity in phonotactic constraints. Evanston, IL: Northwestern University ms.
Gafos, Adamantios. 1996. The articulatory basis of locality in phonology. Baltimore, MD: Johns Hopkins University dissertation.
Gazdar, Gerald; Ewan Klein; Geoffrey Pullum; Ivan Sag. 1985. Generalized phrase structure grammar. Cambridge, MA: Harvard University Press.
Goad, Heather. 1993. On the configuration of height features. Los Angeles: University of Southern California dissertation.
Goldsmith, John. 1976. Autosegmental phonology. Cambridge, MA: MIT dissertation [published 1979, Garland: New York].
Goldsmith, John. 1981. Subsegmentals in Spanish phonology: An autosegmental approach. Linguistic Symposium on Romance Languages no. 9, ed. by W. W. Cressey and D. J. Napoli. Washington, D.C: Georgetown University Press. 1-16.
Goldsmith, John. 1990. Autosegmental and metrical phonology. Cambridge, MA: Blackwell.
Greenberg, Joseph. 1950. The patterning of root morphemes in Semitic. Word 6. 162-181.
Haiman, John. 1972. Phonological targets and unmarked structures. Language 48.2. 365-377.
Halle, Morris. 1983. On distinctive features and their articulatory implementation. Natural Language and Linguistic Theory 1. 91-105.
Halle, Morris. 1995. Feature geometry and feature spreading. Linguistic Inquiry 26.1. 1-46.
Halle, Morris, Bert Vaux, and Andrew Wolfe. 2000. On feature spreading and the representation of place of articulation. Linguistic inquiry 31.3. 387-444.
Harris, James. 1969. Spanish Phonology. Cambridge, MA: MIT Press.
Harris, James. 1983. Syllable Structure and Stress in Spanish: A Nonlinear Analysis. Cambridge, MA: MIT Press.

Harris, James. 1984. Autosegmental phonology, Lexical Phonology, and Spanish nasals. Language sound structure, ed. by M. Aronoff and R. T. Oehrle. Cambridge, MA: MIT Press. 67-82.
Hayes, Bruce. 1990. Diphthongization and coindexing. Phonology 7. 31-71.
Hulst, Harry van der. 1988. The geometry of vocalic features. Features, Segmental Structure and Harmony Processes 2, ed. by H. v. d. Hulst and N. Smith. Dordrecht: Foris. 77-125.
Hulst, Harry van der; Norval Smith. 1985. Vowel features and umlaut in Djingili, Nyangumarda and Warlpiri. Phonology 2. 277-303.
Hulst, Harry van der; Norval Smith. 1987. Vowel harmony in Khalkha and Buriat (East Mongolian). Linguistics in the Netherlands, ed. by F. Beukema and P. Coopmans. Dordrecht: Foris. 81-91.
Hulst, Harry van der; J. van de Weijer. 1991. Topics in Turkish phonology. Turkish Linguistics Today, ed. by H. Boeschoten and L. Verhoeven. Kinderhook, NY: E.J.Brill. 11-59.
Hyman, Larry. 1988. Underspecification and vowel height transfer in Esimbi. Phonology 5. 255-73.
Inkelas, Sharon, Orhan Orgun, and Cheryl Zoll. 1997. The implications of lexical exceptions for the nature of grammar. Derivations and constraints in phonology, ed. by Iggy Roca. Oxford: Oxford University Press. 393-418.
Itô, Junko. 1986. Syllable theory in prosodic phonology. Amherst, MA: University of Massachusetts dissertation.
Itô, Junko. 1989. A prosodic theory of epenthesis. Natural Language and Linguistic Theory 7. 217-259.
Itô, Junko; Armin Mester. 1993. Licensed segments and safe paths. Constraint-Based Theories in Multilinear Phonology, ed. by C. Paradis and D. LaCharité. 197-213.
Itô, Junko, and Armin Mester. To appear. Markedness and word structure: OCP effects in Japanese. Cambridge, MA: MIT Press.
Jakobson, Roman, Gunnar Fant, and Morris Halle. 1952. Preliminaries to Speech Analysis. Cambridge, MA: MIT Press.
Jakobson, Roman; Morris Halle. 1956. Fundamentals of language. The Hague: Mouton.
Kager, René. 1999. Optimality Theory. Cambridge: Cambridge University Press.
Kaun, Abigail. 1995. The typology of rounding harmony: an Optimality Theoretic account. Los Angeles: UCLA dissertation.
Keating, Patricia. 1988. Palatals as complex segments: x-ray evidence. UCLA working papers in phonetics 69: UCLA.
Kiparsky, Paul. 1982. Lexical phonology and morphology. In Linguistics in the Morning Calm, edited by I. S. Yang. Seoul: Hanshin, pp. 3-91.
Kiparsky, Paul. 1999. Paradigm effects and opacity. Stanford, CA: Ms., Stanford University.
Kirchner, Robert. 1993. Round and back vowel harmony and disharmony: An optimality theoretic account. Los Angeles: UCLA ms.
Korn, David. 1969. Types of labial vowel harmony in the Turkic languages. Anthropological linguistics 11.3. 98-106.
Ladefoged, Peter, and Ian Maddieson. 1996. The sounds of the world's languages. Cambridge, MA: Blackwell.
Lass, Roger. 1976. English Phonology and Phonological Theory. Cambridge: Cambridge University Press.

Lass, Roger; John M. Anderson. 1975. Old English phonology. Vol. 14, Cambridge Studies in Linguistics. Cambridge: Cambridge University Press.
Leben, Will. 1973. Suprasegmental Phonology. Cambridge, MA: MIT dissertation.
Lees, Robert. 1961. The phonology of Modern Standard Turkish, Uralic and Altaic Series, vol. 6. Bloomington: Indiana University Publications.
Levin, Juliette. 1985. A Metrical Theory of Syllabicity. Cambridge, MA: MIT dissertation.
Lombardi, Linda. 1990. The nonlinear organization of the affricate. Natural Language and Linguistic Theory 8. 375-425.
Lombardi, Linda. 1991. Laryngeal features and laryngeal neutralization. Amherst, MA: University of Massachusetts dissertation [published 1994, Garland: New York].
Lombardi, Linda. 1995. Laryngeal features and privativity. The linguistic review 12. 35-59.
Maddieson, Ian; Peter Ladefoged. 1989. Multiply articulated segments and the feature hierarchy. UCLA working papers in phonetics 72: UCLA.
Mascaró, Joan. 1976. Catalan phonology and the phonological cycle. Cambridge, MA: MIT dissertation.
Mascaró, Joan. 1983. Phonological levels and assimilatory processes. Barcelona: Universitat Autònoma de Barcelona ms [of talk given at given at GLOW, 1983].
Massaro, Dominic W. 1998. Perceiving talking faces: from speech perception to a behavioral principle. Cambridge, MA: MIT Press.
McCarthy, John J. 1979. Formal Problems in Semitic Phonology and Morphology. Cambridge, MA: MIT dissertation.
McCarthy, John J. 1981. A prosodic theory of nonconcatenative morphology. Linguistic Inquiry 12. 373-418.
McCarthy, John J. 1985. Features and tiers: Semitic root structure constraints revisited. Handout of talk delivered at the University of Illinois at Urbana, October 1985.
McCarthy, John J. 1986. OCP Effects: Gemination and antigemination. Linguistic Inquiry 17. 207-263.
McCarthy, John J. 1988. Feature geometry and dependency: A review. Phonetica 43. 84-108.
McCarthy, John J. 1991. On guttural transparency. Handout of talk presented at the Distinctive Features Workshop, 1991 LSA Institute, Santa Cruz, CA.
McCarthy, John J. 1994. The phonetics and phonology of Semitic pharyngeals. Papers in Laboratory Phonology III, ed. by P. A. Keating. Cambridge: Cambridge University Press. 191-233.
McCarthy, John J. to appear. The fundamentals of Optimality Theory. Cambridge: Cambridge University Press.
McCarthy, John J; Alan Prince. 1993a. Generalized alignment. Yearbook of morphology, ed. by G. Booij and J. v. Marle.
McCarthy, John J., and Alan Prince. 1993b. Prosodic Morphology I: Constraint Interaction and Satisfaction: Ms., University of Massachusetts, Amherst, and Rutgers University, New Brunswick, NJ.
McCarthy, John J.; Alan S. Prince. 1995. Faithfulness and reduplicative identity. University of Massachusetts Occasional Papers in Linguistics 18. Amherst, MA: GLSA, Department of Linguistics, University of Massachusetts. 249-384.
Mester, Armin. 1986. Studies in tier structure. Amherst, MA: University of Massachusetts dissertation.

Mohanan, K. P. 1983. The Structure of the melody. Cambridge, MA: MIT ms.
Ní Chiosáin, Máire. 1994a. Irish palatalisation and the representation of place features. Phonology 11.1. 89-106.
Ní Chiosáin, Máire. 1994b. Barra Gaelic vowel copy and (non-)constituent spreading. Proceedings of the West Coast Conference on Formal Linguistics 13. Stanford, CA: CSLI Publications (Distributed by University of Chicago Press).
Ní Chiosáin, Máire; Jaye Padgett. to appear. Markedness, segment realization, and locality in spreading. In Linda Lombardi, ed., Constraints and representations: segmental phonology in Optimality Theory, Cambridge University Press.
Odden, David. 1991. Vowel geometry. Phonology 8. 261-289.
Ohala, John J. 1990. The phonetics and phonology of aspects of assimilation. Papers in Laboratory Phonology, ed. by J. Kingston and M. Beckman. Cambridge: Cambridge University Press. 258-75.
Ohala, John J.; James Lorentz. 1977. The story of [w]: an exercise in phonetic explanation for sound patterns. Proceedings of the the Berkeley Linguistics Society 3. Berkeley, CA: BLS, UC Berkeley.
Ohala, John J.; Manjari Ohala. 1993. The phonetics of nasal phonology: theorems and data. Nasals, nasalization, and the velum, ed. by M. K. Huffman and R. A. Krakow. New York: Academic Press. 225-49.
Padgett, Jaye. 1991. Stricture in feature geometry. Amherst, MA: University of Massachusetts dissertation.
Padgett, Jaye. 1995a. Stricture in feature geometry. Revised version of Padgett 1991, published by CSLI Press, Stanford, CA.
Padgett, Jaye. 1995b. Partial class behavior and nasal place assimilation. In Proceedings of the South Western Optimality Theory Workshop 1995. Tuscon, AZ: The University of Arizona Coyote Papers, pp. 145-183.
Padgett, Jaye. 1995c. Feature classes. University of Massachusetts Occasional Papers in Linguistics 18. Amherst, MA: GLSA, Department of Linguistics, University of Massachusetts.
Painter, Colin. 1970. Gonja: a phonological and grammatical study. Bloomington, IN: Indiana University Press.
Pierrehumbert, Janet. 1993. Dissimilarity in the Arabic verbal roots. Proceedings of the Norh East Linguistics Society 23. 367-81.
Polgárdi, Krisztina. 1994. Turkish vowel harmony in Optimality Theory. Leiden: HIL/Leiden University ms.
Prince, Alan. 1983. Relating to the grid. Linguistic Inquiry 14:19-100.
Prince, Alan; Paul Smolensky. 1993. Optimality theory: constraint interaction in generative grammar. Rutgers University, New Brunswick, NJ, and University of Colorado, Boulder ms.
Pulleyblank, Douglas. 1993. Vowel harmony and Optimality Theory. Actas do workshop sobre fonologia: University of Coimbra, Portugal. 1-18.
Recasens, Daniel. 1991. Fonètica Descriptiva del Català. Barcelona: Institut D'Estudis Catalans, Biblioteca Filològia XXI.
Recasens, Daniel. 1993. Fonètica i fonologia. Barcelona: Biblioteca Universitària 18, Enciclopèdia Catalana.

Reinecke, B. 1972. The structure of the Nkonya language. Leipzig: VEB Verlag Enzyklopädie.
Sagey, Elizabeth. 1986. The Representation of Features and Relations in Nonlinear Phonology. Cambridge, MA: MIT dissertation.
Sagey, Elizabeth. 1987. Non-constituent spreading in Barra Gaelic. Irvine, CA: School of Social Sciences, University of California, Irvine, ms.
Schane, Sanford. 1972. [back] and [round]. A festschrift for Morris Halle, ed. by S. R. Anderson and P. Kiparsky. New York: Holt, Rinehart and Winston. 174-184.
Schane, Sanford. 1984. The fundamentals of particle phonology. Phonology 1. 129-55.
Scobbie, James. 1991. Attribute value phonology. Edinburgh: University of Edinburgh dissertation.
Selkirk, E.O. 1991a. Major place in the vowel space: vowel height features. Amherst, MA: University of Massachusetts ms.
Selkirk, Elisabeth. 1991b. A two-root theory of length. University of Massachusetts Occasional Papers in Linguistics 14: Papers in Phonology. Amherst, MA: GLSA, Department of Linguistics, University of Massachusetts.
Skorik, Petr IAkovlevich. 1961. Grammatika chukotskogo iazyka. Moscow: Akademiia Nauk SSSR.
Smolensky, Paul. 1993. Optimality, markedness, and underspecification. Paper presented at the Rutgers University Optimality Workshop, New Brunswick, NJ. Revised version available on the Rutgers Optimality Archive: http://ruccs.rutgers.edu/roa.html.
Steriade, Donca. 1981. Parameters of metrical vowel harmony rules: the underlying representation of harmonic vowels. Cambridge, MA: MIT ms.
Steriade, Donca. 1982. Greek Prosodies and the Nature of Syllabification. Cambridge, MA: MIT dissertation.
Steriade, Donca. 1987. Locality conditions and feature geometry. Proceedings of The Norh East Linguistics Society 17. Amherst, MA: GLSA, Department of Linguistics, University of Massachusetts. 595-617.
Steriade, Donca. 1995a. Underspecification and markedness. Handbook of Phonological Theory, ed. by J. Goldsmith. Cambridge, MA: Blackwell. 114-74.
Steriade, Donca. 1995b. Positional neutralization. Los Angeles: Ms., UCLA.
Stevens, Kenneth N. 1998. Acoustic phonetics. Cambridge, MA: MIT Press.
Stevens, Kenneth; Samuel Jay Keyser. 1989. Primary features and their enhancements in consonants. Language 65. 81-106.
Stevens, Kenneth N.; Samuel J. Keyser; Haruko Kawasaki. 1986. Toward a phonetic and phonological theory of redundant features. Invariance and variability in speech processes, ed. by J. S. Perkell and D. H. Klatt. Hillsdale, NJ: Lawrence Erlbaum. 426-49.
Straight, H. Stephen. 1976. The acquisition of maya phonology: variation in Yucatec child language. New York: Garland Press.
Svantesson, Jan-Olof. 1985. Vowel harmony shift in Mongolian. Lingua 67:283-327.
Thráinsson, Höskuldur. 1978. On the phonology of Icelandic preaspiration. Nordic Journal of Linguistics 1. 3-54.
Trigo, L. 1988. On the phonological derivation and behavior of nasal glides. Doctoral dissertation, MIT.
Trubetzkoy, N. 1939. Grundzüge der phonologie. Güttingen: Vandenhoeck and Ruprecht.

Vago, Robert. 1980. The sound pattern of Hungarian. Washington, DC: Georgetown University Press.
Walker, Rachel. 1998. Nasalization, neutral segments, and locality: Ph.D. dissertation, University of California, Santa Cruz.
Welmers, William. 1962. The phonology of Kpelle. Journal of African Languages 1. 69-93.
Welmers, William E. 1969. The morphology of Kpelle nominals. Journal of African languages 8:73-101.
Yip, Moira. 1989. Feature geometry and cooccurrence restrictions. Phonology 6. 349-374.
Zimmer, Karl. 1969. Psychological correlates of some Turkish morpheme structure conditions. Language 45. 309-321.

Endnotes

[*] Many people have contributed to the improvement of this work in its long gestation. I would like to thank especially Jill Beckman, Donka Farkas, Junko Itô, Ewan Klein, Linda Lombardi, Armin Mester, Máire Ní Chiosáin, Alan Prince, and Geoff Pullum.

[1] Diagrams such as (3)b might seem to violate a line crossing prohibition, unless it is borne in mind that the relevant features (here [labial] and [dorsal]) occupy separate tiers in the representation (McCarthy 1985, Sagey 1986). Hence the features spread on different planes, something difficult to convey in a two-dimensional image.

[2] Compare the feature classes argued for by Anderson and Ewen (1987), in important respects distinct from those posited within Feature Geometry. Classes posited within Dependency Phonology are generally motivated by evidence which is itself distinct from that motivating Feature Geometry, for example, lenition, fortition, and sonority hierarchies.

[3] Thoughout this paper the familiar features [back] and [round] of Chomsky and Halle (1968) are used for the articulatory dimensions of backness and roundness, though this is not crucial to the arguments.

[4] There is also substantial evidence for a feature class Height, as shown in (7) (and perhaps including [ATR]). See Hyman (1988), Clements (1991), Odden (1991), Selkirk (1991a), and Goad (1993).

[5] Another kind of constraint is widely employed already in the Optimality Theory literature to achieve feature spreading: the constraint Align. (See McCarthy and Prince 1993a on alignment, and for assimilation as alignment Kirchner 1993, Smolensky 1993, Pulleyblank 1993, Cole and Kisseberth 1994, and Akinlabi 1997). Alignment could be made to serve our purposes, but an interesting point arises here: alignment as typically construed incorporates directionality into constraint statements, while directionality of spreading in fact follows from independently motivated factors in many cases. The source of vowel harmony—accounting for its apparent directionality—is typically a vowel in a more prominent position, for example a stressed or word-initial syllable, as Steriade (1995a) notes. Given this fact, it would be odd to stipulate the direction of spreading. A similar point was made by classical licensing accounts in phonology. The constraint employed here therefore abstracts away from directionality.

[6] Though I do not demonstrate the point, the analysis here holds to Richness of the Base (see Prince and Smolensy 1993 on this notion). The choice of an underspecified suffix vowel here is a matter of convenience.

[7] Recall that the representations assumed imply that all consonants in a harmony span bear the harmonizing feature. It is assumed for concreteness that the root-final lateral underlyingly shares its [round] value with the preceding vowel, a state of affairs that arguably obtains if consonants bear the vocalic features of tautosyllabic vowels generally (see Gafos 1996, Ní Chiosáin and Padgett to appear), and assuming Lexicon Optimization (Prince and Smolensky 1993). However, the analysis can be carried through without these assumptions, for instance, assuming that only

vowels are specified for color features generally, and that consonants are skipped by harmony (except for just those consonants and color features that are able to cooccur, such as the lateral and [back]).

[8] Cahill and Parkinson (1997b) disagree with the distribution reported in Padgett (1995b) of partial and total nasal place assimilation to labiovelars in Gã, citing published sources and their own fieldwork. In particular, they claim that partial assimilation occurs within morphemes as well as across morpheme boundaries. The four forms cited to illustrate this are relatively long, and more investigation into their potential morphological complexity could be warranted. Of the two native speakers I have consulted with (both linguists), one (Paul Kotey, p.c.) states that the facts of assimilation in Gã do not depend on morphological context; he further claims that it is basically always total, e.g. ŋ͡m-k͡pai 'my cheeks', ŋ͡mk͡pai 'libation', but allows that it might be partial (regardless of morphological context) perhaps in some dialects, or in casual speech. (Cahill and Parkinson similarly suggest some variation between partial and total assimilation.) The other speaker (Comfort Wentum, p.c.) maintains that there is indeed a distinction as stated in Padgett (1995b), ŋ-k͡pai vs. ŋ͡mk͡pai. Though several sources support the existence of partial nasal place assimilation in the language, clarification of the precise facts is warranted.

[9] On the hardening effect evident in e.g. [ŋgw] and [ŋgj], see Padgett (1991/95a).

[10] The notion of perceptual recoverability is independent of the mode of perception. The [m] of [ŋ͡mg͡b] is recoverable, unlike the [g] of [mg͡b], because the labial nasal is visible. Visual cues are known to have a very significant effect on a listener's perception (see for example Massaro 1998).

[11] Binary features such as [back] might themselves be seen as classes that consist of two members, e.g. {[+back], [-back]}. Given the independently required notion of feature class in phonology, the debate about whether all features are privative (single valued) or not may be moot. As Frisch (1996) notes, we might declare that all features are privative, as many have, without sacrificing the recurrent unity of [back] and [front]: they make up a privileged set or feature class Back. Under this conception, [back] blocks spreading of [front] (and vice versa) not because of line crossing, as is commonly assumed with binary [back], but because of feature incompatibility: spreading [back] onto a segment specified for [front] violates a (presumably universal) feature cooccurrence condition *[back, front]. This explanation works only assuming that the relevant segment cannot be skipped entirely, and hence relies on a relatively strict notion of locality.

[12] Most Russian roots are monosyllabic, e.g. greb- 'dig', koz- 'goat', tolk- 'explain'. They include both consonants and vowels, unlike the all-consonantal roots of Semitic (as traditionally conceived), and hence 'non-adjacent' here means separated by consonant or vowel. See Padgett (1991/95a) for details omitted here about the facts and analysis.

[13] The 'dorsal' group includes the postalveolars [č], [š] and [ž]; these consonants group with the velars in their cooccurrence behavior, occurring significantly infrequently with them. This is because within roots these segments derive historically from velar segments in most cases. It is not of direct concern here whether we regard them as synchronic velars (at some level of representation) or simply view this part of the generalization as opaque in contemporary Russian.

[14] Note how this account for neutralization to default place requires a serial derivation: first the Place node is delinked, then it is presumably re-inserted along with [coronal], since [coronal] is dependent on Place. Such an account cannot translate directly into a non-serial theory based on output constraints: a constraint such as *Place would rule out coronal [n] as well as non-coronal [m], [ɲ], etc. In fact, this is true whether *Place refers to a node in Feature Geometry or a feature class as in Feature Class Theory. Perhaps for this reason, neutralization to [coronal] place is not treated as a feature class generalization at all within existing accounts in Optimality Theory, but instead is taken to follow from the ranking of constraints such as *labial and *dorsal over faithfulness. (See Prince and Smolensky 1993, Smolensky 1993.) None of these comments apply to cases where place is prohibited altogether, as with the mora nasal of Japanese (section 5).

[15] Cahill and Parkinson also state that convincing evidence for partial class spreading must involve two or more (but not all, naturally) features of a class spreading, rather than one out of two as in Turkish or Kɔɔni. No basis is offered for this judgement, however, and the reasoning is unclear. It is hard to find cases fitting the description Cahill and Parkinson suggest, simply because in most cases for independent reasons it cannot happen that three or more members of a feature class occur within a single segment. For example, to see a case of nasal place assimilation involving two out of three features spreading, we would first need to find a triggering consonant having three major places of articulation. Such consonants are at best extremely rare.