Chapter 3

Radical CV Phonology: the categorial gesture

Harry van der Hulst

1 Goals

'Radical CV Phonology' is a variant of Dependency Phonology (Anderson and Jones 1974, Anderson and Ewen 1987). The symbols C and V do *not* refer to skeletal units in the sense of Clements and Keyser (1983), but to two phonological features, which play a pivotal role in this chapter and in the theory I develop here. Radical CV Phonology shares with Dependency Phonology most of its 'leading ideas' and tries to further develop the execution of these in specific domains. In this chapter the domain I focus on is *segmental structure* and, specifically, non-place properties.

The organization of this chapter is as follows. In Section 2 I will offer an overview of the framework of Dependency Phonology as presented in Anderson and Ewen (1987), henceforth AE, limiting myself to the segmental domain.[1] In Section 3 I will then present a critique of AE's model and in Section 4 I proceed with outlining the Radical CV alternative. Section 5 contains a summary of the main points of my proposal and specifies areas for further research.

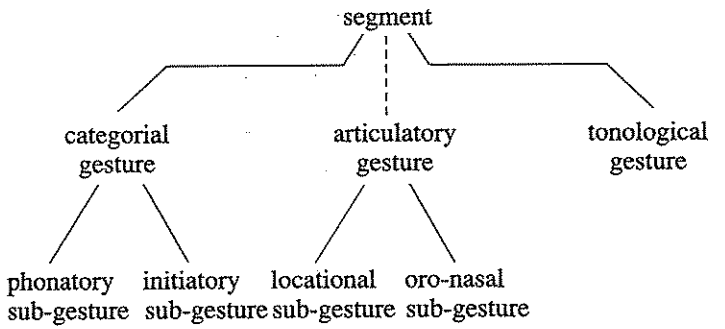
2 An introduction to Dependency Phonology

The most fundamental principle of Dependency Phonology (henceforth DP) is the idea that units (or constituents) which are combined to form higher level units (or constituents) enter into a head – dependency relation. With specific reference to the level of segmental organization, we can formulate further leading ideas of DP as follows: the primes of phonology ('features') form constituents *within phonological segments*, which are called *gestures* (comparable to 'feature classes dominated by a class node' in

Clements 1985a, but not to the 'gestures' in Browman and Goldstein 1989 which are defined in their model as the primitive actions of the vocal tract articulators). A further central claim is that the features are *privative*, and are called *components* (comparable to 'unary features'). The term *component* will not be used here, however, and instead I will use the term *element* (cf. Kaye et al. 1985, Harris and Lindsey this volume, Brockhaus this volume).

DP recognizes two major gestures, the *categorial* and the *articulatory* gesture, and, in addition, a *tonological* gesture. Both major gestures contain two sub-gestures and all four sub-gestures contain a number of elements. The topic of this chapter will be on the categorial gesture, which in standard DP contains a *phonatory* sub-gesture (for elements expressing manner or stricture properties) and an *initiatory* sub-gesture (for elements expressing airstream properties and glottal states). The articulatory gesture contains the *locational* sub-gesture (with elements for place properties) and an *oro-nasal* sub-gesture containing just one element (viz. nasal), as in (1) below.

(1) *Anderson and Ewen (1987)*



The relevance of feature grouping has long been recognized in DP. Already in Lass and Anderson (1975) and Lass (1976) a number of specific arguments are put forward that support the view that the matrix characterizing the segment should be split up into at least two sub-matrices, or gestures, the phonatory and articulatory gestures of Lass and Anderson (1975), or the laryngeal and oral gestures of Lass (1976). This subdivision into phonatory/laryngeal and articulatory/oral feature sets reflects the fact that phonological rules and processes can refer precisely to (e.g. delete) either of these gestures, the other gesture being unaffected (cf. the so-called 'stability effects' of Autosegmental Phonology). Lass (1976) discusses cases of reductions of full consonants to the glottal consonants [h]

and glottal stop, [ʔ], as occurring for instance in many varieties of Scots (cf. also Lass 1984: 113–5), which show the independence of the laryngeal features *vis-à-vis* the oral features, a proposal also made in Thrainsson (1978) on the basis of Icelandic preaspiration data and subsequently in various versions of Feature Geometry (cf. Pulleyblank this volume). It is these two gestures which, together with the latter added initiatory and oro-nasal sub-gestures, and the recently introduced tonological gesture, are the primary ingredients of the most recent DP feature tree, defended in AE, and represented in (1).[2]

The proposals which AE make for the tonological gesture are sketchy; I will return to them below (AE 1987: Section 7.5). First, let us have a look at the content of the four sub-gestures which have been more fully developed, starting off with the phonatory sub-gesture. This gesture contains two elements, |V| and |C|, which AE define as follows (recall that AE use the term component instead of element):

> |V|, a component which can be defined as 'relatively periodic', and |C|, a component of 'periodic energy reduction'. (p. 151)

They then continue:

> |V| and |C| differ from the [Jakobsonian] vocalic and consonantal distinctive features in that the presence of, say, |V| in a segment does not necessarily imply that the segment is in a simple binary opposition to an otherwise identical segment not containing |V|. Rather . . . the more prominent a particular . . . component . . . the greater the preponderance of the property characterised by that component. Notice too that |V| and |C| can characterise segments either alone or in combination. (p. 151)

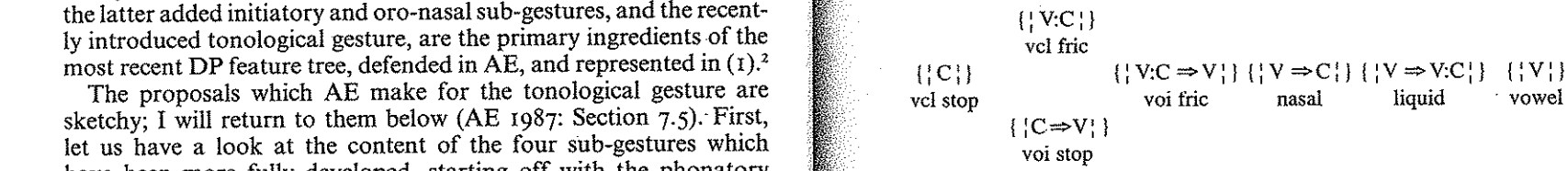
'Prominence' of elements is expressed in terms of a head-dependent relation. In DP, elements can not only be joined by simple, symmetrical combination, but they can also enter into a relationship in which either element is more important, the other element being dependent on it. In addition, two elements can even entertain a relation in which neither feature is dominant, a relationship which DP call 'mutual/bilateral dependency'. Thus we arrive at the set of dependency relationships in (2).[3]

(2)

(a) {|X; Y|} or {|X ⇒ Y|}: Y is dependent on X
(b) {|Y; X|} or {|Y ⇒ X|}: X is dependent on Y
(c) {|X: Y|} or {|X ⇐⇒ Y|}: X and Y are mutually dependent

These dependency relations hand DP the tools to express a number of major segment classes in terms of combinations of |V| and |C|, as in (3), where 'vcl' is 'voiceless', 'voi' is 'voiced' and 'fric' is 'fricative':

(3)



Underneath the actual representations I have indicated what classes of segments they represent. AE argue that the representations reflect a sonority ranking, going from left to right, in which the classes of voiceless fricatives and voiced stops are claimed to have equal sonority. Further distinctions (leading to separate representations for laterals, strident fricatives, etc.) will be discussed in Section 3.

We see here that, as stated in the above quote, the precise phonetic interpretation of the elements |C| and |V| is determined by their status in a structure. Roughly, the phonetic impact of the dependent occurrence of an element is less than the impact of that same element as a head. Note also that we can, if we wish, associate traditional feature names to these interpretations. For example, in the above array of structures, an ungoverned |V| can be glossed as [(+)sonorant], whereas a governed |V| forms the equivalent of [(+)voice]. This particular example reveals that DP manages to express distinct but clearly related phonological categories in terms of a single primitive appearing in different structural positions, where traditional feature systems must stipulate a relation in the form of redundancy rules like [+sonorant] − > [+voice]. In DP [+sonorant] and [+voice] are manifestations of one and the same element, viz. |V|. Thus the relation between these two phonetic events is 'built in' into the basic vocabulary.

In order to characterize the same segment classes in a feature system of the SPE (Chomsky and Halle 1968) type we would need the features [voice], [consonantal], [continuant] and [sonorant] (Clements 1990), where DP uses just two single-valued features, the elements |C| and |V| and their interdependencies. However, pure reductionism has not been AE's primary motivation for replacing major class and manner features by CV-complexes. Their foremost

claim is that their approach is more adequate than traditional binary theories in at least three respects.

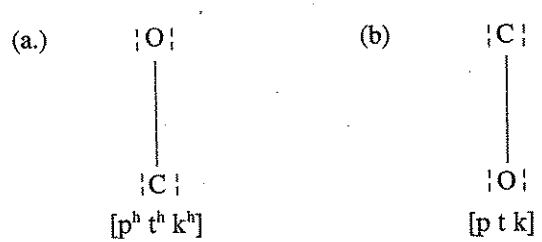
First, by replacing binary features with structures of varying complexity, representations more adequately reflect the relative markedness of phonological major class and manner categories. In (3), the categories *vowel* and *voiceless stop* are the least complex which reflects their relative unmarked status. Fricatives are more complex than stops, and voiced obstruents are more complex than voiceless obstruents. This again reflects well-known and widely accepted claims regarding the relative markedness of these categories.

Second, as stated earlier, AE also claim that the array of structures provides an adequate characterization of the notion of relative sonority. Degree of sonority corresponds to the amount of 'V-ness' that a representation contains. We could likewise define *strength* in terms of the amount of 'C-ness'.[4]

Third, the structures composed of |C| and |V| provide a more adequate basis for the expression of phonological processes. With reference to (3) AE note that these structures reflect an asymmetry in the behaviour of 'voicedness', as opposed to 'unvoicedness'. If we assume (as most phonologists do) that phonological rules can only cause phonetic events by manipulating phonological units, the structures in (3) express that languages can spread 'voicing' but not the absence thereof. If this is empirically correct, representations as in (3) are superior to binary feature systems in which [+voice] and [−voice] have the same status.[5]

I now turn to the second sub-gesture of the categorial gesture, viz. the initiatory sub-gesture. This sub-gesture contains the 'glottal opening' element |O| and two elements used for the description of different types of airstream mechanisms, |G| (for 'glottalicness') and |K| (for 'velaric suction'). These three elements can each enter into a dependency relation with an element or elements of the phonatory sub-gesture, as in (4), in which the contrast between aspirated and unaspirated voiceless stops is represented in Dependency terms (cf. Ewen 1980: 9.4, Ewen 1986: 204).

(4)



I emphasize here that DP allows, then, that representations of the two sub-gestures display variable dependency relations. A similar relationship can also be observed between the |G| element of the initiatory sub-gesture and the |C| of the phonatory sub-gesture, used to differentiate between glottalic ingressive ({|C; G|}) and egressive ({|G; C|}) sounds. Below we will see that dependency relations between sub-gestures are also allowed in the articulatory gesture between the nasality element |n| and the features of the locational sub-gesture (cf. 6). So in addition to dependency between individual elements DP allows dependency between representations of sub-gestures belonging to the same gesture.

Proceeding with this sketch of DP, let us turn to the daughters of the articulatory gesture, which are the locational sub-gesture and the oro-nasal sub-gesture. The place elements, listed in (5), belong to the former:

(5) *DP place elements*

|i|	'palatality, acuteness/sharpness'	|l|	'linguality'
|u|	'roundness, gravity/flatness'	|t|	'apicality'
|a|	'lowness, sonority'	|d|	'dentality'
|@|	'centrality'	|r|	'retracted tongue root'
|T|	'Advanced Tongue Root (ATR)'	|L|	'laterality'

The heart of the set of place elements is formed by the familiar 'aiu' subset, which plays a key role in the representations of vowels and consonants. Two further elements are added for vowels, centrality and ATR, as well as a set of elements which are mainly or exclusively used for consonants.

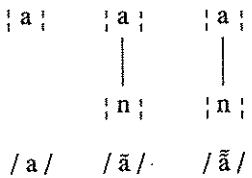
It is irrelevant for our present purposes that DP uses precisely these elements for the representation of the place of articulation of the vowels and consonants. I refer to AE (Chapter 6) for details and to van der Hulst (1993, 1994) for an alternative in the spirit of the present chapter.

Then, finally, there is the oro-nasal sub-gesture, which contains precisely one element, |n|, for 'nasality'. Recall that there also is a phonatory characterization of nasality: {|V ⇒ C|}. One might wonder whether DP really needs a nasality element, or, if it turns out that such an element is necessary, whether this element should have a sub-gesture entirely for itself. With respect to the first question AE argue that nasal consonants not only form a natural class with other sonorant consonants by sharing certain characteristics in their categorial (particularly phonatory) representations, they also form a natural class with nasalized segments, which may

have different specifications in the categorial gesture. In order for this latter natural class to be reflected by the DP representations of the segments in question, Dependency Phonologists argue that we need a separate element, |n|.

The oro-nasal sub-gesture falls under the same gesture as the locational sub-gesture. AE claim that there is at least one language in which the nasality element shows dependency relations with elements of the locational sub-gesture, viz. Chinantec (cf. Ladefoged 1971: 34; Catford 1977: 138). In this language there appears to be a distinction between two different degrees of nasalization, to be represented in DP terms as in (6) (cf. AE 1987: 250):

(6)



The question as to whether the nasality element should occupy a (sub-)gesture of its own, is rather more difficult to answer. In DP phonetic considerations have always played a central role in the justification and motivation of its primitives and hierarchical organizations. Although Catford (1977) recognizes only three functional aspects in the specification of speech (correlating with DP's phonatory, initiatory and locational sub-gestures), Ladefoged (1971) distinguishes four aspects required in speech specification; he adds the oro-nasal aspect. On the basis of Ladefoged's sub-classification into four aspects, AE (1987: 148) conclude that 'it seems possible, then, to account for the oro-nasal process as a distinct sub-gesture within the articulatory gesture', and that hence a subdivision into two sub-gestures, just as in the categorial gesture, 'is perhaps not inappropriate for the articulatory gesture'. Notice, though, that the motivation for a separate oro-nasal sub-gesture does not appear to be overwhelming.

Proposals within Feature Geometry (cf. Pulleyblank this volume) have sometimes also adopted a separate node for the feature nasal (cf. Sagey 1986, 1988). Piggott (1990, 1992) proposes a 'velic class node' dominating only [nasal]. In addition, he adopts a node 'spontaneous voicing', which also may dominate a feature nasal. The duplication of nasality in Piggott's model clearly bears a resemblance to the way DP treats nasality.

Finally, we briefly look at AE's proposals for the tonological gesture. In their excursus on representations for tonal distinctions, AE (1987: 273) make the intriguing suggestion that the elements |i| and |u| (as part of the tonological gesture) could be employed for high and low tone, respectively:

> we propose that the appropriate representations for the two tonal components are . . . |i| and |u|. In other words, we are suggesting that |i| and |u| in the tonological gesture bear the same relation to |i| and |u| in the articulatory gesture as |V| in the categorial gesture does to |a| in the articulatory gesture . . . That is, |i| involves (relatively) 'high frequency' and |u| (relatively) 'low frequency'; whether this is interpreted as high (or low) F_0 or as concentration of energy in the higher (or lower) regions of the spectrum depends on the context – i.e. gesture – in which it occurs.

What is most noticeable in this proposal is the idea to use the same elements, viz. |i| and |u| in two different gestures. In my own proposals I will make quite crucial use of this strategy. To emphasize that this strategy can be traced back to AE's own proposals, I will here also quote AE (1987: 215) on their suggestion concerning the identity of |a| and |V|:

> there is clearly a relationship between |a|, as a component within the articulatory gesture, and |V|, as a component of the categorial gesture. Consider the acoustic glosses which we have given the two components: |V| corresponds with maximal periodicity, and |a| with maximal sonority. Vowels, by virtue of their periodicity are the most sonorous of the categorial segment-types, while open vowels are the most sonorous within the class of vowels . . . The open unrounded vowel, then, might have {|V|} both as the representation of the categorial gesture and of the articulatory gesture.

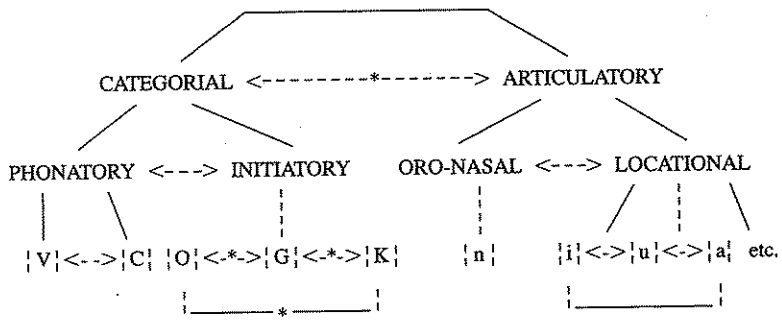
The importance of these quotes is to show that AE themselves suggest the strategy to employ the same elements in different (sub-)gesture, thus deriving similarities in phonetic interpretation, while attributing the differences to the fact that the '(sub-)gestural location' of an element has a bearing on the phonetic interpretation as well. Thus elements are interpreted taking into account not only their position in the head-dependent relation but also taking into account the sub-gesture they are part of. It is precisely this line of reasoning that I fully explore in my own alternative.

3 A critique of classical Dependency Phonology

3.1 Inter- and intrasub-gestural dependency

We have seen that DP explores the possibility of allowing sub-gestures to enter into dependency relations. The possibility of entering elements of the initiatory and phonatory sub-gestures into a dependency relationship is not, however, fully exploited: while it is apparently necessary for |O| and |G| to be able to entertain dependency relations with |C|, |K| cannot 'look beyond' the initiatory sub-gesture, there being no DP representations in which (combinations of) |C| and/or |V| entertain non-symmetrical relations with |K| alone. In addition, intrasub-gestural relationships are not exhaustively employed either, since we do not find dependency relations between the features contained in the initiatory sub-gesture. Schematically, all this is summarized in (7); a '*' indicates that no dependency relations are proposed between the units connected by the bidirectional arrow:

(7)



In (7) I also encode that there are no dependency relationships between the two main gestures: there are no circumstances under which segment-types are distinguished by means of a difference in the dependency relation holding between the elements of the categorial and articulatory gestures.

It is unclear why AE use precisely the dependencies illustrated in (7) and no others. In an attempt to restrict the DP model, Davenport and Staun (1986) have argued to dispense with *inter*sub-gesture dependency. They show that once the glottal opening element |O| is assigned to the phonatory sub-gesture and a new element |I| ('initiator velocity', expressing the direction of airflow) is assigned to the initiatory sub-gesture, no need remains for dependency relations between the phonatory and the initiatory sub-gestures. In the present chapter, I do not want to discuss the specifics of Davenport and Staun's (1986) modifications of the DP framework, but my own proposals agree with theirs in disallowing variable intersub-gestural dependency.

AE also exploit the possibility of allowing variable dependency between the two sub-gestures of the articulatory gesture (cf. 6). I tend to look upon two distinctive degrees of nasalization with great suspicion, but apart from that, it is very questionable whether there is any need for an articulatory sub-gesture containing the element |n|. If, as I argue in the next subsection, no such sub-gesture is postulated, it follows, for that reason alone, that no intersub-gestural dependency relations can be postulated within the articulatory gesture.

The upshot of the above points is that in my own model I will dispense with variable dependency relations between sub-gestures. In the next subsections I will offer critical remarks concerning the various DP sub-gestures and mention some alternatives proposed by others which anticipate my own proposals. I will not, however, discuss the locational sub-gesture since this is treated separately in van der Hulst (1994).

3.2 The oro-nasal sub-gesture

First, although there is abundant evidence for suggesting that nasality can spread autosegmentally, and can hence function independently of other elements, this does not in itself suggest that the element expressing nasality should occupy its own (sub-)gesture. Since the oro-nasal sub-gesture dominates precisely one element, it is impossible to make out on empirical grounds whether, in a case of nasal harmony, it is the oro-nasal sub-gesture that spreads or rather the |n| element individually: in either case we derive the same result. Hence some other argument should be found that could support the relevance of the oro-nasal sub-gesture. We have seen that AE only provide a general phonetic line of reasoning based on suggestions of Ladefoged.

Another question that comes up is why, in the tree in (1), the oro-nasal sub-gesture should be grouped together with the locational features under the articulatory gesture. Perhaps, intuitively, this assumption makes some sense in that the feature [nasal] can only apply to place of articulation features, and cannot have scope over the elements grouped under the categorial gesture. Yet, there does not appear to be any phonological evidence from the area of assimilation processes that could corroborate or disconfirm the constituent-hood of the velic and locational features.

Noting that DP expresses nasality in two ways, Davenport (1993) proposes to dispense with the element |n| altogether. This

implies that the categorial characterization of nasality 'survives', although Davenport's proposal is that nasality is not expressed in the phonatory sub-gesture (i.e. not in terms of specific |C|/|V| combination), but as a separate element |N| in the initiatory sub-gesture. So, in a sense, Davenport's proposal is a compromise between the two 'old' ways of expressing nasality in DP. I refer to Davenport who shows that the dual representation of nasality leads to unsatisfactory results in DP, but whatever these arguments are, it will be clear that if we can demonstrate that a single expression for nasality (in whichever sub-gesture) is sufficient, one of the ways nasality is expressed in DP must be eliminated, whether the dual representation creates 'problems' or not. My own proposals regarding nasality are in agreement with Davenport's. As becomes apparent in the next subsection, nasality will be represented only once in the Radical CV model.

3.3 The initiatory sub-gesture

Davenport and Staun (1986) maintain an initiatory sub-gesture, which contains elements for airstream distinctions: |I| 'egressive airflow', |G| 'glottalicness' and |K| 'velaric suction'; |O| which forms part of this sub-gesture in AE has been moved to the phonatory sub-gesture in their model. Furthermore, we have just seen that Davenport (1993) proposes to add an element |N| 'nasal' to the initiatory sub-gesture. In my own proposal the equivalent of Davenport's initiatory sub-gesture contains no elements like |I| and |K|, but it agrees with his proposals in expressing glottal and nasal/oral distinctions in a single sub-gesture.

3.4 The phonatory sub-gesture

I will now turn to a more extensive evaluation of the organization of the phonatory sub-gesture and argue that the 'syntax' of CV combinations is not clearly defined in AE's version of DP, a point also emphasized in den Dikken and van der Hulst (1988), who offer an alternative which can be seen as the earliest predecessor of the proposals I advance in the next section. For convenience I repeat here the set of distinctions built from |C| and |V| which AE propose as a kind of core set:

(8)

{|C|}	{|V: C|}	{|V: C ⇒ V|}	{|V⇒ C|}	{|V ⇒ V: C|}	{|V|}
vcl stop	vcl fric	voi fric	nasal	liquid	vowel
	{|C ⇒ V|}				
	voi stop				

Given the combinations which are employed one may wonder why many other possible combinations of |C| and |V| are *not* used in AE's model, e.g.:

(9)

{|C ⇒ V: C|} {|V: C ⇒ C|} ...

As AE do not fail to observe themselves, |V: C| represents [continuant] and it would seem therefore that DP has three and not just two categorial features, and, we might add, redundancy statements ruling out combinations of [consonantal] (|C|) and [continuant] (|V: C|), for which there seems to be no use (cf. 1987: 9).

AE do add a number of additional, more complex representations to capture further distinctions:

(10)

{|V: C ⇐⇒ V|}	{|V ⇒ V: C ⇒ C|}	{|V: C ⇐⇒ V ⇒ C|}	{|V: C ⇒ C|}
fricative trill	lateral	voiced lateral fricative	non-sibilant fricative

The argumentation that AE provide in favour of these representations is crucially dependent on the representations in (8) which form the starting point. Fricative trills may pattern with voiced fricatives in conditioning phonological processes for which AE discuss 'Aitken's Law' as an example. Given the representation in (10), the relevant natural class can be represented as {V: C ⇒ V} (cf. 1987: fn. 3). Lateral liquids, of course, must be distinguished from r-sounds, which motivates the second structure in (10):

> laterals are phonetically unique, as far as the phonatory sub-gesture is concerned, in having effectively two manners of articulation. While there is a stricture of open approximation at one or both sides of the mouth (at least for sonorant laterals), there is also closure in the centre of the oral tract. . . . Essentially, then, the |C| node characterizes a secondary . . stricture type within the phonatory sub-gesture. (AE 1987: 163)

The extra dependent |C| in the third representation, then, also adds laterality to the fricatives (p. 164).

The fourth structure reflects the distinction between sibilant and non-sibilant fricatives:

> /s/ may be interpreted as the optimal fricative phonetically; acoustically it shows the 'simplest' combination of consonantal and vocalic properties, while the other fricatives involve energy reduction in various frequency bands. In comparison with the sibilants, then, the other fricatives display extra /C/-ness. (p. 166)

Even though AE are careful in motivating the structures in (8) and (10), one starts having serious doubts concerning the restrictiveness of their approach. Assuming that a theory of segmental structure aims at characterizing a closed set of well-formed representations which matches the set of attested phonological distinctions, I must conclude that classical DP does not do very well in this respect. The 'syntax' underlying combinations of elements (and sub-gestures) is not explicitly defined, i.e. we do not know what the total set of possible dependency structures is. This implies that AE make no serious attempt to come to grips with the notion 'possible phonological segment'.

Despite this criticism, I believe that the basic ideas of DP as well as the specific proposed structures are extremely interesting, even though some of the perhaps most fruitful ideas (like using the same elements in different sub-gestures) are left unexplored. At the same time, due to the absence of restrictions and explicit hypotheses, the theoretical status of DP as a research programme is rather weak. This is perhaps the main reason why this model has not been taken up by many phonologists who now, independently, develop ideas which are quite similar to those characteristic of Dependency Phonology (cf. den Dikken and van der Hulst 1988, van der Hulst and van de Weijer 1994 for a discussion of this point).

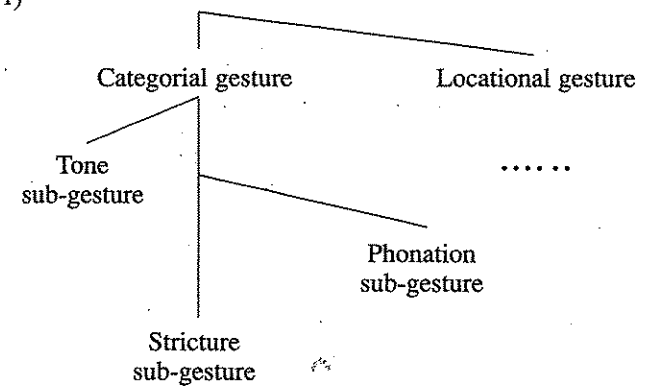
3.5 Summary and preview

In the next section I will propose a different architecture for the categorial gesture, arguing that this gesture comprises three sub-gestures which I will call: *Stricture*, *Phonation* and *Tone*; I use capital letters when referring to the sub-gestures in Radical CV Phonology. Thus (for the moment ignoring details concerning the internal organization of the sub-gestures), I refer to Anderson and Ewen's 'phonatory sub-gesture' as the Stricture sub-gesture. I then use the term Phonation sub-gesture for some of the properties they express under their 'initiatory sub-gesture' (abandoning the latter term), and I incorporate their tonological gesture into the categorial gesture as the Tone sub-gesture.

I will only deal with Anderson and Ewen's articulatory gesture where I am concerned with the representation of nasality. In standard DP, nasality has a dual representation, as we have seen. It is represented in terms of a categorial (more specifically, phonatory) characterization (for nasal consonants) and as a separate element |n| under the oro-nasal sub-gesture of the articulatory gesture (for nasal consonants and nasal vowels). I will propose to represent nasality in categorial terms only, expressing it in the Phonation sub-gesture. Thus I arrive at the position that the articulatory

gesture deals exclusively with locational or place distinctions (cf. van der Hulst 1993, 1994)

(11)



We have seen that in standard DP, sub-gestures, like elements, may enter into dependency relations with each other, such that different dependency relations give rise to different segment types. Apart from rearranging its content, a second aspect of my proposal for the categorial gesture is that its sub-gestures enter into a *fixed* dependency relation: the Stricture sub-gesture is the head and the other two sub-gestures are dependents.

A third aspect of my proposal concerns the choice of elements. In standard DP, each sub-gesture has its own set of elements. The use of the same elements in different sub-gesture is hinted at but by no means fully explored. In the proposal advanced here I claim that *all* sub-gestures contain exactly two elements. A logically independent claim will be that in each case we are dealing with the same pair: |C| and |V|. In van der Hulst (1993, 1994) I discuss how this works out for the Locational gesture. In this chapter I will demonstrate that not only Stricture distinctions, but also Phonation and Tonal distinctions can be represented in terms of |C|, |V| and a fixed set of CV-combinations. The main goal of this exercise is not to arrive at a totally new set of distinctive categories but rather to *re*construct the more or less accepted set of distinctive features in such a way that (a) the set is not a random list but instead a well-defined subset of the logically possible |C|/|V| combinations and (b) relations between separate features are not arbitrary since they turn out to involve (partially) identical |C|/|V| combinations occurring in different sub-gestures.

4 An alternative approach: Radical CV Phonology

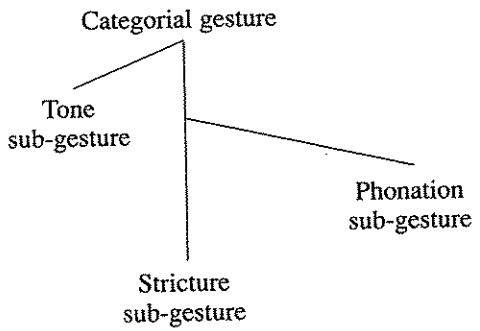
4.1 Outline

In this section I will propose a strict and uniform syntax to form CV-combinations. I will propose CV-structures for *all* categorial distinctions (which explains the name of this theory: *Radical CV Phonology*).

The phonetic interpretation of the two elements is, as can be expected, fairly general. Nevertheless, I will suggest that these elements do have a phonetic (i.e. acoustic and articulatory) interpretation: C denotes articulatory events which are referred to as *closure*, *stricture* or *contraction* (and their acoustic effects). The phonetic interpretation of V involves the opposite or the absence of these C-type events, leading to a relative high degree of *sonorancy*. Depending on the structural position of C and V (in terms of dependency and 'hosting' sub-gesture) specific interpretations (compatible with the general interpretations) arise. By mainly using articulatory glosses I do not intend to disagree with Harris and Lindsey (this volume) who claim that the primary meaning of elements is a mental acoustic image. I do not, however, subscribe to their view that all elements are independently pronounceable. This may be true of locational elements in so far as they occur in vowel structures, but that is simply a result of the fact that pure vowels (i.e. a, i, u) have no categorial elements in them. A discussion of this issue would take me beyond the scope of the present chapter, however.

I will propose that the categorial gesture consists of three sub-gestures which enter into a fixed dependency relation:

(12)



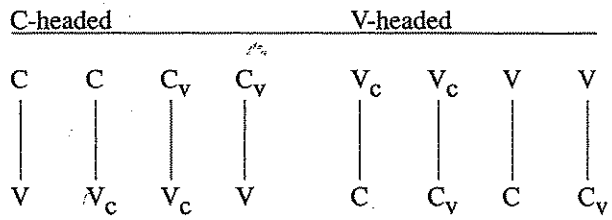
In each of these we find the four simplest structures which can be composed from C and V:

(13)

(a)	{|C|}	{|C ⇒ V|}	{|V ⇒ C|}	{|V|}
(b)	C	C_v	V_c	V

(13a) is the standard DP notation, while (13b) gives a simplified notation which I will henceforth use. I will assume, then, that elements do not enter into a dependency relation with themselves within a sub-gesture, since it will turn out that the four-way distinction in (13) is sufficient for all sub-gestures, given that within each sub-gesture we allow instead combinations of the four structures in (13) can occur. I will represent these combinations as in (14):

(14)

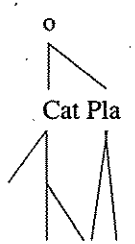


I will refer to the structures in (13) as simple structures and to those in (14) as complex structures. I assume that complex structures consist of two simple structures with non-identical heads. It will turn out, however, that the distinction between C- and V-headed complex structures is distinctive in the Stricture sub-gesture only. For the other two sub-gestures a total of four complex structures will be sufficient.

In van der Hulst (1993, 1994) the current approach is extended to cover place distinctions. It is argued that the place gesture is composed of two sub-gestures. The categorial and the place gestures, as we may expect, enter into a (fixed) dependency relation in which the categorial gesture is the head:

(15)

(15)



The argumentation for taking the categorial gesture as the head is based on the fact that categorial distinctions (and specifically stricture distinctions) determine the distribution of segments in the syllabic organization. Being head properties we expect them to be 'visible' in the root node. A further indication comes from spreading behaviour. I assume that the head-dependent asymmetry is manifested in spreading processes in such a way that dependent properties can spread independently, while heads can only spread together with their dependents. It is well known that stricture properties do *not* spread, while place properties do. This confirms the head status of the categorial gesture. Notice, however, that I do not claim that the dependent categorial sub-gestures Tone and Phonation are incapable of spreading. In fact, the two layers of dependency in (12) are meant to reflect that Tone properties are more likely to be 'prosodic' (i.e. autosegmental) than Phonation properties. Tone elements form, so to speak, the outermost shell of the categorial gesture.

The diagram in (15) is *not* meant to express *linear order* of elements within the segment. I will assume that linear order is only specified at the root level, but in this chapter I will not discuss 'supra root' gestures (cf. van de Weijer and van der Hulst, in preparation).

Concluding this section, let me note that a structure as in (15) is not unlike the kinds of structures which have been proposed within *Geometrical Phonology* (Clements 1985a, Sagey 1986, 1988, McCarthy 1988, Pulleyblank this volume). My approach differs from that line of work in that I assume that the adoption of structural relations can, and if possible *must*, be counterbalanced by a reduction of the number of phonological primes. Implicit to this point is of course the criticism of Geometrical Phonology that this approach has taken for granted that the hierarchical relations must simply be *added* to the set of features which stems from the SPE tradition. This criticism is not undermined by the fact that certain changes *vis-à-vis* the SPE system have been adopted, since it seems that these changes are not at all determined by the adoption of (a specific) hierarchical structure.

The addition of structure to the segment allows us to recognize different traditional features as 'allofeatures' of the same 'featureme'. Obviously, in order to explore the reduction strategy that I suggest, we need independent support for the particular grouping we assume and both phonetic and phonological constraints on assigning 'allofeatures' to a single 'featureme'. I suggest the following criteria. First, the two features must be in complementary distribution in the sense that they occur in different sub-gestures. Second, one might argue that the two phonetic events corresponding to allofeatures must be *similar*; this would be the phonetic constraint.

Third, we must be aware of (historical) phonological processes which reveal their affinity; this is the phonological constraint.

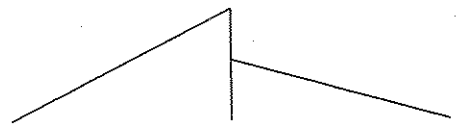
For a phonologist, the reduction strategy and the constraints I mention here must have a familiar ring. After all, it is common practice to argue along similar lines when we try to establish the minimal set of phonemes for a particular language. Two sounds will be attributed to a single phonological category (a single phoneme) if they occur in complementary distribution (i.e. occur in different structural positions) and if, in addition, there is phonetic similarity, and, furthermore, both are involved in phonological alternations.

In the next subsection, I will discuss whether there is a reasonable match between the structures allowed by our syntax and the distinctions which are generally considered contrastive in the analysis of segment inventories. In this enterprise, I must rely on a certain consensus with respect to which phonetic properties are potentially distinctive. Such a consensus is apparent from the fact that certain features appear to be widely accepted, or that certain relations between features have been taken to be well-established. Despite consensus there is, of course, a lot of uncertainty as well. In the model proposed here I will be forced to make decisions which await further empirical underpinning.

4.2 The proposal

The following diagram anticipates most of my proposals regarding the match between the elements in the Categorial gesture and traditionally recognized distinctive features:

(16) Categorial gesture



tone		*stricture*		*phonation*	
C	high tone	C	stop	C	constricted glottis
C_v	low tone	C_v	continuant	C_v	spread glottis
V_c	high register	V_c	sonorant cns	V_c	nasal voice
V	low register	V	vowel	V	oral voice

In each sub-gesture I will create further distinctions in terms of the combinations given in (15). I will now discuss the sub-gestures one by one starting with the head sub-gesture: Stricture.

4.2.1 The Stricture sub-gesture

The elements C and V, when part of the Stricture sub-gesture, correspond to the following articulatory events and their acoustic effects:

(17)

C = relatively high degree of stricture (as in obstruents)
V = unimpeded outflow of air (as in sonorants)

A dependent C differentiates between two types of 'high degree of stricture' and a dependent V between two types of 'unimpeded outflow of air':

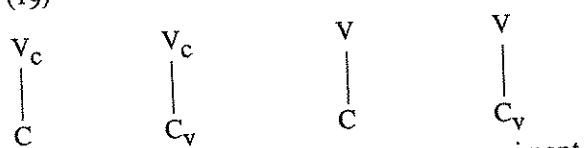
(18)

C = absolute stricture (as in stops)
C_v = non-absolute stricture (as in fricatives)
V = unimpeded outflow of air (as in vowels)
V_c = unimpeded outflow but not necessarily centrally or uninterrupted (as in sonorant consonants)

Both V and V_c express what has been called *spontaneous voicing*. In (18) we express directly that stops and oral vowels are unmarked with respect to their stricture: stops are unmarked obstruents (as opposed to fricatives) and oral vowels are unmarked sonorants (as opposed to nasal vowels); the *oral* outflow is clearly the default option. This is in accordance with the fact that the prototypical unmarked syllable consists of a stop followed by an oral vowel, as well as with many other well-known generalizations stemming from the study of language acquisition, language change and aphasia (cf. Jakobson 1941 for this line of reasoning).

I now turn to the interpretation of complex stricture structures, starting with the V-headed ones. There are four possible structures (cf. 19). I have added the proposed interpretations below the complex structures:

(19)



lateral liquid | fricative rhotic liquid | flap | approximant

In the first complex structure, the dependent C expresses 'the secondary central stricture' which lateral liquids have: the governed C stands for (central) closure. In (18) we have seen that ungoverned C represents stops. To say that the same stricture occurs in laterals (and nasals, which I will represent below as 'nasal liquids') allows us to capture the natural class of stops, liquids and nasals: 'bare' C (whether head or dependent). In general terms we expect all classes denoted by a common substructure to be natural. In this particular example, traditional feature systems can only succeed in grouping stops, nasals and liquids by attributing the redundant [−continuant] to laterals and nasal consonants.[6] At the same time, since the bare C occurs in different structural positions in stops as opposed to lateral and nasal sonorants, we can 'choose' to differentiate between stops and sonorants by assuming that processes can make reference to bare C *as a head*, i.e. to stops alone. It is reasonable to expect that the reverse is excluded, i.e. that one cannot refer to a property in a 'weak' (i.e. dependent) form and ignoring its stronger manifestation, but this is a point that needs further investigation.

There is of course a phonetic difference between stops and nasals on the one hand and laterals on the other hand: the oral stricture is *total* in the case of stops and nasals, and *partial* (i.e. central) in the case of laterals. This does not take away that an element such as C has a constant phonetic interpretation, viz. *contact stricture* (as opposed to non-contact stricture). The reason why in nasal consonants C is realized as a total closure is that nasals, but not laterals, allow the air to escape through the nasal cavity. Hence nasals are sonorant in having free outflow of air, but since the nasal escape route is chosen, the oral closure is 'free' to be optimal. In laterals there is no nasal outflow of air and the oral stricture must therefore be partial.

The class of rhotics is very diverse. Still, phonologically, there appears to be a single category (cf. Lindau 1980), except when we find a so called fricative rhotic next to a 'normal' rhotic (as in Czech). A system opposing a normal rhotic to a lateral can be represented as having both V_c (the rhotic) and $V_c \Rightarrow C$ (the lateral). In systems which have no lateral/rhotic contrast, the liquid can often vary from rhotic to lateral depending on contextual factors. The precise factors which condition this allophonic variation are not always clear (cf. Bhat 1974). In some cases the choice depends on the manner properties of surrounding segments: the rhotic occurs in intervocalic position and the lateral elsewhere. This suggests that rhotics have a weaker constriction than laterals since the environment V-V counts as a weakening or lenition

context (cf. 21(c)). The structures in (19) are adequate in this respect. In other cases the rhotic appears in a post-coronal context (Ewe, Ganda; cf. Halle and Clements 1983) which implies that place factors are involved. (Cf. van de Weijer 1993b for a study of this type of allophony.)

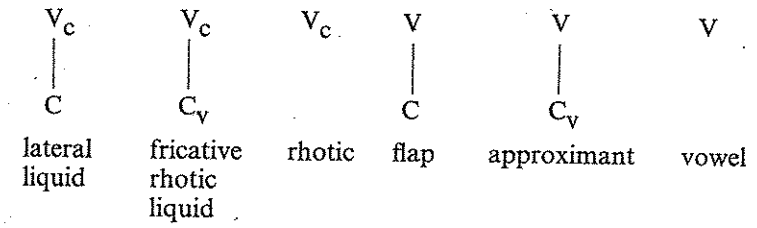
The second structure in (19) will be taken as the representation of a fricative /r/, which AE also argue for (cf. above). This r-type shares with continuants obstruent the substructure C_v and may thus be expected to pattern with (voiced) fricatives, as, in fact, it does in Aitken's law (cf. above).

For the third structure I propose the interpretation 'flap' or 'tap'. I follow Maddieson (1984) in the claim that taps and flaps (although different from an articulatory point of view) do not involve distinct phonemic categories. The taps or flaps have clear consonantal properties in that they involve a contact stricture, but they fall in the class of sonorants in being spontaneously voiced. Below, I demonstrate how a flap can be seen as an expected natural weakening of stops, just like approximants are weakening products of fricatives.

Finally, for the fourth structure I propose to interpret this as *approximant*, assuming that approximants may be categorially different from vowels and not just positionally different (i.e. in terms of syllabic position). The approximants /y/ and /w/ in Dutch, for example, have traditionally always been analysed as sonorant consonants and the phonology of Dutch seems to offer no reason to treat these sounds as vowels at any level of representation. Another classic case where a lexical distinction must be made between a vocalic and consonantal labial approximant is found in French in the pair *l'oiseau*, *le whisky*, where the shape of the definite article reveals that in the former case the initial approximant is a vowel, while in the latter case it must be analysed as a consonant.

The representations in (19) predict that rhotics are more sonorous than laterals, because the latter have an additional C. Hankamer and Aissen (1974: 137–8) discuss this matter and conclude that the relative ordering of laterals and rhotics on the sonority scale is not universally fixed. In certain cases rhotics are clearly more sonorous than laterals, e.g. syllable final *-rl* is allowed but *-lr* is not (for example in German), but sometimes, at least according to Hankamer and Aissen, it is the other way around. I tentatively suggest that in that case the 'lateral' is not a liquid but a weaker sound falling in the category of flaps. Example (20) represents the full sonority scale for sonorants, showing increasing sonority going from left to right:

(20)



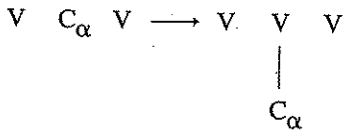
The representations proposed here for taps/flaps and approximants provide a good basis for dealing with lenition processes:

(21)

(a) stop → flap

(b) fricative → approximant

(c) *Weakening*



In this view, then, weakening is interpreted as an assimilatory process. Another view is expressed in the version of element theory found in government phonology, where lenition is seen as the loss of complexity (cf. Harris and Lindsey this volume).

A further advantage of our sonority scale (when compared to that of AE) is that nasal consonants have no place on it. This is preferred because nasals never occur as steps in weakening chains, as Davenport (1994) observes, who, as we have seen, also removes nasality from the stricture sub-gesture.

We will now turn to obstruents. The table in (22) specifies the four possible complex structures, differing from those in (19) by being C-headed:

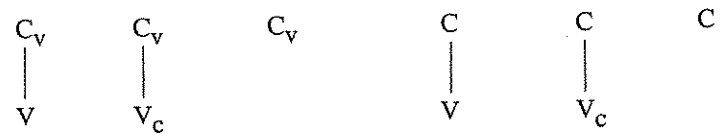
(22)

fricative	fricative	stop	stop
strident	lateral	strident	lateral

As specified in (22), I assume that we have to sub-categorize stops and fricatives in terms of *stridency* (or sibilancy) and *lateral-ity*. To claim that strident and lateral(ized) are relevant categories for fricatives is not controversial, but I would like to suggest here that the same two categories apply to stops as well: both the third and the fourth structure can be taken to represent *affricates*. The fourth is a *laterally released stop*, also referred to as *lateral affricate*. The third structure is a 'simple' affricate, which Jakobson et al. (1952), in fact, also characterize as a *strident* stop. From an articulatory point of view stridency seems to involve an extra barrier creating greater turbulence of the outflowing air. Dependent V represents the high acoustic energy (i.e. stridency) which results from this turbulence. This suggests that we must interpret V-ness not simply as sonority, but as something more abstract, viz. acoustic energy. Sonority adds to acoustic energy and so does the extra noise associated with stridency. Lombardi (1990) and van de Weijer (1994) offer other views on the representation of affricates, which involve the use of combination features like [stop] and [continuant].

We thus arrive at the following sonority scale for obstruents:

(23)



strident	lateral	plain	affricated	lateral	plain
fricative	fricative	fricative	stop	stop	stop

We note that the finer distinctions in (20) and (23) (i.e. the ones represented by complex structures) are mainly relevant for coronal place. This fact, however, finds no formal expression in this model. It is not obvious that this correlation *must* be formally expressed. Geometrical proposals which make features such as [lateral] and

[strident] formally daughters of the feature coronal run into problems anyway, as has been pointed out in various studies (e.g. Shaw 1990, Rice and Avery 1991). In this model, where the Location gesture is formally dependent on the Categorial gesture, we expect the latter to determine the former, but this geometrical relation does not exclude the other possibilities. Hence (as pointed out in van de Weijer 1993a) the implication (formulated in Shaw 1990): if lateral then coronal, or if tap/flap then coronal, reflects the dependency relation which I posit between the Categorial and the Locational gesture.

4.2.2 The phonation sub-gesture

We now turn to the 'complement' sub-gesture: Phonation. The proposal is that the elements C and V, when part of the Phonation sub-gesture, correspond to *glottal stricture* and *voicing*, respectively:

(24)

- C = glottal stricture (as in glottal stop, glottalized consonants and ejectives)
- C_v = glottal opening as in aspirated obstruents and 'voiceless' sonorants
- V = oral voice (as in sonorants and voiced obstruents)
- V_c = nasal voice (as in nasal consonants and nasalized vowels and approximants)

In (24) I express the fact that glottal (stop) and oral voice are unmarked. This is supported by the fact that the glottal stop is the most widely attested hiatus filler or default consonant and by the clear fact that nasal vowels are more marked than oral vowels.

By specifying C and C_v on obstruents I characterize aspirated and glottalized obstruents. I assume that glottalized stops and ejectives fall in the same phonological category. This point follows from the survey given in Ladefoged (1973) which tells us that these phonation types never contrast in a single language. Lombardi (1991, 1993) draws the same conclusion. C and C_v, then, represent two opposing glottal states which may be found on both obstruents and sonorants.

An advantage of the proposal I make here is that the 'features' stop/continuant and constricted/spread are considered to be instances of the same elements, i.e. C and C_v, respectively. This allows us to find a formal expression for the well-known phenomenon that stops reduce to *glottal* stop and fricatives to [h], if we assume that stops and fricative when losing their place acquire a

phonation type which 'mirrors' their original stricture type; another instance of enhancement perhaps (cf. Padgett 1991 who disputes the empirical basis for this point.)

Like C and C_v, V and V_e may also occur on both obstruents and sonorants. Added to obstruents they represent *voice* and *(pre)nasalisation*, respectively. On this view, then, (non-spontaneous) voice is a privative property, a claim which has both been supported (Mester and Itô 1989, Cho 1991a, 1991b, Lombardi 1991, 1993) and rejected (Pulleyblank this volume).

Attributing the C-phonation types to sonorants leads to glottalized and 'voiceless' sonorants. The glottalized category is obviously needed, whereas the second will be taken to lead to aspirated sonorants. This latter claim is also made in AE where they use |O| (aspiration) to represent voiceless sonorants. Cho (1990) and Lombardi (1991, 1993) make the same claim in a different model and provide a number of relevant examples.

If we say that both V and V_e may occur on sonorants as well, we allow that oral vowels and liquids have simple V. This adds, redundantly one might say, non-spontaneous voicing to these sonorants. The fact that vowels and liquids are V-headed in the Stricture sub-gesture already accounts for their spontaneous voicing.

If we can specify sonorants redundantly with the Phonation elements V we predict that they may trigger voicing assimilation. Yet, it has often been pointed out that even though sonorants are voiced, they do not trigger voicing assimilation processes. Rice and Avery (1989), however, mention the case of Sanskrit where sonorant consonants do in fact cause voicing on preceding obstruents. This shows that we should, in fact, allow sonorants to carry the element expressing (non-spontaneous) voice in specific cases.

Nasalized vowels and nasal sonorant consonants have V_e in the Phonation sub-gesture. More specifically, I will say that a lateral liquid structure, if provided with V_e phonation, *is* a nasal consonant. This predicts that /l/'s, when nasalized, turn into the nasal consonant /n/ which is an attested alternation type (cf. 29).

Notice that whereas nasal vowels are more marked than oral vowels, it is commonly assumed that the unmarked sonorant *consonant* is the nasal consonant. Perhaps this is so because in nasal consonants the phonation type V_e enhances the stricture type V_c. As Stevens and Keyser (1989) point out: enhancement (i.e. more of the same) is a common phenomenon, which may outweigh economy.

With the proposed interpretations we arrive at the following combinations of Stricture and simple Phonation structures (C_α stands for bare C or C_v; V_α stands for bare V or V_e) for obstruents and sonorants, respectively:

(25) *Obstruents*

(a) C_α V — voiced

(b) C_α V_c — (pre)nasalized

(c) C_α C — glottalized

(d) C_α C_v — aspirated

Sonorants

(e) V_α V — voiced

(f) V_α V_c — nasal(ized)

(g) V_α C — glottalized

(h) V_α C_v — aspirated

The fact that voicing and prenasalization on obstruents are represented as closely related elements finds support in the observation that (pre)nasalized obstruents are sometimes interchangeable with voiced obstruents or function as such in forming the 'voiced' counterpart of a voiceless obstruent series. We can say that in such cases the dependent phonation element is bare V, where, in the absence of a contrast with V_e, its interpretation varies between voicing and prenasalization.

Steriade (1991) suggests that fricatives cannot be *pre*nasalized because they have no 'stop phase'. This, however, does not require us to exclude nasality on fricatives, as Steriade points out herself. The result of this is simply a fully nasalized fricative.

I now turn to the interpretation of *complex* phonation structures. Complex phonation types can, in principle, be V- or C-headed, as in the Stricture sub-gesture, leading to eight different possible structures (cf. 20, 23). In actual fact, however, we need at most four:

(26)

V_e & C	V_e & C_v	V & C	V&C_v
nasal glottalized	nasal aspirated	creaky voice	breathy voice

The first and second structures represent glottalized and 'voiceless' nasals. Technically speaking, we expect to find these complex structures also on obstruents, leading to glottalized or aspirated (pre)nasalized obstruents. Such cases are rare, at best, and it is therefore tempting to suggest that the first two cases prefer to 'enhance' segment types which are V-headed in terms of their stricture.

Breathy and creaky voice are phonologically represented as combinations of aspiration plus oral voice and glottalization plus oral voice. This is a very common way of representing these glottal states, going back to Halle and Stevens (1971). Recent support for this view is offered in Lombardi (1993). In this case we may note that these two cases have a strong preference for obstruents.

If we need just four categories this means that the dependency relation holding between the two simple structures which are combined is not distinctive. This is why I represent the combinations in (26) without specifying a dependency relation. We are not denying that a dependency relation holds in every combination (a fundamental claim of DP), however. Instead, we simply say that the dependency relation is not distinctive and therefore not phonologically specified. There may be reasons for choosing one way or the other on the basis of more phonetic considerations.

Another way to go would be to encode the 'bias' of these structures in terms of the dependency relation. If the first two are only relevant for sonorant consonants (i.e. because there are no prenasalized glottalized or aspirated obstruents) we may encode this by their V-headedness, assuming, as before, that in the unmarked case structures in the different sub-gestures are in an enhancement relation (i.e. have the same head type). If we take this option we will want to represent the two other cases in (26) as C-headed, since creaky and breathy voice appear to be distinctive on obstruents only. This matter needs further investigation.

Following Ladefoged (1973), Lombardi (1991, 1993) proposes to regard *implosives* as falling in the same phonological category as consonants with creaky voice, which also includes *preglottalized* and *laryngealized* consonants (cf. Greenberg 1970). This is a welcome move. Ejectives have been analysed as glottalized sounds and in van der Hulst (1994), it is proposed (inspired by Traill 1991) that clicks are complex segments (more specifically 'double root segments', cf. van de Weijer and van der Hulst in preparation, van de Weijer 1994). Finding a place for

implosive *within* the current system, then, makes it unnecessary to look any further for 'airstream features'. To eliminate such a category was also the goal of Halle and Stevens (1971) and the present proposal shares some characteristics with their feature system.

Interestingly, Greenberg reports that in Mayan languages implosives may have a *nasal release*. Greenberg also notes that implosives cannot be directly preceded by nasals, which then typically show up as plain voiced obstruents. This shows that the phonation of implosives is, as it were, 'absorbed' by the nasal voice of the preceding nasal. Implosives may also be in contrast with prenasalized stops as in Kambera (Marian Klamer, personal communication), showing that it would be wrong to attribute both to the same category. Precisely how such phonetic effects and cooccurrence restrictions must be expressed forms a topic for further research.[7]

Summarizing, I have proposed to represent six oral and three nasal phonation types:

(27)

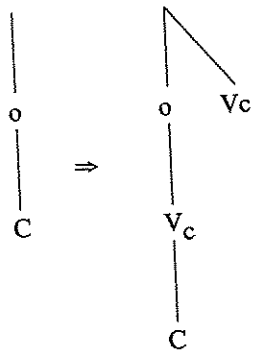
(a)	V	voiced
	C	constricted glottis
	C_v	spread glottis (i.e. aspirated)
	V ⇒ C	creaky voice, laryngealized, implosive
	V ⇒ C_v	breathy voice, murmured
(b)	V_c	nasal
	V_c ⇒ C	glottalized nasal
	V_c ⇒ C_c	aspirated nasal

Lombardi (1991, 1993) supports these types as phonological categories even though she does not count nasality among phonation.

Although in this chapter I will not investigate the formulation of processes, it is clear that in most cases we are dealing with the 'unification' of structures; for the notion of unification employed here, cf. Scobbie (1991) and Coleman (this volume). I will illustrate this with some examples taken from Rice and Avery (1989).

Nasalization of stops (as a process) may apparently give rise to plain nasals. In Korean, for example, stops assimilate to nasals:

(28)

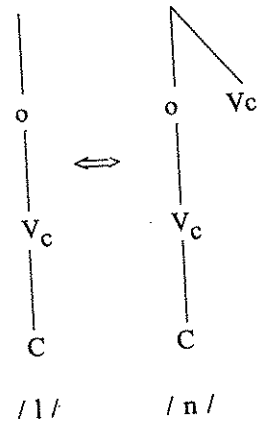


stop nasal

The nasal structure subsumes the stop structure. Note that we must allow, then, that the addition of CV-information may involve demoting a C to dependent status.

Another case I made reference to above was the alternation between /l/ and /n/, the latter occurring in a nasal context. This, for example, occurs in Yoruba (cf. Akinlabi 1992: esp. fn. 6):

(29)

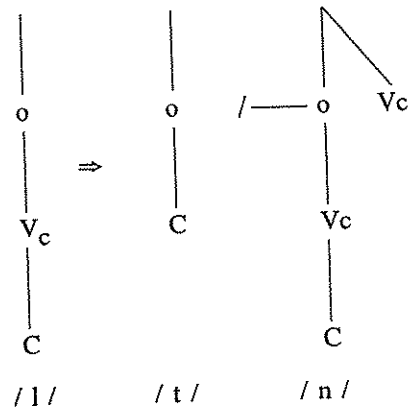


/ l / / n /

This is a case which shows the appropriateness of representing /l/ as being subsumed by /n/. Rice and Avery (1989) propose exactly the opposite, i.e. laterals have more structure than nasals. This, I would claim, makes it more difficult to deal with /l/ – /n/ alternations of this type.

Rice and Avery also discuss the case of Kuman where /l/ is changed into /t/ before /n/:

(30)



/ l / / t / / n /

We can view this as a dissimilation phenomenon, since /l/and /n/ share the substructure V_c, which is lost for /l/.

In this Section I have suggested various areas which call for further research. This implies that the representations and their associated interpretations must be taken as tentative. Yet, the main point, i.e. that we *derive* a set of structures representing a set of distinctive phonological categories which all other phonological models simply enumerate in the form of a distinctive feature list, should not be obscured by the fact that we, like all other writers on the subject, currently have insufficient data involving inventories and phonological alternations to decide on the precise set of phonation categories and their phonological representation.

4.2.3 The tonal sub-gesture

We finally turn to the 'specifier' sub-gesture Tone. Current proposals in the area of tonal phonology distinguish two 'ingredients': tone (proper), sometimes called 'melody' (Yip 1989), and register. This view goes back to Gruber (1964), and has been extensively supported in Yip (1980, 1989, 1993), Snider (1988, 1991), Inkelas et al. (1987), Hyman (1993), Bao (1991) and Duanmu (1991). In this section I will show that in this case too, such proposals fall out rather straightforwardly from the general principles which, in our view, underlie the inventory of distinctive categories. I propose that the elements C and V, when part of the Tone sub-gesture, can be taken to represent *tone* (proper) and *register*. What we have to decide is whether V represents tone and C register, or vice versa.

To find the right correlation, let us take a look at one example which involves the interplay between phonation and tone/register.

In his work on the genesis of the Vietnamese tonal system, Haudricourt (1954) proposes the following scenario:

(31)

I. *First phase* (1st century)

pa	pas > pah	paT > pa?	(T = stop)
ba	bas > bah	baT > ba?	

II. *Second phase* (6th century)

mid	falling	rising
pa	pah > pà	pa? > pá
ba	bah > bà	ba? > bá

III. *Third phase* (12th century)

mid	falling	rising	
pa	pà	pá	high register
pa	pà	pá	low register

The first phase is an automatic result of debuccalization, i.e. loss of locational elements. Then, the loss of final aspiration and glottal stop produced a tonal distinction and the loss of the voice distinction (in the onset) produced a register distinction. We thus find the following correlations:

(32)

aspiration	C_v	→ low tone
glottalization	C	→ high tone
oral voice	V	→ low register

If we assume that Tone sub-gesture structures prefer to be in an enhancement relation with Phonation, we may choose for the following interpretations for our simple Tone structures:[8]

(33)

C = high tone
C_v = low tone
V = low register
V_c = high register

Note that, whereas low register derives from voice in this view, high register comes into being as a consequence of having low register. I am unaware of evidence for or against correlating high register with nasality.

Phonetically, tone results from vocal cord vibration. The correlation between tone and glottalization/aspiration, then, has a clear physiological basis in so far as both aspects involve vocal cord activity.

In my view, register does not involve vocal cord activity. Rather register involves a more global characteristic which refers to a resonating cavity, the pharyngeal cavity, which is relatively small in case of V_c and relatively large in case of V. Likewise, V as part of the Phonation sub-gesture also refers to a resonating cavity, the oral cavity. The action of increasing or decreasing the volume of the pharyngeal cavity may involve a whole collection of articulatory activities. This interpretation of register is familiar from studies on Asian languages (cf. Matisoff 1973, Trigo 1991).

Simple tonal systems with a two-way 'H/L' opposition may be represented with either tone or register. In view of the fact that low 'tone' is usually regarded as unmarked, one is tempted to interpret this as an indication that register is used to express the opposition in those cases.

I now turn to the interpretation of complex phonation structures. Again it would seem that we need just four possibilities and we therefore have not specified a dependency relation (cf. 26):

(34)

V_c & C	V_c & C_v	V & C	V & C_c
high	high-mid	low-mid	low

What we arrive at here is simply a reconstruction of the current view on tonal representations.

Again it seems possible to assume that a three-tone system (like a two-tone system) can be represented in terms of register alone. I do not want to exclude, however, the possibility of register *and* tone playing a role in three-tone systems. In that case the mid tone would either be specified with a low or with a high tone, depending on whether it is part of the high or the low register. Such differences appear to be empirically motivated (cf. van der Hulst and Snider 1993 for a discussion of relevant cases). Of course, most current feature systems for tone have all the 'flexibilities' that we encounter here (cf. Duanmu 1991). In four tone systems, tone and register *must* both play a role and such systems are no longer ambiguous like the more impoverished systems.

There are, in addition, tonal systems which have more than four distinctive tones *per tone bearing unit* (especially in East-Asian languages), but it is usually claimed that the additional distinctions *must* involve *tonal contours*, i.e. *two* tones per tone-bearing unit. Hence such cases do not call for more tone structures, but rather for the possibility of one tone bearer being associated to more than one tone structure.[9]

A question that must be addressed at this point is whether the

possibilities in (33) and (34) can be used contrastively. Implicit in the preceding discussion was the point that once register and tone are active, the absence of neither can be used contrastively with the options in (34). Hyman (1993) makes a proposal which is similar to the one proposed here, allowing, however, precisely this contrast. He argues that tonal distinctions are made by specifying H, L or HL (which for him gives a mid tone). He then allows either of these possibilities to combine with high, low or *no* register. Hyman, then, allows the *absence* of register to be contrastive with the presence of both low and high register. I refer to his study for motivation of the resulting set of contrasts. I doubt whether we must reckon with the *contrastive* absence of register-specifications. Formally, there seems no reason why the presence or absence of tone (in our case) cannot be used distinctively. In the Phonation sub-gesture, after all, we do allow (24) and (26) (which correspond to (33) and (34) respectively) to be contrastive. This issue, then, needs more investigation.

A final question I will address is whether we wish to say that the tonal structures may be combined with both V-headed and C-headed stricture. This boils down to the question whether we want to say that obstruents may bear tone and/or register. Kingston and Solnit (1988) argue in favour of a tonal node for obstruents (or rather consonants in general) based on empirical evidence coming from case studies of tonogenesis in a variety of East-Asian languages.

As I have demonstrated in (30), the loss of certain consonantal phonation types may produce tone and register distinctions on neighbouring vowels. In our view this could be represented in two steps. First, we could 'move' the structures in (33) from the Phonation to Tone sub-gesture. In that phase, then, consonants (including obstruents) acquire a simple copy of their phonation type under the Tone sub-gesture. Subsequently, in the second phase, the phonation type gets deleted and the tone structure spreads to the neighbouring vowel. Kingston and Solnit strengthen the argument for assigning tones to consonants by showing that consonantal tone structure cannot always be regarded as simple copies of their phonation properties. Here, as in a number of cases involving tone and phonation, further study must shed more light on this question.

Finally, let us observe that the distinctive use of headedness for complex structures is only necessary for Stricture; we have not used it for phonation and tone (cf. 26 and 34). Given that Stricture is the head sub-gesture, we would indeed expect that this sub-gesture makes use of the largest variety of structures.

5 Conclusion

In spoken languages, most morphemes have a phonetic form, an acoustic event brought about by a complex articulatory activity. This phonetic form of morphemes is represented as a structured set of discrete categories, organized in a particular way. These categories correspond to mental representations of the phonetic 'sub-events'.

Virtually all theories of phonological structure are based on the assumption that there is a universal set of such categories which are basic, i.e. which cannot be decomposed in smaller sub-events and which are called *distinctive features* or *elements*. From this set languages employ a subset to make lexical distinctions.

Currently, there is no full consensus with respect to the extension of the set of elements, but the general view is that whatever the set is, its members must be enumerated. The extension of the set, then, is essentially random from a theoretical phonological point of view. The elements which are currently postulated have been established inductively, i.e. on a case-by-case basis. Whenever some language appears to make distinctive use of a previously unattested phonetic 'sub-event', a new feature is added to the list. In Feature Geometry models the claim has been made that features are organized in 'second order' categories (like a laryngeal feature category, a place feature category and so on), but the number of such categories, and their extension, is again random, from a theoretical point of view (cf. Clements 1985a, McCarthy 1988, den Dikken and van der Hulst 1988, to appear).

The central thesis of this chapter is that the set of features and feature classes is random in one specific respect only, viz. with respect to the choice of the phonetic dimensions from which the discrete categories are derived. For example: from a theoretical point of view it is random that there are tone features, phonation features and stricture features. What is considerably less random, in my view, is the specific constellation of discrete categories which is 'extracted' from these domains and perhaps it is not random either that categories are selected from *three* domains, since this produces the typical '[spec[head complement]]' pattern. Taking as a point of departure a common core of distinctive features which emerges from roughly half a century of phonological research, including more recent proposals regarding ways in which these features are organized in second order categories (i.e. place, manner, laryngeal, etc.), I have shown that the set of 'phonological features' has a quite specific structure. The above proposals for the representation of non-place properties, especially those for

phonation types and tonal distinctions, are still rather tentative. I believe, however, that the claim that the distinctions in these various sub-gestures do *not* call for three sets of totally independent phonological primes is convincing. Our approach may not be entirely successful in neatly characterizing the set of allowed combinations, yet it does succeed in deriving the specific set of 'features' and a fair amount of the relations which the tradition posits between these features. In other feature models these feature sets and the relations among the features have to be stipulated. 'Redundancy rules' of the following sort are quite typical:

(35)

(a) [+son] → [+voice]
(b) [+voice] → [+low register]
(c) [+lateral] → [−continuant]

In our model these rules find no equivalent, since in all cases we are dealing with different interpretations of the same primes.

Implicit in our approach is the claim that there is *no* innate set of features. Rather what is innate is a capacity to parse a limited set of discrete categories from the available phonetic 'scales', and a limited syntax for combining these categories.

There are two obvious lines for further research. First, the present proposal needs to be tested against a richer body of data involving both segmental inventories and phonological processes. The proposal advanced here accommodates a lot of claims in the area of non-place properties, but it also leaves a number of issues undecided. Second, the approach outlined must be applied to the place dimension as well. A third area of investigation concerns the relation between categorial features, specifically stricture features, and syllable structure. It will be clear that the current approach embodies interesting perspectives, especially if we decide to represent the syllable organization itself in terms of C and V (for onset and rhyme, respectively).

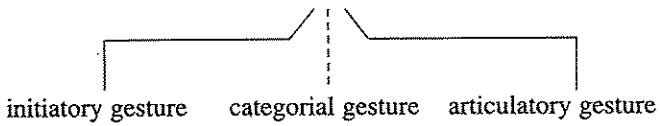
The first and third task form part of research in progress (cf. van der Hulst 1993, in preparation), whereas an attempt to perform the second task can be found in van der Hulst (1994).

Notes and Acknowledgements

This chapter is an extraction from a longer work, *Principles of Radical CV Phonology*, which contains work in progress (van der Hulst 1993). In a paper, which forms the complement of this chapter, I discuss the representation of place properties (van der Hulst 1994; cf. also van der Hulst to appear). Earlier versions of van der Hulst (1993) have circulated under the title 'Book of Segments'. During the period 1980–93 I presented a number of talks based on this material, using subtitles such as 'The molecular structure of phonological segments' and 'On the nature of phonological primes and the structure of segments'. I received useful comments on these oral presentations from Mike Davenport, John Harris, Rob Goedemans, John McCarthy, Michael Kenstowicz, Jacques Koreman, Aditi Lahiri, Simone Langeweg, David Nessly (who suggested the name of this model), Iggy Roca, Wendy Sandler, Keith Snider and Moira Yip. I am grateful to Marcel Den Dikken, Colin Ewen, Helga Humbert, Norval Smith and Jeroen van de Weijer who discussed with me some of the material presented here. I would like to thank Marian Klamer, Jeroen van de Weijer, Rint Sybesma and the editors of this volume for their comments on a prefinal version of this chapter.

1. This section draws on a similar section in den Dikken and van der Hulst (1988).
2. The tree in (1) is the, at present, final stage of a discussion within the DP framework about featural hierarchitecture, beginning with Lass and Anderson (1975), whose phonatory gesture was later renamed categorial gesture, and whose bipartite division was extended with a third main gesture, the initiatory gesture, introduced in Anderson and Ewen (1980) and Ewen (1980):

(1)



This tripartite split was not felt to suffice either, however, essentially because it was considered to be 'somewhat understructured' (Ewen 1986: 205). As a result, DP introduced sub-gestures within gestures, and eventually developed the tree in (1).

3. In the notational system of DP, an element enclosed between vertical lines represents just the element in question, while the representation {|x|} is used to exhaustively characterize a particular sub-gesture of a segment. DP also employs the notation {x}, which is used to express that the segment class in question is characterized by the element |x|, but not exhaustively so. A DP representation such as x;y denotes that |x| governs |y|, or that |y| is dependent on |x|. This representation is equivalent to the alternative representation with double arrows or the vertical notation used, e.g. in (12), below, so that x;y equals x ⇒ y as well as x.

4. In the next section I will call this claim into question with respect to the representation of nasals.
5. This discussion gets more complicated once we allow rules which may delete elements.

6. In most traditional accounts (cf. Clements 1985a, Levin 1988, Hegarty 1989) it has been argued that nasals and laterals are [– cont] just in case they pattern with stops. Such an *ad hoc* strategy is clearly superfluous in this model.
7. On the phonology and representation of implosives in a feature geometrical framework, see Lloret (1992).
8. Interestingly, C and C_v, used here for H and L tones, are also the representations for coronal and labial in the locational gesture. In a sense I therefore agree with AE's proposal to use |i| (coronal) and |u| (labial) for H and L tone, respectively.
9. Duanmu (1991) claims that tonal contours always involve two tone-bearing units.

PART 2

STRUCTURES