Phonological Primes: Gestures or Features?*

George N. Clements

1. Introduction

In their studies of the articulatory patterns that underlie speech, Browman and Goldstein (see e.g. Browman and Goldstein 1989, 1990, this issue) have brought to light some of the important respects in which the physical activity of speaking contributes to the way phonological systems are structured. From this point of view, their work has several interesting implications. One is that much of what has usually been considered as lying in the domain of discrete phonological rules may be better understood from the point of view of nondiscrete or gradient properties of articulatory organization. Thus, in various papers Browman and Goldstein have produced evidence suggesting that many types of prosodically-conditioned reduction, contextual allophony and casual speech variation reflect dynamic properties of the activity of speaking and are better modeled at the level of physical speech production than at that of more abstract categorical representation. Somewhat more ambitiously (and controversially), they propose that the dynamic properties of speech production, often viewed by phonologists as having little interest for the study of grammatical organization, play a large or even predominant role in shaping the structure of what is interpreted: the phonological system of rules and representations itself.

In this issue of *Phonetica*, Browman and Goldstein propose that a gesture-based model of phonology and phonetics can provide an alternative to models taking segments and features as their basic units. Their explicit intention is to show that "gestures are basic units of contrast among lexical items as well as units of articulatory action" and to "help clarify the differences among gestures, features, and segments." This commentary will address both of these goals. It will examine several areas in which articulatory phonology as presently conceived by Browman and Goldstein appears insufficient to account for some of the generalizations that are usually thought to lie in the domain of phonological (and phonetic) theory, and will suggest ways in which a theory of this sort might be extended in order to accommodate these generalizations. It will finally consider the status of articulation-based models of phonetic interpretation in phonological theory as a whole.

---

* Editors note: This paper will appear in a forthcoming theme issue of *Phonetica* . All instances of "this issue" in the text and references refer to this issue of *Phonetica.*

2. On the notion "gesture"

As Browman and Goldstein describe them, gestures are abstract characterizations of the formation and release of local constrictions in the vocal tract. The constrictions themselves are defined in two ways. First, gestural scores display the temporal duration of individual gestures as well as the extent of their mutual overlap. Second, tract variables specify the location and degree of the constriction formed by the coordinated set of articulators that create it. In this conception, gestures are defined not in terms of articulator movements as such, but in terms of temporary, local constrictions of the vocal tract.

There is an important division of labor between the linguistic gestural model, whose function is to construct the score, and the task dynamic model, whose function is to execute it. However, Browman and Goldstein offer only brief and informal discussion of how these two models relate to each other. Their account of phonological rules and casual speech processes implies both that gestures can be rephased by the linguistic gestural model, affecting the structure of the score, and that parameter values can be assigned to tract variables by the task dynamic model, affecting the dynamic properties of gestures. Since adjustments at both of these levels can affect the nature of the output, one would like to know more about how they are allowed to interact, and how their interaction is constrained. In particular, assignment of parameter values (which eventually include stiffness and damping ratios, see Browman and Goldstein 1990) has the potential of bring two or more related gestures "out of synch" and thus of producing output inconsistent with the score. I discuss related problems concerning gestural coordination in section 4.

Since gestures not only drive the task dynamic model but are also primitives of the representational system, they play a role similar to that of features in more familiar phonological models. Specifically, Browman and Goldstein make use of five gestures, which are defined with respect to the lips, the tongue tip, the tongue body, the velum, and the glottis. Each of these is assigned to a separate tier, and may be present or absent in a given gestural score. Thus their function is entirely analogous to that of the articulator nodes labial, coronal, dorsal, nasal, and laryngeal in models such as that of Sagey (1986). In addition, the task-dynamic model supplies the dynamical parameter values which determine the degree, shape and location of the constriction formed by each set of articulators associated with a gesture. Although these values are not inherently categorical, they are said to function as such since their ranges are subject to the constraints provided by quantal articulatory-acoustic relations and/or adaptive dispersion principles. Elsewhere (see Browman and Goldstein 1989), they have been allowed to take on a small set of discrete values, forming a subset of the places and manners of articulation recognized by traditional phonetic description. Thus, the dynamical parameters provide information analogous to that provided by binary-valued features such as [continuant], [anterior], [distributed] and so forth. Finally, constriction degree values - corresponding roughly to features such as [continuant], [sonorant], and [consonantal] - are computed from local and global configurations of the vocal tract at any given point in time by the subsystem of tube geometry (Browman and Goldstein 1989).

This brief review should be sufficient to show that the gestures and parameter values proposed in Browman and Goldstein's articulatory phonology capture roughly the same set of contrasts defined by phonological feature theory. The next two sections will compare gestures with phonological features more closely. Section 3 will argue that as far as their intrinsic properties are concerned, gestures differ from features primarily, if not uniquely, in that they are allowed to take on gradient values, a fact which renders them less appropriate than features for the representation of lexical contrasts. Section 4 will offer reason to believe that gesture-based representations, or scores, must introduce a richer notion of hierarchical structure if they are to capture a full range of phonological regularities.

3. Intrinsic Properties of Gestures and Features

In most current versions of feature theory, including the one assumed here, features are defined in terms of acoustic and aerodynamic as well as articulatory properties. Thus, for example, [spread glottis] is defined both in terms of an acoustic effect (aspiration) and the articulatory means used to produce that effect (glottal opening). Of course, feature theories differ among themselves in terms of the set of features proposed, and the properties which are taken as definitional of any given feature.

It might appear that gestures differ from features in being defined in exclusively articulatory terms. However, this is not quite accurate. The term "gesture," in Browman and Goldstein's system, is not used in its ordinary-language sense of "movement of a body part to convey emphasis or meaning," but is used in a technical sense to refer to local constrictions in the vocal tract, as we have seen. In addition, by allowing gestural parameter values to be constrained by acoustic models such as quantal theory and dispersion theory (as well as by aerodynamic models such as tube geometry), Browman and Goldstein assign acoustic and aerodynamic considerations a potentially significant role in their model. This requires a reevaluation of their claim that gestures "do not correspond to features:" to the extent that gestures are defined or constrained in part by acoustic and aerodynamic considerations, the intrinsic difference between gestures and features, as general descriptive categories, is reduced. The major differences will lie primarily in the range of values that may be assigned to them at any given level of representation, and in the way they are organized in representations, or scores.

One fundamental respect in which gestures differ from features lies in the "quantitative variation in a gesture's dynamic parameters" that we observe under prosodic or other contextual conditions, discussed by Browman and Goldstein in some detail. Dynamic parameters, including duration,[1] may be assigned any value within their (nonbinary) range. Thus, for example, gestural shrinkage has the effect of scaling down the metrical properties of a gestural event, as in the gradient reduction of the glottal gesture that reduces or eliminates aspiration in speech output. From their discussion, it is clear that Browman and Goldstein do not regard gradient operations of this sort as constrained by quantal considerations, which presumably operate at the more abstract levels of system organization and lexical contrast. In contrast, features are usually defined as binary, or in some cases one-valued, and thus contrast categorically rather than quantitatively.

In this respect, features seem to provide a more adequate unit than gestures for expressing regularities at the abstract level of lexical contrast. Underlying lexical contrasts do not require the full range of gradient parameter values needed for the description of output regularities. Instead, gestural parameters of constriction degree and location, etc., regularly behave in a categorical fashion in lexical representations and early phonological rules, and durational information is usually reduced to a simple long vs. short contrast. To explain this observation, gesture-based models must restrict the parameters defining gestural events to effectively binary values; but this is exactly what is claimed by feature theory, in which only categorial distinctions (plus vs. minus specfications of binary features, presence vs. absence of privative features) are available.

Just the same considerations hold with respect to gestural phasing relations and overlap. To account for the diversity of ways in which gestures can be timed with respect to each other, Browman and Goldstein allow several types of phasing relations, varying according to where separate gestures are aligned with each other, as well as three degrees of overlap (minimal, partial, complete). But again, this range of choices projects to an excessively large number of theoretical lexical contrasts. As far as overlap is concerned, only two types appear to be required: full vs. partial. No further distinction, such as one between partial and minimal overlap, seems required for the expression of lexical contrasts. Thus, for example, tone languages do not distinguish between two types of contour tones differing only in the relative duration assigned to their first and second components, nor are any languages known to have phonemic contrasts between "partially" and "minimally" prenasalized stops. This restriction of contrasts in overlap to two is exactly what is predicted by current feature-based phonological theories, which allow features to be related to each other in either a one-to-one or a many-to-one relation (see e.g. Clements 1985, in press, Sagey 1986, McCarthy 1988). Thus the difference between plain nasal stops and prenasalized stops can be represented as follows (simplifying irrelevant detail):

```
   [n]:              [ⁿd]:

   root              root
    |                / \
[+nasal]      [+nasal] [-nasal]
```

In the first of these representations, the single feature [+nasal] is aligned, or associated, with the other features of the segment, represented by the root node. In the second, two sequenced features [+nasal], [-nasal] are so aligned. Since association is a discrete relation, this mode of representation allows no way of distinguishing two pairs of sequenced features in terms of differences in their mutual temporal overlap with other features, and thus predicts (correctly) that such distinctions are irrelevant at the level of lexical representation.[2]

A stronger case might be made for lexical contrasts involving phase relations, i.e. the relative timing of overlapping gestures. For example, a lexical distinction between pre- and post-aspirated stops could be expressed quite naturally in terms of the relative phasing of the glottal opening gesture with stop arrest or closure in the first case, stop release in the second. Another candidate discussed by Browman and Goldstein is the distinction between voiceless and voiced aspirated stops in languages like Hindi. They suggest that the difference between these two categories can be treated as a built-in difference in phasing relations, with the glottal gesture timed later in voiced aspirated stops than in voiceless ones.

However, the theory still predicts too large a range of possible lexical contrasts in phasing relations. Browman and Goldstein currently recognize three points at which consonantal gestures can be coordinated with other gestures: onset of movement toward the target, achievement of the target, onset of movement away from the target. Allowing that two gestures can be aligned with each other at any pair of these points, this allows up to nine ways in which two consonantal gestures can be coordinated in lexical representations; and as still further gestures overlap, the number of possible phasing relations among them increases exponentially. In contrast, current feature-based theories of phonology place strong inherent constraints on the ways in which any pair of features or nodes can be coordinated in lexical representation, as we have seen, and provide no way of representing differences in phasing relations in lexical representations.

If gestures are to be used to express lexical contrasts, then, strong constraints must be placed on the way gestures can be coordinated with each other. One direction currently being explored by some phoneticians working in feature-based frameworks involves introducing a level of discrete structural positions representing "articulatory landmarks" such as stop closure and release into the representational system, to which glottal features (and perhaps others) can associate (Huffman 1989, Keating 1990). The introduction of such landmarks into gestural phonology would result in a more constrained theory of gestural coordination, though it is still unclear whether they are required for the expression of lexical contrasts. A somewhat different, though not incompatible approach draws on the view that features are defined in part in terms of acoustic and aerodynamic goals, and that these goals must be effectively realized in the output, at least in the absence of further cues to their presence. Thus, glottal features would be defined in part in terms of goals such as vocal fold vibration, aspiration, etc., rather than uniquely in terms of the glottal constriction. In this view, the fact that glottal opening in aspirated stops is not completely overlapped by the stop closure would follow automatically from the requirement that the functional goals associated with features must be effectively realized, in this case forcing the alignment of glottal opening with the arrest or release phase of a stop consonant to produce audible aspiration (i.e. with an "acoustic landmark" in the sense of Stevens 1991, or a "transition" in the sense of Hertz 1991). Similarly, the lower amplitude and extra phasing lag of the glottal opening gesture in voiced aspirated stops as compared to voiceless ones can be explained by the fact that these properties are conducive to achieving the goal of glottal vibration during the period of stop closure (Davis 1991). These and other approaches are likely to receive continuing attention in feature theory, and might have useful implications for the development of gesture-based models.

4. The hierarchical organization of phonological units

By taking gestures as primitives, Browman and Goldstein are able to describe a large range of phenomena in terms of relations among successive and overlapping gestures (constrictions). However, since the location and degree of the constrictions do not occupy separate tiers in gestural scores, Browman and Goldstein's system strongly predicts that they will always be synchronized with each other and with the activity of the articulator whose movement they constrain. Here again the gesture-based model appears to make incorrect predictions.

In many languages, we find that certain segments assimilate place of articulation - a term which I use to designate the active articulator and its constriction location - from a neighboring segment, without assimilating its constriction degree. A common process of this type consists of the assimilation of nasals to following consonants. Thus in Yoruba, the syllabic nasal which forms the progressive aspect prefix assimilates in place to a following stop, fricative, or nasal, and optionally (or speaker-variably) to a liquid. Before nonassimilated liquids and the glides /w, j, h/, on the other hand, it is realized as syllabic [ŋ] (Ward 1952; R. Ṣonaiya, personal communication).

a. obstruents:			nasal place of articulation:
N + b:	m-bɛ	'be well'	bilabial
N + f:	ɱ-fɔ	'be washing'	labiodental
N + t:	n̪-tɛ	'be spreading'	dental-alveolar
N + d:	n-dɛ	'be setting a trap for'	alveolar
N + s:	n-se	'be cooking'	alveolar
N + ʃ:	ń-ʃe	'be doing'	alveo-palatal
N + ɟ:	ń-ɟa	'be fighting'	alveo-palatal
N + k:	ŋ-ka	'be reading'	velar
N + g:	ŋ-ge	'be cutting'	velar
N + kp:	ŋm-kpa	'be killing'	labio-velar
N + gb:	ŋm-gbona	'be getting hot'	labio-velar
b. sonorants:			
N + m:	mmu	'be drinking'	bilabial
N + n:	nna	'be beating'	alveolar
N + l:	nlɔ ~ ŋlɔ	'be going'	alveolar or velar
N + r:	nra ~ ŋra	'be buying'	alveolar or velar
N + j:	ŋjɔ	'be coming out'	velar
N + w:	ŋwa	'be coming'	velar
N + h:	ŋhɔ	'be scratching'	velar

([ŋm, kp, gb] are doubly articulated stops). Since the prefix's place of articulation is always predictable from the following consonant (the prefix does not occur before vowels), it should not be specified in its lexical representation. Its place of articulation is either assigned from the following consonant, or realized with a default velar constriction. In all cases, the nasal is realized with full closure.

A gestural interpretation of facts such as these must allow the constriction location of the following consonant to overlap the nasal gesture, but in the current model this cannot be done without simultaneously overlapping its constriction degree. In previous discussion of a similar case, Browman and Goldstein suggest that if the assimilated nasal is indeed realized with complete oral closure, the following consonant's oral gesture must change its constriction degree when it overlaps the velum lowering gesture (Browman and Goldstein 1989, 242). However, this proposal only displaces the problem, since if the overlapping gesture is produced with complete closure, the following consonant (which shares the gesture) should also be realized with complete closure. Similar cases of assimilated stop + continuant sequences can be cited from other languages, posing a genuine problem for Browman and Goldstein's characterization of gestures.

One may always question whether such descriptions are based on accurate observations. I know of no quantitative or instrumental study of nasal closure in nasal + consonant sequences in Yoruba. However, at least one such description of nasal clusters has been supported by instrumental analysis. Shona, as described by Doke (1931), has the homorganic nasal sequences [mb mv nd nz ȵʐ ɲdʒ ŋg], where the symbol [ȵ] designates a heavily rounded alveolar nasal appearing only before the voiced alveolar-labialized fricative [ʐ]. As far as the labial sequences are concerned, Doke states (p. 54):

> The bilabial nasal in Shona is formed just as in English, by complete contact of the lips, the air passing through the nose . . . Apart from the use of **m** immediately before vowels, it appears in the compounds **mb** (homorganic), **mv** (semi-homorganic), . . . Shona does not employ the denti-labial nasal **ɱ**, as do Zulu, Lamba, Bemba, etc., homorganically before **f** or **v**, but the full bilabial nasal in the combination **mv**.

Although the sequence [mv] is only "semi-homorganic" in Shona, we probably want to describe it as fully homorganic at the phonological level just as we do the other nasal clusters, and account for the fact that labial stops are predictably bilabial by adjusting the constriction location in the articulatory realization (without more information, of course, we cannot propose a definitive analysis). As far as the alveolar sequences [nz ȵʐ] are concerned, Doke does not explicitly state whether the nasals are produced with complete closure. However, he reproduces palatograms for these sequences from several Shona dialects, all of which show complete contact between the tip or blade of the tongue and the alveolar ridge. This closure must be attributed to the nasal, since (as other palatograms show) the fricative alone leaves an unobstructed passage through the center of the vocal tract. In Shona, then, nasals unambiguously assimilate constriction location but not constriction degree from a following continuant.[3]

There is also evidence that segments can assimilate constriction degree, but not constriction location. In Browman and Goldstein's model, vowel height is modelled in terms of constriction degree, which has the values [narrow], [mid], and [wide] (Browman and Goldstein 1989, 225-6), while place of articulation depends on the constriction location (palatal, velar, uvular, etc.) of the dorsal articulator. We would thus expect vowel height and place of articulation to assimilate as a single unit, never separately. In a number of languages, however, vowel height assimilates separately from place of articulation, and vice-versa (see Odden 1989 for a review of cases). The following examples are from Kimatuumbi, in which noninitial vowels in the stem assimilate to the height, but not the place of articulation of the initial vowel, provided both are nonlow. (Underlying noninitial vowels are represented below with upper-case letters, indicating their archisegmental status.)

underlying:	surface:	example (stem):	
i + I	i + i	-yipilya	'thatch with for'
i + U	i + u	-libulwa	'be ground'
u + I	u + i	-utika	'be pullable'
ɪ + U	ɪ + ʊ	-tɪkʊlya	'break with'
ʊ + I	ʊ + ɪ	-ʊʊgɪlwa	'be bathed'
ʊ + U	ʊ + ʊ	-kʊmbʊlya	'beat with'
ɛ + I	ɛ + ɛ	-chɛɛngɛya	'make build'
ɔ + I	ɔ + ɛ	-bɔɔlɛlwa	'be de-barked'
ɔ + U	ɔ + ɔ	-bɔmɔlwa	'be destroyed'

Once again, these facts are problematical for Browman and Goldstein's account of the gesture.

A solution to these problems can be found if place of articulation and constriction degree are allowed to occupy separate tiers of their own, where they can spread to other points in the representation independently of each other. This conception "unpacks" the notion of gesture without undermining it, since gestures can still be defined in terms of the lowest node superordinate to place of articulation and constriction degree. Suppose, to be specific, that we define consonantal and vocalic gestures in terms of the following configurations:

```
        cons                     voc
       /    \                   / \
  stricture  \         stricture   \
             place                 place
```

The "cons(onantal)" and "voc(oidal)" nodes in these figures designate gestures, and features are assigned to the stricture and place nodes to characterize them in terms of constriction degree and place of articulation. "Cons" and "voc" are not abstract labels, but define the range of stricture values appropriate to consonants and vocoids, respectively, and thus function similarly to the feature [±consonantal]. They link to the higher-level root node, to which their values may percolate. It is assumed that all nodes in this figure are assigned to different tiers in a representation (or score), effectively segregating consonants and vowels.

The "place" node dominates the set of oral tract articulator features labial, coronal, dorsal, and perhaps radical, or their gestural counterparts (L, TT, TB, etc.). These features are further specified for constriction location and shape by features such as [anterior] and [distributed]. The "stricture" node dominates stricture features, or their gestural counterparts: [±continuant] in the case of consonants, and vowel height features in the case of vocoids. Vocoids might also be redundantly assigned [+continuant], if the evidence warrants it. Clements (1990) presents evidence that vowel height features may spread independently of each other, and must thus be arrayed on further independent tiers.

These representations have the properties we need to characterize gestures in a way compatible with Browman and Goldstein's general conception, while allowing stricture and place of articulation to spread or overlap independently of each other. However, they appear to require modification of Browman and Goldstein's framework in two respects. First, in their current presentation Browman and Goldstein do not recognize a separate tier for "place of articulation," as is required by the representations above. However, there is much evidence that phonological rules may target the full set of oral place (articulator) features as a whole, rather than targeting only individual features. Thus in the Yoruba examples above, the rule assigning place of articulation to the nasal prefix applies to all oral places of articulation without exception, and spreads both components of the doubly articulated stops [kp, gb]. This is expressed in feature frameworks by spreading the place node leftward, but cannot be directly expressed in a gesture-based framework which does not provide a place node.

Second, in their current presentation Browman and Goldstein do not explicitly allow units on different tiers to be linked by association lines or other devices for indicating their membership in higher-level units.[4] Indeed, they appear to take a skeptical view toward the recognition of higher-level groupings of gestures, or coordinative structures, stating that "the only hierarchical unit for which we currently have evidence is that of the oral gestures in a (syllable-initial) consonant cluster." But as we have just seen, there is reason to group the components of the doubly-articulated stops [kp, gb] of Yoruba into a hierarchical structure (the place node), since both components spread or overlap as a unit. And once we analyze gestures into separate place and constriction degree components, it is necessary to indicate the connection between these components in some fashion. In the feature-based representations proposed above, this is done by connecting them with association lines, although other devices can be imagined.

The use of such associations or connections generalizes to much of the other data discussed by Browman and Goldstein. For example, they analyze the casual speech deletion of schwa in words like <u>beret</u> in terms of the overlapping of the two gestural components of [r] (tongue tip constriction and tongue root constriction) with the labial gesture. But the reason the two gestures of [r] overlap or spread to the labial as a single unit is most likely that they characterize a single phonological segment, represented in current phonological models by association to the root node. If the two gestures are regarded as merely accidental constellations with no intrinsic connection between them, the fact that they behave as a single temporal unit goes unexplained. Browman and Goldstein have presented many examples of sets of gestures that function as temporal units, but in none of these cases do the coordinated gestures <u>fail</u> to constitute a segment under most types of phonological analysis. Thus it would appear that there is motivation for grouping gestures into higher-level hierarchical units such as the segment, even for the treatment of casual speech phenomena.

Similar evidence can be given for recognizing further hierarchical units, such as the mora, the syllable, and and the phonological word. In short, if articulatory phonology is to be a viable candidate for a general theory of phonology and phonetics, it must provide an explicit way of expressing hierarchical relations among phonological units. [5]

5. Discussion

We have made two general points. First, to the extent that gestures and features differ in their empirical properties, features seem more appropriate to express the discrete nature of lexical contrast: at this level, what we require is not a photograph of the vocal tract, but a rough map. Second, if gesture-based models are to be adequate to the expression of many phonological (and perhaps phonetic) generalizations, they require additional hierarchical organization of the sort postulated in current feature-based models of representation.

What, then, is the relation between gestural phonology and feature-based phonology? Are the two completely incompatible? I have suggested above that if the notion "gesture" is suitably revised along the lines suggested above, gesture-based phonologies will differ little if at all from feature-based phonologies at the more abstract and systematic levels of phonological representation at which units behave in a discrete fashion. The apparent incompatibility between the two modes of representation can perhaps be resolved in terms of Browman and Goldstein's further observation that "increase in overlap among gestures in fluent speech is a general gradient process that can produce apparent (perceived) discrete alternations" (my italics). What this remark suggests is that speech is produced in a gradient fashion, but perceived (and thus represented) categorically. If this view is correct (and there is much evidence that it is), then the status of gesture-based models in a global theory of phonology and phonetics may be clarified. Gesture-based models (and others with similar goals) are models of speech production, which address the complex and important problem of "how to bridge the gap between the discrete segments of the phonological system and the fluid change in time and space that is the final result of phonetic processes," to quote Huffman's apt characterization (1989, 139).

In sum, Browman and Goldstein have developed an elegant and comprehensive theory of articulatory structure and its relation to phonological structure which can serve as a basis for proposing and testing specific hypotheses about how abstract phonological representations are related to physical phonetic output. Indeed, it has already proven its value in this respect. On the other hand, it has not successfully handled certain aspects of phonological patterning in which current feature theories seem to offer a closer fit to linguistic reality. To the extent that the future development of their approach is able to address problems of this sort successfully, it will take a step toward relieving phonological feature theory of the task of accounting for phenomena that lie outside its proper domain, and help lay the basis for a better definition of the way in which abstract phonological structure is transmitted through the medium of speech.

Notes

[1] The status of duration in Browman and Goldstein's current model is not altogether clear. On the one hand, they state that "quantitative temporal information is provided not by specifying time directly but by specifying the parameters of the gestural regimes and their phasing" (Browman and Goldstein 1990, 310). On the other, the present paper states that a gestural score "displys the duration of the individual gestures as well as the overlap among the gestures." If the gestural score constitutes the input to the task dynamic model which contributes to determining temporal information, as Browman and Goldstein propose, temporal information should not be available to the score. For purposes of the present discussion I take their more recent statement as the definitive one.

[2] See Steriade (1990) for further discussion.

[3] An exceptional case was noted for a speaker of the Ndau dialect, for whom "when **nz** was initial, the **n** did not effect complete contact, there being a space showing that in reality it was nasalized **z** (z̃) which was produced. Complete contact was effected when not initial, as in **hanzu**" (Doke 1931, 265). However, the same speaker realized /tʃ/ as [ʃ] word-initially, and incomplete contact was found before the (similarly fricativized) realization of word-initial /ɲdʒ/, suggesting that this speaker used incomplete closure quite generally in initial position.

[4] They have elsewhere suggested, however, that association lines might be used to express phasing relations between overlapping <u>successive</u> segments (Browman and Goldstein 1989).

[5] Browman and Goldstein have elsewhere (1989) tentatively proposed that autosegmental and prosodic structure can be interfaced with gestural scores on distinct planes, but this suggestion is not taken up in the present paper.

References

Browman, C.P. and L. Goldstein (1989) "Articulatory Gestures as Phonological Units," Phonology 6.2, 201-51.

Browman, C.P. and L. Goldstein (1990) "Gestural Specification using Dynamically-defined Articulatory Structures, Journal of Phonetics 18.3, 299-320.

Browman, C.P. and L. Goldstein (this issue) "Articulatory Phonology: an Overview," Phonetica.

Clements, G.N. (1985) "The Geometry of Phonological Features," Phonology Yearbook 2, 225-252.

Clements, G.N. (1990) "The Hierarchical Representation of Vowel Height," unpublished ms. Ithaca, N.Y.: Cornell University.

Clements, G.N. (in press) "Place of Articulation in Consonants and Vowels: a Unified Theory," to appear in B. Laks and A. Rialland, eds., L'Architecture des Représentations Phonologiques, Paris: Editions du C.N.R.S.

Davis, K. (1991) The Acquisition of the Voicing Contrast by Speakers of English and Hindi, unpublished Ph.D. dissertation. Ithaca, N.Y.: Cornell University.

Doke, C.M. (1931) A Comparative Study in Shona Phonetics. Johannesburg: University of the Witwatersrand Press.

Hertz, S.R. (1991) "Streams, Phones, and Transitions: Toward a New Phonological and Phonetic Model of Formant Timing," Journal of Phonetics 19.

Huffman, M. (1989) Implementation of Nasal: Timing and Articulatory Landmarks, Ph.D. dissertation (also UCLA Working Papers in Phonetics 75), Los Angeles: Dept. of Linguistics, UCLA.

Keating, P.A. (1990) "Phonetic Representations in a Generative Grammar," Journal of Phonetics 18.3, 321-34.

McCarthy, J. (1988) "Feature Geometry and Dependency: a Review," Phonetica 45,84-108.

Odden, D. (1989) "Vowel Geometry," unpublished ms. Columbus, Ohio: the Ohio State University.

Sagey, E. (1986) The Representation of Features and Relations in Nonlinear Phonology, Ph.D. dissertation. Cambridge, Ma.: MIT.

Steriade, D. (1990) "Gestures and Autosegments: Comments on Browman and Goldstein's 'Gestures in Articulatory Phonology'," in J. Kingston and M.E. Beckman, eds., Papers in Laboratory Phonology I: Between the Grammar and the Physics of Speech, pp. 382-97. Cambridge: Cambridge University Press.

Stevens, K.N. (1991) "Phonetic Evidence for Hierarchies of Features," paper presented at the Third Conference on Laboratory Phonology, Los Angeles: U.C.L.A.

Ward, Ida C. (1952) An Introduction to the Yoruba Language. Cambridge: Cambridge University Press