

Linguistic Society of America

Features, Gestures, and Igbo Vowels: An Approach to the Phonology-Phonetics Interface

Author(s): Elizabeth C. Zsiga

Source: *Language*, Vol. 73, No. 2 (Jun., 1997), pp. 227-274

Published by: Linguistic Society of America

Stable URL: <http://www.jstor.org/stable/416019>

Accessed: 17/06/2010 11:19

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at <http://www.jstor.org/page/info/about/policies/terms.jsp>. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Please contact the publisher regarding any further use of this work. Publisher contact information may be obtained at <http://www.jstor.org/action/showPublisher?publisherCode=lsa>.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.



Linguistic Society of America is collaborating with JSTOR to digitize, preserve and extend access to *Language*.

<http://www.jstor.org>

FEATURES, GESTURES, AND IGBO VOWELS: AN APPROACH TO THE PHONOLOGY-PHONETICS INTERFACE

ELIZABETH C. ZSIGA

Georgetown University

This article examines two processes that affect vowels in Igbo: harmony and assimilation. Through these two processes, the relationship between autosegmental features and articulatory gestures is explored. Vowel harmony is argued to be featurally represented, but acoustic evidence shows that vowel assimilation is gradient and best represented in terms of articulatory gestures. Neither representation is adequate in itself to describe the full range of phonological and phonetic data; rather, I advocate a mapping procedure that takes advantage of the resemblances between autosegmental and gestural representations without collapsing the two. A complete account of Igbo vowel harmony and assimilation is provided, demonstrating the need for two kinds of representation, and illustrating the suggested feature-to-gesture mapping.*

He only says 'good fences make good neighbors.'

[But] why do they make good neighbors? Isn't it
Where there are cows? . . .

Before I built a wall I'd ask to know

What I was walling in or walling out.

(Robert Frost, *Mending Wall*)

1. FEATURES AND GESTURES. Researchers in phonology and phonetics disagree over how different phonological and phonetic representations should be from one another and whether describing the relationship between those representations is easy or hard, straightforward or complex. On the one hand, articulator-based features arranged to mirror the structure of the vocal tract seem to make the correspondence between phonetic and phonological representations transparent. Many researchers have pointed out the similarities of models of feature geometry to the organization of the vocal tract (e.g. Clements 1985, Sagey 1986, McCarthy 1988, Browman & Goldstein 1989a, Padgett 1991, Keyser & Stevens 1994). Even theories proposing geometries that do not conform closely to the vocal tract, such as Clements 1991 and Hume 1992, or that assume no feature geometry at all (Cole & Kisseberth 1994, Padgett 1995), maintain features based on active articulators such as [labial], [coronal], and [dorsal].

Yet even assuming the most transparent, articulator-based inventory of phonological features, many researchers investigating the phonology/phonetics interface have argued that phonological and phonetic representations are fundamentally different. It is difficult to reconcile the kind of representation needed to capture the categorical alternations of the phonology with the kind of representation needed to capture the gradient, continuously varying pro-

* This research was supported by NIH grant HD-01994 to Haskins Laboratories. The author gratefully acknowledges the help and support of Louis Goldstein, Draga Zec, and John McCarthy at all stages of this work, as well as the helpful and insightful comments of two anonymous referees and the editors of *Language*.

cesses of speech. Pierrehumbert (1990), for example, argues that phonological representations must be qualitative and symbolic while phonetic representations must be quantitative and physical, and that an account of the mapping between them remains elusive. Keating (1990a) argues that despite the feature geometry's emphasis on articulation, phonological features must be mapped into both the articulatory and acoustic domains (as was originally envisioned by Jakobson et al. 1951), and that even within the articulatory domain, a given phonetic parameter may be influenced by more than one feature, and a single feature may influence many parameters. This point is supported and elaborated by Kingston and Diehl (1994), who argue that phonetic adjustments are not automatic but independently controlled and language-particular, and by Keyser and Stevens (1994), who discuss the recruitment of articulators that are not phonologically specified as active in order to acoustically enhance the specified phonological features.

In contrast, Browman and Goldstein, in the theory of ARTICULATORY PHONOLOGY (1986, 1989a,b, 1992), have made maximal use of the correspondence between the structures of the vocal tract and the representations needed to describe phonological patterning. They propose a single unit, the ARTICULATORY GESTURE, as the basic unit both of phonological representation and of its physical actualization. Each gesture involves a group of articulators that act together to form and release a constriction in the vocal tract, such as labial closure, glottal opening, or velar frication. The gestural approach relies on the proposal that articulators may act in concert to account for the finding that more than one physiological parameter may be active in the realization of a given phonological contrast.¹ According to Browman and Goldstein, not only are articulatory gestures 'characterizations of discrete, physically real events that unfold during the speech production process' (1992:156), they also suffice for the description of phonological contrast. As autosegmental representations may contrast on the basis of the presence or absence of a feature or segment, gestural representations may contrast on the basis of the presence or absence of a gesture (*had* vs. *add*), in the articulator set used (*bad* vs. *dad*), or in the parameters of constriction location or degree (*shad* vs. *sad* vs. *tad*). The organization of gestures in time can also produce phonological contrasts: *bomb* and *mob* have the same gestures, but they are organized differently. Thus, in articulatory phonology, there is no interface between phonology and phonetics, because phonological and phonetic representations are the same.

Following Browman and Goldstein, I adopt the view that the units of phonological and phonetic representation are indeed basically the same. I depart from the articulatory phonology approach, however, and maintain the traditional view that the two kinds of units cannot be exactly alike. I adopt the articulatory gesture as the basic unit of phonetic representation and then use the correspon-

¹ For further information on the details of Articulatory Phonology, the reader is referred to Browman and Goldstein (1992) and references therein. References on the functional grouping of articulators include Kelso et al. 1984, Shaiman 1989, Gracco & Abbs 1988, and Kollia et al. 1992. References on gestural dynamics include Saltzman 1986, Saltzman & Kelso 1987, Saltzman & Munhall 1989, and Hawkins 1992.

dences between features and gestures to argue for a mapping between the two types of representation. Igbo vowel harmony will serve as an example of the kind of categorical alternation that a featural representation is designed for, but that a gestural representation does not handle well. Igbo vowel assimilation, however, will be shown to be a gradient and variable process that a featural representation cannot capture, but that is well described in a gestural approach.

Although featural and gestural representations are both based on articulators, I will argue that the crucial difference between them is that gestures have inherent quantitative specifications, most importantly, exact temporal relations. A gestural representation may be seen as a featural representation fully specified for temporal relations, or conversely, an autosegmental representation may be seen as a gestural representation for which specific timing information is not available. The quantitative specifications, which are available to the phonetics but not the phonology, replace the categorical hierarchical nodes and association lines, which are available to the phonology but not to the phonetics.

The idea of filling in quantitative specifications in the course of a derivation is not new. Chomsky and Halle (1968) proposed that in the phonetics numbers should replace plus and minus featural specifications (so that an /o/ in English might be [3 back], [2 round], and [4 high]). Kiparsky (1985) and Mohanan (1986), in the theory of lexical phonology, proposed that postlexical phonological processes could access quantitative information, while lexical phonological processes could not. I agree with these and other researchers (including Keating 1988, 1990a, Cohn 1990, and Pierrehumbert 1990) in reaffirming that qualitative and quantitative representations are both needed, and that they should be kept separate. This article differs from previous work in proposing a simple and specific mapping procedure that incorporates the gestural structures of articulatory phonology.

Consider how such a mapping would work in the Igbo verb root [kpa], 'to weave by hand', which begins with a voiceless labio-velar stop. (For simplicity, laryngeal specifications, which may differ dialectally, are not shown. See Ladefoged et al. 1976 and Ladefoged & Maddieson 1996 for discussion of laryngeal configurations in Igbo double stops.) The feature-geometrical representation of this syllable, assuming for concreteness the feature geometry proposed by Padgett (1991), is as shown in Figure 1A.

In an autosegmental representation, only two temporal relations can be expressed: simultaneity and precedence. If two features are on one tier, either *a* precedes *b* or *b* precedes *a*. Features are organized into sets (segments) by direct or indirect linking to a root node. Features associated to a single root node but on different tiers are, for the purposes of the phonology, simultaneous, even if not executed simultaneously in actual speech (Sagey 1986, 1988).² For example, in the representation of [kpa], the consonant precedes the vowel,

² Sagey (1988) in fact proposes that phonology recognize that features have duration, and that association lines be seen as representing not simultaneity but overlap in time. The extent of the overlap, however, remains undetermined; whether all instances of the articulations or only one are shared can not be discovered by the phonology. 'The internal detail', Sagey writes, 'is not accessible to or manipulable by phonological processes' (p. 112).

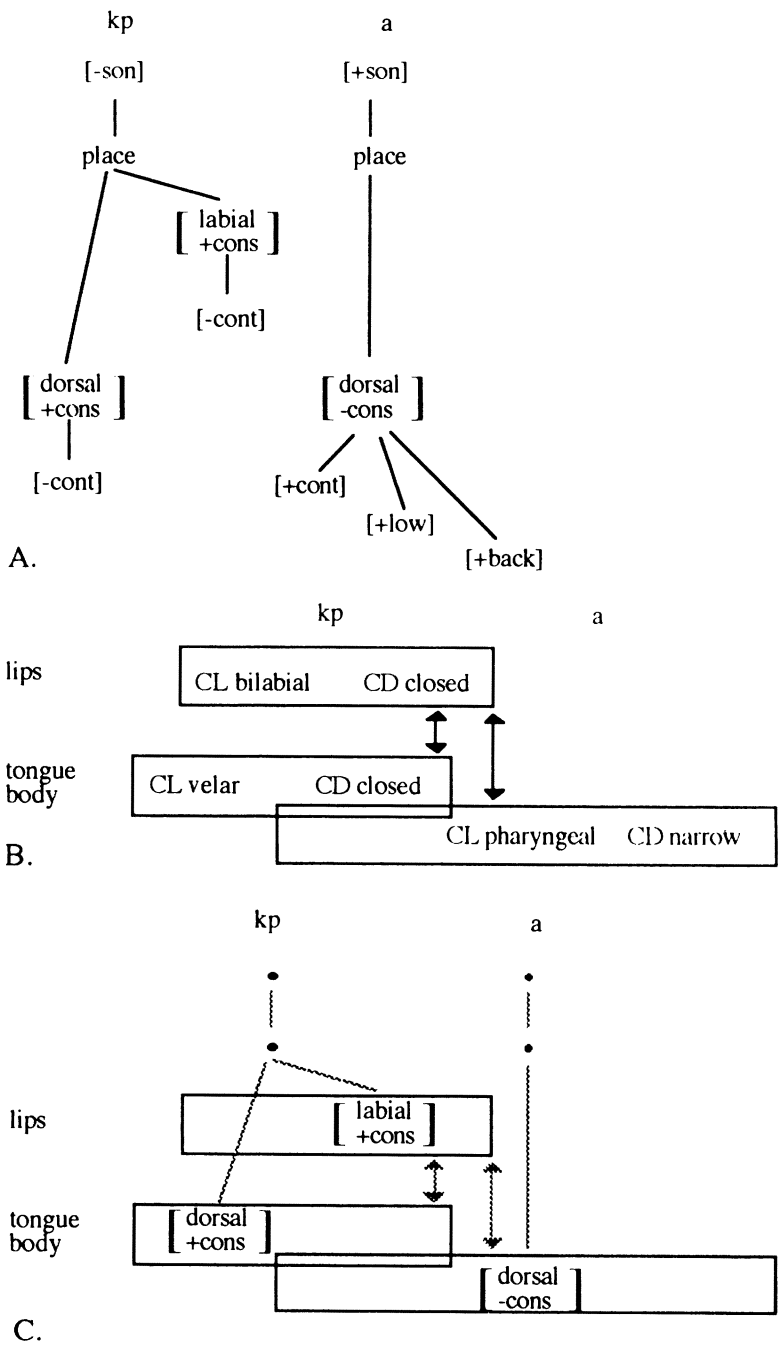


FIGURE 1. The syllable [kpa]. A. Feature-geometrical representation, following Padgett (1991). B. Gestural representation, following Browman and Goldstein (1989). C. Feature-to-gesture correspondence.

because the root node of the one precedes the root node of the other. The two parts of the labiovelar stop are not ordered with respect to each other, however, because the features [labial] and [dorsal] are on different tiers, linked (via a class node) to a single root node. The exact timing and duration of the articulations is not part of this representation.

This featural representation will be mapped into three gestures: labial and dorsal closing gestures for [kp], and a pharyngeal constriction gesture for [a]. The gestural representation for this syllable, following the conventions of a GESTURAL SCORE (Browman & Goldstein 1986), is shown in Figure 1B. In a gestural score, the horizontal dimension represents time, and the length of the boxes indicates the duration of the gesture: the period of time that the articulator set is actively controlled. Constriction location (CL) and degree (CD) are shown. Timing between gestures is specified by coordinating a specific point in one gesture (such as onset, achievement of target, or release of target) with respect to a specific point in some other gesture.

In [kpa], velar release is timed to occur during the labial closure. (Ladefoged & Maddieson 1996 argues that this is the crucial timing relation necessary to distinguish labio-velars from both clusters and single labial and velar stops.) The vowel gesture, in turn, is timed to reach its target position as the labial gesture is released. These specified temporal relations are indicated by arrows in Figure 1B. Note that no root or class nodes are used to organize gestures, so that all association is indicated in terms of phase relations rather than links to abstract hierarchical nodes (Browman & Goldstein 1990). There is no direct gestural correlate of the segment: [kp] differs from [k#p] only in having a specific temporal relation.

Temporal organization aside, features and gestures correspond closely. Comparison of 1A and 1B shows that [–cont] corresponds to a closed constriction degree, while the articulator features [labial] and [dorsal] determine the articulator sets (lips, tongue body) that implement the oral constrictions. Features dependent on the articulator features determine the exact constriction location (so that [low] and [back] map into CL pharyngeal). The feature [cons] is implemented as STIFFNESS, which determines a gesture's velocity. The clear correspondence can be illustrated by laying the two figures on top of one another, as in Figure 1C. Note how direct timing between gestures substitutes for the abstract hierarchical structure. It is in the specification of timing that gestural and featural representations crucially differ. Gestures have specific extent in time, while features do not.

Evidence from vowel harmony and vowel assimilation in Igbo will illustrate the need for both kinds of representations, and will further explicate this proposed mapping. While a gestural approach works well for describing many kinds of phonetic and connected speech processes, I argue that it is inadequate for the expression of categorical alternations. A featural representation, however, is inadequate for the expression of the gradient changes that a gestural approach handles well. No single representation is adequate in itself to describe the full range of phonological and phonetic data.

2. VOWEL HARMONY.

2.1. EVIDENCE FOR FEATURES. Igbo has eight vowels, which can be described as contrasting in the features [high], [round], and [ATR] (advanced tongue root), as shown in 1. Anticipating the discussion below, the [−ATR] vowels are marked with the IPA symbol for retracted tongue root, approximating the Igbo orthographic convention of marking these vowels with an under-dot. To avoid a confusion of symbols, [+ATR] vowels are left unmarked. Tones are also left unmarked to further limit notational complexity. (Although tonal alternations are often seen in the same environment as vowel assimilation, tone will not be a focus here. See Clark 1990 and Liberman et al. 1993 for further discussion of the tonal system of Igbo.)

(1) Igbo vowels

	i	ĩ	e	ǣ	u	ũ	o	õ
high	+	+	−	−	+	+	−	−
round					+	+	+	+
ATR	+	−	+	−	+	−	+	−

Note that the pairs [i] and [ĩ], [u] and [ũ], [o] and [õ], and [e] and [ǣ] are described as featurally alike except for their [ATR] value. Evidence for these featural pairings comes from phonetic studies and from the phonological alternations of vowel harmony.

Ladefoged 1968 provides x-ray tracings of vocal tract configurations for each of the Igbo vowels. For the vowels [i, ĩ, u, ũ, o, õ], the [+/−ATR] contrast is realized simply by advancing or retracting the tongue root: the lips, the jaw, and the highest point of the tongue body are seen to be in the same position for each member of the pair. This is not the case for [e] and [ǣ], however. In addition to retraction of the tongue root for [ǣ] as opposed to [e], there is also a clear difference in the position of the jaw and tongue body: [ǣ] is both lower and further back than [e]. Nonetheless, it is phonologically necessary to describe these two vowels as differing only in the feature [ATR] because [e] and [ǣ] alternate in morphemes subject to harmony, exactly parallel to the other pairs of vowels.

In Igbo, all vowels within a noncompound word are drawn from either the [+ATR] or [−ATR] set, as illustrated in Table 1. Blank cells indicate nonoccurring combinations.³ Generally, [+ATR] and [−ATR] vowels can never co-occur within morphemes. As Table 1 shows, there are no other restrictions on vowel combinations.

The only disharmonic morphemes are a few nouns beginning with [ǣ]. These nouns are apparently remnants of an older nine-vowel system, similar to neighboring Akan and still extant in some western dialects of Igbo, which had an additional midfront vowel as the [−ATR] counterpart of [e], and in which the

³ Unless otherwise noted, all the Igbo words and phrases used as examples in §§2 and 3 were provided by Esther O., a native speaker of a Central dialect of Igbo, residing in New Haven, CT. at the time the data was collected. Esther provided original data, and confirmed examples taken from other sources. The dis-harmonic nouns in Table 1 are from Emenanjo (1978).

	i	u	e	o	i	u	a	o
i	isi 'head'	ihu 'face'	ihe 'thing'	igbo 'Igbo'				
u	ubi 'farm'	ukwu 'big'	uwe 'clothing'	ugo 'eagle'				
e	ezi 'loan'	esu 'millipede'	eze 'teeth'	ego 'money'				
o	obi 'heart'	ozu 'corpse'	ose 'stream'	obodo 'town'				
i					isi 'to tell'	izu 'to buy'	nkita 'dog'	icho 'to want'
u					udi 'kind, sort'	umu 'children'	uwa 'world'	ulo 'house'
a	akpi 'scorpion'	adu 'kola nut'	akpe 'soap box'	alo 'thought'	anyi 'we, us'	mmadu 'people'	ada 'daughter'	ato 'three'
o					oti 'grub'	odu 'tusk'	ota 'shield'	ozo 'another'

TABLE 1. Vowel harmony: Possible vowel combinations in noncompound words

low vowel was transparent to harmony (see Welmers 1973). In all other dialects of modern Igbo, [a] has become the [– ATR] counterpart of [e].

All inflectional affixes alternate so as to agree with the [ATR] value of the stem, as shown in 2. A few aspectual suffixes also harmonize, such as the 'directional' suffix *-ta/-te* shown in 3 (examples from Emenanjo 1978).

(2) [ATR] harmony to inflectional affixes⁴

- a. -a/e si-a si-e
IMP tell! cook!
- b. -ghi/-ghi si-ghi si-ghi
NEG.INDIC did not tell did not cook
- c. i-/i- i-si i-si
INF to tell to cook
- d. o-/o- o-si o-si
AGN the teller the cook
- e. a-/e- a-si e-si
PART telling cooking
- f. o-/o-, -Vla/-Vle o-si-ala o-si-ele
3SG.SUBJ.PERF s/he has told s/he has cooked
- g. a-/e-, -ri/-ri a-si-ri e-si-ri
3SG.INDEF.SUBJ.INDIC someone said someone cooked

⁴ The following abbreviations are used: agentive, AGN; applicative, APPL; associative, ASSOC; complementizer, COMP; definite, DEF; directional, DIREC; imperative, IMP; indefinite, INDEF; indicative, INDIC; infinitive, INF; inflectional prefix (prefix on inflected verbs with overt subjects), INFP; negative, NEG; object OBJ; participial, PART; past, PST; perfective, PERF; plural, PL; possessive, POSS; progressive, PRG; singular, SG; subject, SUBJ. Igbo pronouns have no gender: third person singular pronouns may be translated here as 'she', 'he', 's/he', or 'it' depending on the context from which the example is taken.

(3) One harmonizing aspectual suffix⁵

- a. i-zu-ta
 INF-buy-DIREC
 'to buy for'
- b. i-zu-te
 INF-meet-DIREC
 'to meet with'

Harmony does not apply, however, to most aspectual suffixes, or between the members of a compound, as shown in 4. Thus the [+ATR] verb *bi* 'live' may combine with the [-ATR] aspectual suffixes *kɔ* 'associative' and *rɪ* 'applicative'. The [-ATR] verb *gha* 'turn' can combine with the [+ATR] verb *gbu* 'hurt' to create a compound verb meaning 'cheat'. Inflectional affixes harmonize with the adjacent nonharmonizing morpheme.

(4) Nonharmonizing aspectual suffixes and compounds

- a. i-bi-kɔ-rɪ-ta
 INF-live-ASSOC-APPL-DIREC
 'to live together to one another's advantage'
- b. ibe a-gha-gbu-go m
 Ibe INFP-turn-hurt-PST 1SG.OBJ
 'Ibe cheated me.'

In order to account for these patterns of harmonizing and nonharmonizing morphemes, I have argued elsewhere (Zsiga 1992) that the domain of Igbo vowel harmony is best described as the phonological (rather than morphological) word. That article makes the case that nonharmonizing verb roots and aspectual suffixes form independent phonological word domains, while the harmonizing inflectional suffixes are incorporated into the domain of the adjacent morpheme. Assuming that alternating affixes are unspecified for an [ATR] value, harmony can then be described as a rule that spreads [ATR] within the phonological word (Zsiga 1992), or alternatively, as a constraint prohibiting conflicting [ATR] specifications within the same domain (Cole & Kisseberth 1994), or aligning [ATR] values with domain boundaries (Archangeli & Pulleyblank 1994). In each case, the result is a set of categorical associations between vowels and [ATR] specifications, as shown in 5.

(5) Link ATR



While this featural representation expresses the facts well, a gestural representation would not.

2.2. EVIDENCE AGAINST GESTURES. Evidence that Igbo vowel harmony should be represented with features rather than gestures comes from several

⁵ Examples of other harmonizing aspectual suffixes are found in Emenanjo 1978, Clark 1990, and Zsiga 1992.

areas. First, the harmonic [ATR] alternations are categorical, the substitution of one vowel for another. There is no evidence that a derived [a], as in the perfective suffix in *o-si-ala* (2f above), is phonetically different from an underlying [a], as in *ala* 'land'. In fact, there is debate about whether the initial vowels in nouns should be considered as underlying or as prefixes (see Clark 1990). This debate could not occur if underlying and derived vowels could be distinguished phonetically. A gestural approach (or more generally any approach that relies on exact phonetic specifications) cannot account well for alternations that are always categorical.

Certainly, some assimilations, deletions, and insertions that have been described in terms of phonological rules are better captured by processes that manipulate gestures (see Browman & Goldstein 1986, 1989b, 1990, 1991, 1992 and references therein). In this approach, variation is expressed only in terms of reorganization in the temporal patterning of gestures: gestures may overlap one another to a greater or lesser extent, and gestural magnitude may increase or decrease. 'Gestures are never changed into other gestures, nor are gestures added' (Browman & Goldstein 1992:173). For example, the assimilation or deletion in English of word-final alveolars before a following stop (Gimson 1962) can be described in terms of increased overlap between the tongue tip gesture and a following velar or labial closure gesture, effectively hiding the acoustic consequences of the coronal closure (Browman & Goldstein 1990, Zsiga & Byrd 1990, Byrd 1992, Nolan 1992). An /n/ may be dental, not alveolar, in the pronunciation of *ten things*, as two different gestures that call on the same articulator set overlap, causing the articulators to reach a position in between the two conflicting targets (Browman & Goldstein 1986). In a phrase like *press your point*, the pitch of the fricative sound may change from /s/-like at the beginning to /ʃ/-like at the end, because of blending of the alveolar gesture for the /s/ and the palatal gesture for the /y/ (Zsiga 1995). Hiding and blending follow without stipulation from the dynamics of combining abstract gestures. As Browman and Goldstein argue, if all phonological processes can be described in this way, then articulatory gestures can serve both as 'phonological primes' and as the basic units of articulation, and gestural representations will obviate the need for any other phonological representation.

A number of researchers, however (e.g., Steriade 1990, Clements 1992, Kingston & Cohen 1992, Zsiga 1993, 1995, and Nolan et al. 1996), have argued that articulatory gestures are not adequate phonological representations because much more detail is needed to accurately model the movements of articulators than is needed to express the set of possible categorical contrasts and alternations. Given the power of the theory to describe the details of timing, articulatory phonology does not offer a convincing account of how temporal reorganization in categorical alternations can be constrained. Cohn (1990) and Krakow (1989), for example, have shown that the particulars of timing between nasal and oral gestures in nasalized consonants and vowels will differ from language to language and from position to position. Browman and Goldstein (1992) point out that a gestural account can correctly model the different degrees of vowel nasalization found in English when an adjacent nasal consonant occurs.

pies different syllable positions, as a result of distinct timing relations. But no language uses these specific subsegmental timing relations to express a lexical contrast. No language has two lexical rules of vowel nasalization, one creating fully nasalized, the other partially nasalized vowels, or creating two degrees of partial nasalization. Similarly, while the gestural approach accounts well for the gradient and variable palatalization seen in /s + y/ or /s + ʃ/ sequences across word boundaries in English, it fails to account for why lexical palatalizations, as in the alternations between *press* and *pressure* or *confess* and *confession*, always apply categorically.⁶ Any one representation that is powerful enough to describe gradient processes will not be constrained enough to explain the categorical nature of alternations such as Igbo vowel harmony.

Not only is Igbo vowel harmony categorical, it is also clearly a lexical rule according to the definition proposed by Kiparsky (1985). Harmony is sensitive to morphological structure, such as the difference between inflectional and aspectual affixes, and has lexical exceptions, failing to apply to the disharmonic nouns in Table 1, and applying unexpectedly to some aspectual suffixes as in 3. Even proponents of the strongest versions of articulatory phonology do not suggest that such lexical alternations should be accounted for in terms of changes in gestural magnitude or phasing.⁷

Rather, Browman and Goldstein suggest that words that are related by lexical alternations may be stored as distinct gestural structures in the lexicon, chosen by the speaker depending on the context (Browman & Goldstein 1991:324). For example, the prefix *im-* in *impolitic* and *imperfect* could be stored in the lexicon as an allomorph of *in-*, unlike the variable, cross-word-boundary assimilations in phrases like *my cousin is i[m] politics* or *they skated i[m] perfect step* that may be described in terms of gestural overlap. However, to the extent that phonological alternants must be listed in the lexicon, generalizations about the processes that relate them are not expressed. No distinction is made between the large body of constrained phonological alternations that have clear articulatory bases (and can be represented simply by articulator features associated to hierarchical nodes like the root node) and truly opaque morphological alternations.

Finally, Igbo vowel harmony must be represented with features rather than with gestures because it is not phonetically transparent. The [e/a] alternation shows that more is involved in vowel harmony than simply adding a tongue root gesture. Retracting the tongue root does not automatically change a front vowel into a back vowel or a mid vowel into a low vowel. (This is not the result in nine-vowel systems such as Akan and Western Igbo, for example.) Rather, changing (or adding) an [ATR] specification creates a new, categorically con-

⁶ Nolan et al. (1996) have argued that even some post-lexical cases of /s/ to /ʃ/ assimilation seem to apply categorically; although, as the authors freely admit, the evidence for categorical assimilation is rather weak. Several apparently complete cases of Igbo vowel assimilation are discussed in §3.

⁷ McMahon et al. (1994) propose a weakened version of articulatory phonology that can account for lexical alternations. Such an account of Igbo vowel harmony, however, would still fall short for the other reasons mentioned here.

trastive, combination of features, which is subject to a specific phonetic (gestural) interpretation. If the contrastive possibilities of modern Igbo are expressed in terms of the features [high], [round], and [ATR], then vowel harmony, including the [e/a] alternation, can be expressed simply and generally as an association between an [ATR] feature and all the vowels within a phonological word. A featural representation best accounts for the facts that harmony is categorical, lexical, and phonetically opaque. The interpretation of the [–high, –round, +ATR] vowel as [e] and the [–high, –round, –ATR] vowel as [a] can be safely left to the phonetics.

Vowel assimilation, however, is a very different story.

3. VOWEL ASSIMILATION.

3.1. GRADIENT ASSIMILATION: DESCRIPTIVE EVIDENCE. Vowel harmony applies within the phonological word; across phonological word boundaries, vowel assimilation takes place. Most Igbo words begin and end in vowels, so vowels are continually adjacent. While in careful speech both vowels in a sequence are pronounced, in fluent speech the first vowel takes on the quality of the second. The duration and tone of both vowels are retained.

Assimilation is not normally indicated in the orthography, but in phonological descriptions the output is represented by a doubled vowel, as in 6.

(6) V1V2 → V2V2

nwoke a → nwoka a

man DEF

'this man'

Although this transcription indicates that assimilation is complete and categorical, some accounts in the literature describe a process that is gradient and variable.

Several authors have noted that the high vowels do not assimilate as freely as the nonhigh vowels. Green and Igwe (1963) note that assimilation of a low vowel to a high vowel is more likely than the reverse. Clark (1990) states that [i] and [ɪ] never assimilate, while [u] and [ʊ] may not assimilate in careful speech. Emenanjo (1978:25) states that 'The non-close vowels assimilate completely and automatically while the close vowels either never assimilate completely or do so only conditionally.' According to Emenanjo, /u/ and /ʊ/ assimilate to a following vowel only in very rapid speech and the high front vowels do not assimilate at all, but become glides if both V1 and V2 are on the same tone level and share the same [ATR] value. Emenanjo notes, however, that even in the case of gliding of /i/ and /ɪ/, the duration is the same as that of a vowel sequence.

Welmers (1973:41–42), in his description of Igbo vowel assimilation, states that a remnant of V1 is almost always perceptible.

The only vowel which is completely assimilated to a following vowel in quality is /a/, and even this complete assimilation is heard primarily in very common expressions or in rapid speech ... In all other sequences, the quality of the first vowel merges very quickly into the quality of the second. Thus the second vowel is more prominent, but the first remains identifiable. There may be a slight centralization of a front before a back vowel, or of a back before a front

vowel, but all contrasts are maintained. This has been strikingly demonstrated by a number of native speakers of Igbo used as models in teaching Igbo to speakers of English. Hearing the slight centralization of /e/ and the rapid transition from /e/ to /o/ in a phrase like /ebe ole/ 'which place? where?', students sometimes go to the extreme of saying [ebo ole]—and similarly [eba à] for /ebe à/ 'this place, here'. A native speaker of Igbo may hesitate to reject such exaggerations out of hand, but the invariable reaction is a look of dissatisfaction and a repetition of the phrase in hopes of hearing a more accurate imitation. In short, such vowel sequences in Igbo may be difficult to learn to reproduce accurately, but they are by no means instances of full vowel assimilation.

According to Welmers, then, the description 'V1V2 becomes V2V2' is an unacceptable exaggeration of a process more accurately described as 'slight centralization' of V1 followed by a 'rapid transition' to V2. Welmers describes very clearly a gradient assimilation. Yet the result of the process is close enough to complete assimilation to confuse learners and to make native speakers pause before rejecting a completely assimilated pronunciation.

Some examples of assimilation in Igbo vowel sequences are given in 7–14. These phrases are taken from transcriptions of one Igbo speaker reading aloud a passage from the novel *Isi akwụ dara n'ala* (The palm frond that fell to the ground), by T. Ubesie, as well as some sentences constructed by the same speaker for a study of Igbo syntax. Nearly all possible two-vowel combinations are represented, although there are a few sequences (such as /o/ and /o/ followed by high vowels) which didn't happen to occur in the sample. Where the transcriber (the author) heard these sequences as more similar to a single long vowel than to a sequence of two different vowels, the output of assimilation is shown as a doubled vowel. This step is taken only because transcription systems are limited in the ability to show gradient differences. Phonetic evidence for gradient assimilation in sequences such as these is given in §3.2.

(7) assimilation of a

- a. ɔnwà isii → ɔnwi isii
months six
'six months'
- b. mǎ i-gǎ → mǐ iga
but INF-go
'but to go'
- c. ɔ nǎ e-re → ɔ ne ere
3SG. PRG INFP-sell
'she is selling'
- d. yǎ uwe → yu uwe
3SG.POSS clothing
'her clothing'
- e. ka ɯmɯ nwanyi → kɯ ɯmɯ nwanyi
COMP PL woman
'that women'
- f. nwǎ okorɔbiǎ → nwo okorɔbiǎ
man young
- g. ɔnwǎ ɔbula → ɔnwɔ ɔbula
month each

(8) assimilation of e

- a. *iche* *iche* → *ichi* *iche*
different different
‘various’
- b. *ihe* *abu* → *iha* *abu*
thing two
- c. *ṅke* *ukwu* → *ṅku* *ukwu*
item (of) large size
‘extensively’, ‘in a big way’
- d. *oge* *ufodu* → *ogu* *ufodu*
time some
‘some time’
- e. *ebe* *obi* → *ebo* *obi*
place (of the) heart
‘where the heart (is)’
- f. *onye* *obula* → *onyo* *obula*
one each

(9) assimilation of ɔ

- a. *i-cho-ro* *ego* → *iḥore* *ego*
2SG-want-INDIC money
‘you want money’
- b. *ulo* *akwukwo* → *ula* *akwukwo*
house school
‘schoolhouse’
- c. *ulo* *ukwu* → *uluukwu*
house big
- d. *o-to* *uto* → *otu* *uto*
3SG-is sweet sweetness
‘it is sweet’
- e. *a-cho-ro* *onye* *ozo* → *aḥoro* *onyo* *ozo*
3SG.INDEF-want-INDIC one another
‘someone wants another one’

(10) assimilation of o

- a. *do-ro* *ihu* *ebe* *a* → *doriihu* *eba* *a*
face toward-INDIC face place DEF
‘face towards this place’
- b. *ugo* *e-be-le* *ebe* → *uge* *ebele* *ebe*
eagle INFP-perch-PERF place
‘eagle perched’
- c. *ogologo* *akuko* → *ogologa* *akuko*
length story
‘a long story’
- d. *zu-oro* *uwa* → *zuoru* *uwa*
find-INDIC rest
‘found rest’

(11) assimilation of /i/

- a. o-nwe-ghi onye → onwegho onye
 3SG-has-NEG.INDIC one
 'she has no one'
- b. chi ọjọọ → ọjọ ọjọọ
 dietary evil

(12) assimilation of /i/

- a. ọ-di ihe ọbụọ → ọdi iha ọbụọ
 3SG-be thing two
 'there are two things'
- b. ụmụ nwoke a-na-ghi e-le → ụmụ nwoke anaghe ele
 PL man INFP-are-NEG.INDIC INFP-look at
 'men do not look at'
- c. anyi onwe anyi → anyo onwa anyi
 1PL self 1PL
 'we ourselves'
- d. ndi ọzọ → ndọ ọzọ
 person another

(13) assimilation of /u/

- a. o-nye-lu Idu ewu → onyeli idu ewu
 3SG-give-INDIC Idu goats
 'She gave Idu some goats.'
- b. o-nye-lu eze ihe → onyele ezi ihe
 3SG-give-INDIC chief thing
 'He gave the chief something.'
- c. ugbo a → ugba a
 time DEF
 'this time'
- d. ọ-gba-ta-ra
 3SG-participate-ASSOC-INDIC
 egwu ọ-gba-ra → ọgbatara egwọ ọgbara
 dance 3SG-participate-INDIC
 'She dances very well.'

(14) assimilation of /u/

- a. izuru ihe → izuri ihe
 2SG-buy-INDIC things
 'you bought things'
- b. ka ọ-puru ezi → kọ ọpuru ezi
 COMP 3SG-go-INDIC outside
 'when she got outside'
- c. ụmụ ọkwukwọ → ụmā ọkwukwọ
 children school
 'schoolchildren'
- d. ọrụ ọzọ → ọrọ ọzọ
 job another

As evidenced in 11–14, several cases of high vowels assimilating to a following vowel were found. There were many other high vowel sequences, however, where the first vowel sounded as if it had become a glide (as in 15), or in which no assimilation was heard (as in 16), even though there was no perceptible pause or phrase break.

(15) gliding of high vowels

- a. ndi_ɪ a → ndya_ɪ
people DEF
'the people'
- b. udi_ɪ onodu_ɪ a → udi_ɪ onodwa_ɪ
kind (of) condition DEF
'this kind of condition'
- c. ugbu_ɪ a e-kwu-ru → ugbweekwuru_ɪ
time DEF 3SG.INDEF-say-INDIC
'this time they say'

(16) sequences with no apparent assimilation

- a. iri_ɪ abuo_ɪ
ten two
'twenty'
- b. isi_ɪ akwu_ɪ
head (of) palm
'palm frond'
- c. eriri_ɪ ozo_ɪ ya_ɪ
chain title 3SG.POSS
'his ceremonial chain'
- d. nani_ɪ ada_ɪ
only Ada
- e. ndi_ɪ oru_ɪ bekee_ɪ
people (of) job foreign
'office workers'
- f. ugbu_ɪ a_ɪ
time DEF
'this time'
- g. oru_ɪ ozo_ɪ
job another
- h. i-gwa_ɪ mma_ɪdu_ɪ ihe_ɪ
INF-tell people things
'to tell people things'
- i. otutu_ɪ ajuju_ɪ
various questions

Assimilation of the high vowels was highly variable in this sample. For example, three repetitions of the phrase *na oru ahụ adighi* 'if that job is not ...' produced three different patterns, including no assimilation, gliding, and apparent complete assimilation.

- (17) na oru ahu a-di-ghi
 COMP job DEF INFP-IS-NEG.INDIC
 a. → n_ɔo_ɔru a_ɔhwa_ɔdighi
 b. → na o_ɔra_ɔhu a_ɔdighi
 c. → n_ɔo_ɔra_ɔha_ɔa_ɔdighi

In addition to affecting different vowel qualities differently, Igbo vowel assimilation is sensitive to phonological phrasing: the process occurs only at the boundary between two phonological words, usually only within smaller syntactic constituents such as noun phrases, prepositional phrases, complex verbs, and verb phrases. Assimilation between the subject of a sentence and the verb (Igbo word order is SVO) may occur if the subject noun phrase is short. Assimilation also occurs between the two elements of a compound word. (Recall that the failure of harmony to apply across the boundary between the two parts of a compound shows that they must be considered separate phonological words.)

Vowels are seldom adjacent within noncompound words, because of the overwhelming preponderance of CV syllables, but when such vowel sequences do occur both vowels retain their quality, as shown in 18 and 19.

(18) No assimilation within morphemes

- | | | | | |
|--------|--------|---------|----------|--------|
| a. bja | b. hie | c. abuo | d. awai | e. mai |
| come | tie | two | porridge | wine |

(19) No assimilation across morpheme boundaries

- a. o-so-ele (but dialectally o-so-b-ele or o-so-w-ele)
 3SG-follow-INDIC
 's/he has followed'
- b. o-cho-a_ɔ (but dialectally o-cho-b-a_ɔ or o-cho-w-a_ɔ)
 3SG-desire-INDIC
 's/he has desired'
- c. o-si-ele (but dialectally o-si-y-ele or o-sy-ele)
 3SG-cook-INDIC
 's/he has cooked'

It must be noted, however, that examples of nonidentical vowels within words are not numerous. Further, in 18 each word contains a high vowel, which some authors have argued never assimilate in any case. All of the forms in 19 have variant pronunciations in which an epenthetic consonant may be inserted to break up the vowel sequences. Nonetheless, it seems that the cases where vowel assimilation fails to apply must be accounted for.

The evidence presented thus far demonstrates that the facts of vowel assimilation cannot be captured by a simple rule schema such as that in 6. Vowel assimilation is sensitive to prosodic domains. It seems more likely to affect nonhigh vowels than high vowels. Most importantly, it appears that assimilation is not a categorical alternation. Many researchers (e.g. Green and Igwe 1963, Emenanjo 1978, Clark 1990) agree that for the high front vowels a remnant of

V1 remains after assimilation. According to Welmers (1973), assimilation is almost never complete for any vowel sequence, but is better described in terms of 'centralization' and 'transition'. Further experimental evidence for the gradient and variable nature of Igbo vowel assimilation is presented in §3.2.

3.2. GRADIENT ASSIMILATION: EXPERIMENTAL EVIDENCE. Zsiga 1993 reports the results of an experiment that gathered controlled phonetic data on the process of Igbo vowel assimilation. Three speakers of Igbo were asked to pronounce pairs of words in which two vowels are adjacent at the word boundary. Each of the eight Igbo vowels occurred twice as V1. The second word was either *ato* 'three' or *ozo* 'another', so V2 was either /a/ or /o/. The surrounding consonants were always coronal. These word pairs were placed in a sentence, and three different conditions were recorded: (1) sentences read fluently (SENTENCE CONDITION); (2) word pairs pronounced carefully (PAIR CONDITION); and (3) words in isolation (ISOLATION CONDITION). Vowel assimilation between the two words is expected only in the sentence condition. (Thus, the pair and isolation conditions are together referred to as the control conditions.) Two example sentences are given in 20.

- (20) a. for the word pair *asaŋo ato*
 o si na asaŋo ato bu ohu na otu
 3SG say that seven three be twenty and one
 'S/he said that three sevens make twenty-one.'
- b. for the word pair *oti ozo*
 o si na oti ozo da-ra n-ala
 3SG say that grub another fall-INDIC to-floor
 'S/he said that another grub fell to the floor.'

For the sentence condition, six repetitions of each token were elicited from each subject. Three repetitions of each token were elicited for the control conditions.⁸

The tokens were digitized and formant values were computed by linear prediction analysis. A 20 ms Hamming window was used, with analysis frames at 5 ms intervals. Formants in the vowel sequences were measured at three points: vowel onset, onset + 25 ms, and target (the point at which the vowel reached its most extreme articulation as seen in the formant trajectories). The duration of the vowel sequence was also measured for tokens in the sentence condition.

The three measurement points are illustrated in the spectrograms in Figure 2, made from utterances produced by subject 1. The token in Figure 2A is *ode* 's/he presses', spoken in isolation. In this token F2 remains steady at about 2000 Hz at all three points. Compare this to the token *ode ozo* 's/he presses another', spoken fluently in the sentence context (Figure 2B). If assimilation in this utterance were complete, the spectrogram would show a steady state

⁸ Subject 1 did not complete the experiment. For this subject, condition 3 was used for comparison to the sentence condition, while condition 2 was used for subjects 2 and 3. Few differences were noted between conditions 2 and 3 for the two subjects where both were collected, and results were comparable across the three subjects.

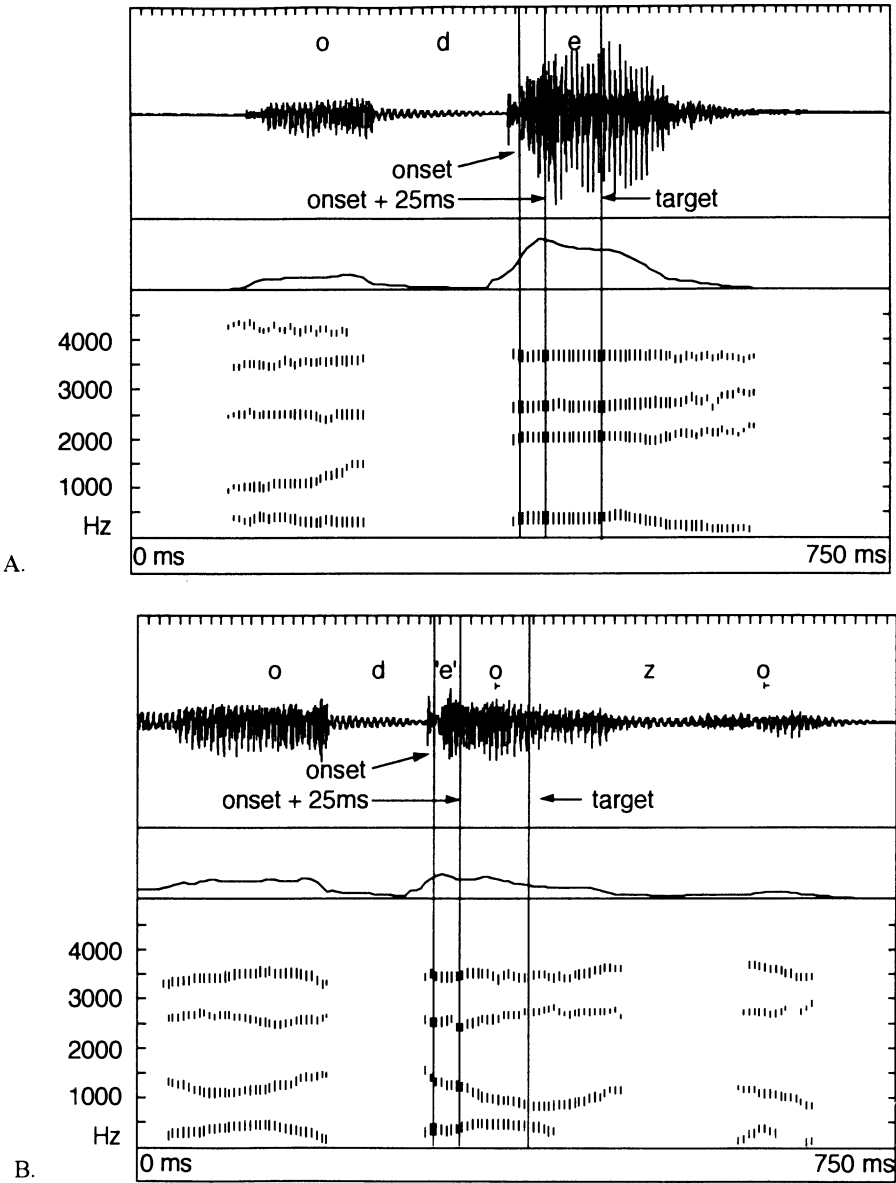


FIGURE 2. Spectrograms showing the points at which formant values were measured. A. The phrase *ode* 's/he presses'. B. The phrase *ode ozo* 's/he presses another'.

mid back vowel, with no trace of the /e/ remaining. There is no smooth transition from coronal to /o/, however. Rather, there is a plateau at around onset + 25 ms, where F2 remains relatively high. F2 is never as high as 2000 Hz, but it is higher than would be expected for /o/. This high F2 is a remnant of partially assimilated /e/.

In order to argue for a gradient assimilation, it must be shown that, at onset and onset + 25 ms, the formant patterns for V1 in the assimilation condition are different from those found for either V1 or V2 in the control conditions, as well as from those of an underlying V2#V2 sequence. That is, in a V1#V2 sequence formant patterns should fall in between those expected for V1 and those expected for V2. Figures 3 and 4 illustrate the range of variation that was

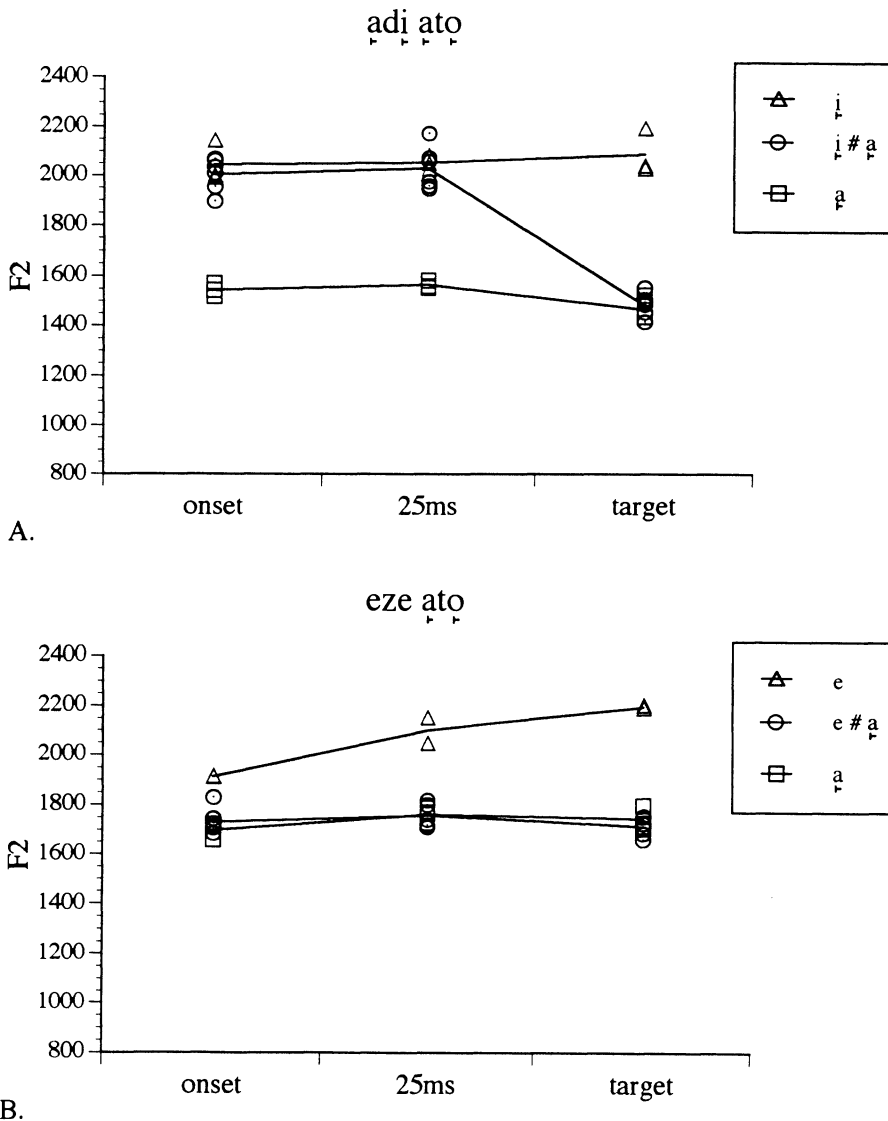
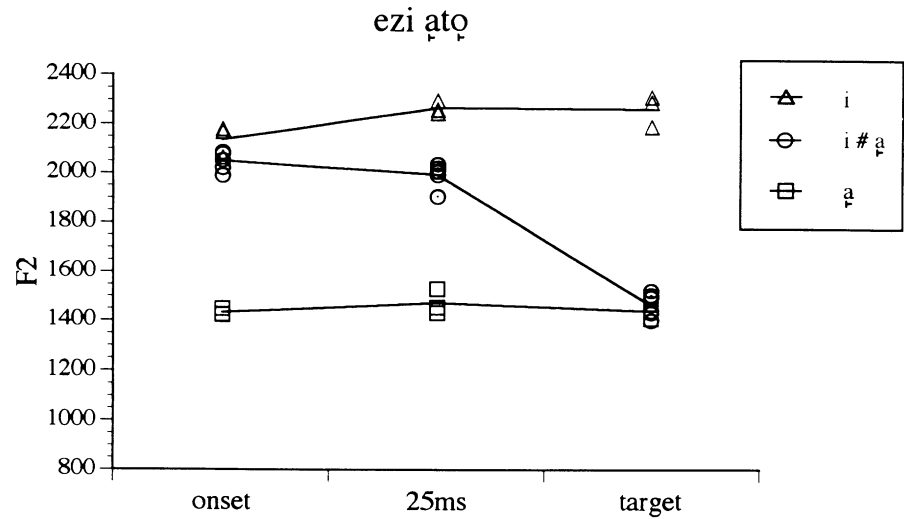
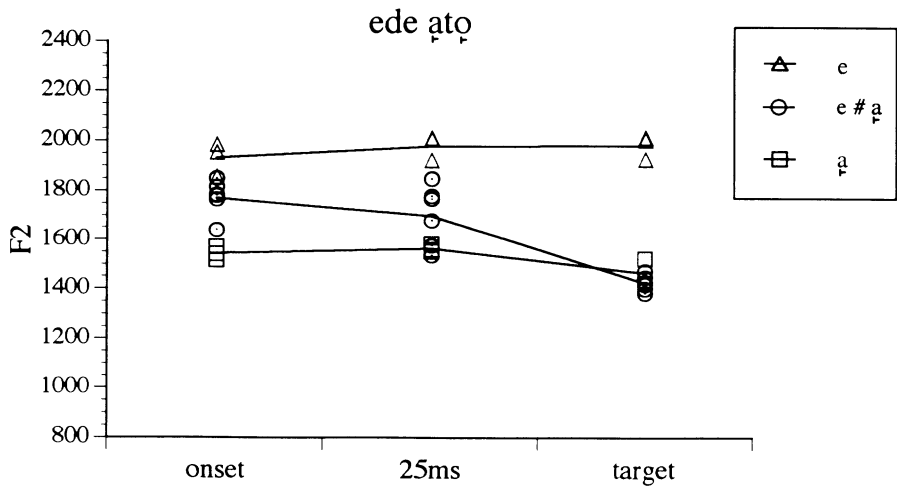


FIGURE 3. A. An /i#a/ sequence (*adi ato* 'if there are three'), compared to a carefully pronounced /i/ (*adi*, 'three') and /a/ (*ada* 'daughter'), as produced by subject 2. No assimilation is apparent. B. An /e#a/ sequence (*eze ato* 'three teeth'), compared to a carefully pronounced /e/ (*eze*, 'teeth') and /a/ (*nkata* 'basket'), as produced by subject 3. Assimilation is apparently complete.

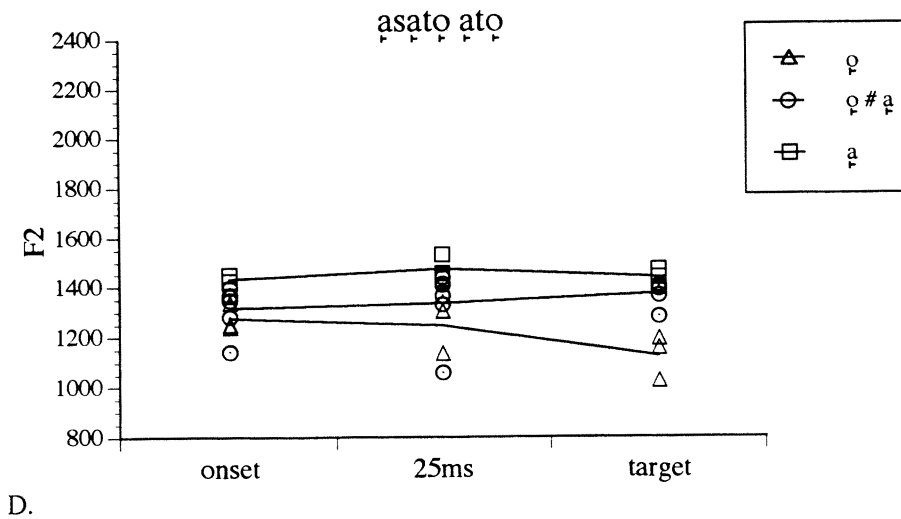
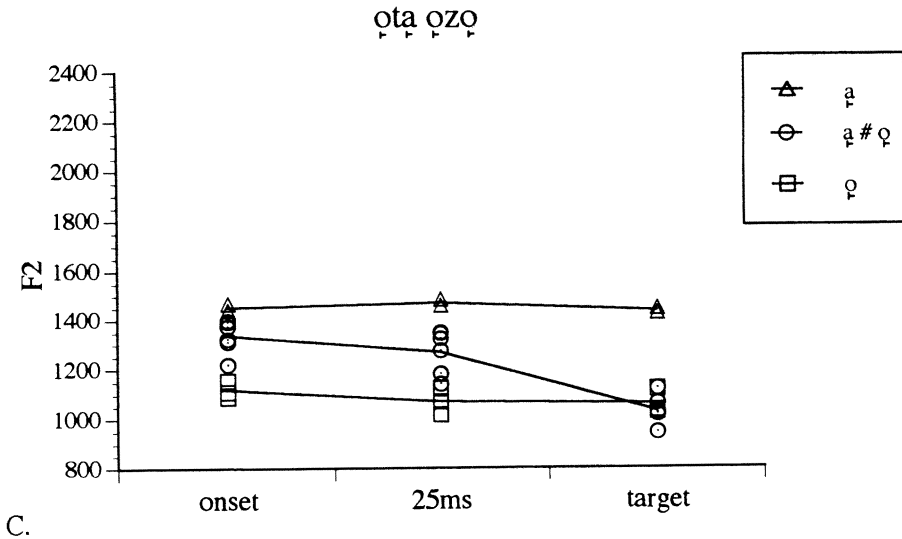


A.



B.

FIGURE 4. Tokens of partial assimilation from subject 2. A. An /i#a/ sequence (*ezi ɹ̥at̥ɔ* ‘three loans’) compared to /i/ (*ezi* ‘loan’) and /a/ (*nk̥at̥ɔ* ‘basket’). B. An /e#a/ sequence (*ede ɹ̥at̥ɔ* ‘three cocoyams’) compared to /e/ (*ede* ‘cocoyam’) and /a/ (*q̥at̥ɔ* ‘daughter’). C. An /a#ɔ/ sequence (*ɔt̥ɔ ɔz̥ɔ* ‘another shield’) compared to /a/ (*ɔt̥ɔ* ‘shield’) and /ɔ/ (*ɔt̥ɔ* ‘mud’). D. An /ɔ#a/ sequence (*q̥at̥ɔ ɹ̥at̥ɔ* ‘three sevens’) compared to /ɔ/ (*q̥at̥ɔ* ‘seven’) and /a/ (*nk̥at̥ɔ* ‘basket’).



found in this data. In the graphs in these figures, values of F2 in a fluently pronounced sequence of two vowels (sentence condition) are compared to F2 in each of the vowels pronounced carefully (control condition).⁹ Measurements at onset, onset + 25 ms, and target are shown. The figures show that assimilation results in a continuum of formant values.

⁹ Evidence for gradient assimilation was found to be at least as strong in measurements of F1. See Table 2.

Figure 3 illustrates the endpoints of the continuum. Figure 3A shows an /i#a/ sequence (*adi ato* 'if there are three') produced by subject 2. This vowel sequence is compared to a carefully pronounced /i/ (from the word *adi*) and carefully pronounced /a/ (from the word *ada* 'daughter'). As in each of the graphs, tokens of the V1#V2 sequence are plotted with circles, while tokens of V1 and V2 in the control condition are plotted with triangles and squares respectively. Mean values in each condition are connected by a solid line. In this /i#a/ sequence, no assimilation is apparent. The beginning of the sequence, at onset and onset + 25 ms, is indistinguishable from /i/, and at target the value expected for /a/ is reached.

In contrast, Figure 3B shows one of the few cases of apparently complete assimilation in the data. The graph shows tokens of an /e#a/ sequence (*eze ato* 'three teeth') produced by subject 3. Again, tokens from the vowel sequence are plotted with circles, while /e/ (*eze*) in the control condition is plotted with triangles, and /a/ (*nkata* 'basket') in the control condition is plotted with squares.¹⁰ For /e/ in the control condition, F2 rises steadily to reach a value of 2200 Hz. Values for the /e#a/ sequence, however, remain steady at about 1700 Hz, indistinguishable from the values for /a/.

If these two patterns were indicative of the entire data set, the conclusion would follow that Igbo vowel assimilation is a variable but categorical rule: sometimes assimilation does not apply, sometimes it does. However, these two cases represent only endpoints of a continuum. In most cases assimilation was partial, gradient, and resulting in a continuously changing output. Several examples are shown in Figure 4. For purposes of comparison, all the data in Figure 4 are taken from a single subject (subject 2), although all three subjects showed similar results.

Figure 4 shows the phrases *ezi ato* 'three loans', *ede ato* 'three cocoyams', *ota ozo* 'another shield' and *asato ato* 'three sevens'. In the /i#a/ sequence (Figure 4A), F2 begins high, but never moves up to the 2300 Hz value that /i/ reaches. Twenty-five ms after onset, the F2 values fall about halfway between /i/ and /a/, and then target for /a/ is reached. This illustrates clearly the pattern described by Welmers (1973): centralization of V1, followed by a rapid transition to V2. The variability of Igbo vowel assimilation is illustrated by the /e#a/ sequences in Figure 4B. Across six repetitions of the same phrase by the same speaker, the degree of assimilation varies from nearly none to nearly complete. While the target value for /a/ is eventually reached in these sequences, the beginning of the sequence differs from both /e/ and /a/. Note how the F2 values at onset and onset + 25 ms are scattered between the values for the control vowels. Some tokens are quite close to the expected values for /e/, others are quite close to or indistinguishable from the expected values for /a/. A similar pattern is seen for the /a#o/ sequence in Figure 4C. The /o#a/

¹⁰ Note that two different words, *ada* and *nkata*, are used as controls for /a/. Since it was found that F2 for all vowels was slightly higher at onset following /d/ than following the other consonants, *ada* is used as a control for words containing /d/, and *nkata* is used for words containing /t/, /s/, or /z/. No other significant difference due to preceding consonant was found (see Zsiga 1993).

sequence in Figure 4D is included to show a case where F2 RISES under assimilation. Since F2 for /a/ is higher than for /o/, in this case the line representing V2 is at the top of the graph. Although the F2 values for /a/ and /o/ are similar,¹¹ and one token seems to show no assimilation, other values for the /a#o/ sequence fall in between between the control vowels, rising from values typical of /o/ at onset to values typical of /a/ at target.

Across the data set, formant values in a fluently pronounced V1#V2 sequence were scattered between those found in the control condition for V1 and those found in the control condition for V2. These patterns were found in both F1 and F2, for all three subjects. Assimilation took place, but it was variable and partial.

The statistical analyses that tested for the generality of these patterns across the data set revealed a number of interesting results, especially with respect to the effects of assimilation on the different vowels. The tests for gradient assimilation were divided into three parts:

1. What difference is expected between V1 and V2?
2. Is the beginning of the V1#V2 sequence different (in the expected direction) from V1 in the control context? That is, is there evidence of any assimilation at all?
3. Is the beginning of the V1#V2 sequence different (in the expected direction) from a V2#V2 sequence? That is, is there evidence that assimilation is not complete?

Tests were conducted separately for each subject. Results for F1 and F2 at onset + 25 ms are shown in Table 2. For the full set of statistical analyses for both F1 and F2 at onset and onset + 25 ms, see Zsiga 1993.

To determine what difference is expected between V1 and V2, an analysis of variance was performed for each subject on the target formant values of the eight vowels in the control conditions. Values from /a#a/ and /o#o/ in the sentence condition were also included in the analysis. The main effect of vowel was highly significant for each subject ($p < .0001$), and post hoc pairwise comparisons showed that, for the most part, vowels differed in the expected directions (significance level $p < .05$). For all subjects in both formants, /a/ in the control condition did not differ from /a#a/ in the sentence condition, and /o/ did not differ from /o#o/. Otherwise, all vowels showed F1 values significantly lower than /a#a/ and all vowels except /a/ showed F1 values significantly lower than /o#o/. Thus, in all cases except for assimilation of /a/ to /o/, F1 is expected to rise. All front vowels showed F2 values higher than /a#a/, and almost all vowels (other than /o/) showed F2 values significantly higher than /o#o/. Thus for most vowel combinations, under assimilation, F2 would be expected to fall (as in Figure 4A–C). Some back vowels showed F2 values lower than /a#a/; in these cases F2 would be expected to rise under assimilation (as in Figure 4D). Surprisingly, however, in a number of cases the F2 values of the back

¹¹ Because of the similarity in formant values, the difference between the sentence and control condition did not quite reach significance ($p = .08$).

TABLE 2. Results of statistical tests for partial assimilation.

Igbo phrase	Subject 1		Subject 2		Subject 3	
	F1	F2	F1	F2	F1	F2
ezi ɔ̀ɔ̀ 'three loans'	**	**	**	**	**	**
akpati ɔ̀ɔ̀ 'three boxes'	**	**	**	**	**	**
ihe ntumadi ɔ̀ɔ̀ 'three accidents'	**	**	**		**	**
adi ɔ̀ɔ̀ 'if there are three'	**	**	**		**	**
eze ɔ̀ɔ̀ 'three teeth'	**	**	**		—	—
ede ɔ̀ɔ̀ 'three cocoyams'	—	**	*	**	—	**
esu ɔ̀ɔ̀ 'three millipedes'			**		**	
adu ɔ̀ɔ̀ 'three kolas'			**		**	
atu ɔ̀ɔ̀ 'three chewing sticks'	**		**		**	
mmaɔ̀ ɔ̀ɔ̀ 'three people'	—		—		—	—
eso ɔ̀ɔ̀ 'he follows three'	—		**		—	—
onɔ̀ edo ɔ̀ɔ̀ 'three gold pieces'	**		**		*	
ato ɔ̀ɔ̀ 'three threes'	*	—			*	—
asato ɔ̀ɔ̀ 'three sevens'	—	—	—		—	—
ozi ɔ̀ɔ̀ 'another errand'	**	**	**	**	*	**
ugodi ɔ̀ɔ̀ 'another padlock'	**	**	**	**	**	*
oti ɔ̀ɔ̀ 'another grub'		**	**	**	**	**
odi ɔ̀ɔ̀ 'there is another one'	**	**	**	**	**	**
ose ɔ̀ɔ̀ 'another stream'	—	*	**	**	—	**
ode ɔ̀ɔ̀ 'he presses another'	*	—	**	**	**	**
ozu ɔ̀ɔ̀ 'another corpse'	**	**	**	**	**	
otu ɔ̀ɔ̀ 'another organization'	**	—	**	—	*	
odu ɔ̀ɔ̀ 'another tusk'	**		**	—	**	
odu ɔ̀ɔ̀ 'another tail'	**		**	—	**	
okoso ɔ̀ɔ̀ 'another toy top'	—	**	**	*	*	
obodo ɔ̀ɔ̀ 'another town'	—	**	**	**	—	
ota ɔ̀ɔ̀ 'another shield'	—	—	—	**		—
oda ɔ̀ɔ̀ 'another permit'	—	—		**		*



assimilation could not be tested:
no difference between control V1 and control V2.



no assimilation:
no difference at onset + 25 ms between control V1 and V1#V2 sequence

vowels were not found to be significantly different from the values for /a#a/ or /o#o/. For subject 1, /u/ and /o/, and for subject 2 /u/, /u/, and /o/ were not distinct in F2 from /a#a/; and for subject 3, /u/, /u/, and /o/ are not distinct from /o#o/. These cases are shown as solidly filled cells in Table 2. No further analysis of assimilation could be performed for these cases, since, if F2 values for V1 and V2 are not significantly different, no prediction can be made for these vowels about the expected direction of formant change under assimilation.

Further examination of the data suggested that the lack of significant differences in F2 among these back vowels was due to unexpectedly high F2 values for these vowels in the control conditions. (Compared to, for example, the values for Igbo vowels given in Ladefoged & Maddieson 1990, these anomalous tokens appear to be fronted.) For these back vowels in this experiment, F2 did not fall to the expected value for /u/ or /o/, but remained high after the release of the coronal consonant. This strong effect of a preceding coronal consonant, such that the expected acoustic consequences of a high back vowel tongue shape are not reached, suggests that these word-final vowels were reduced. (No large effect on F2 would be expected for the front vowels or /a/, for which F2 was already high.) Word-final reduction, independent of vowel assimilation, is discussed further in §3.3.

For those sequences where significant differences in formant values between V1 and V2 WERE found, a second statistical test was conducted to seek evidence of assimilation in these tokens. Formant frequencies for V1 in the sentence condition were compared to formant frequencies in the control condition in an analysis of variance.

So that main effects could be interpreted directly as evidence for assimilation, vowel sequences were divided into subsets according to the identity of V2 and the predicted direction of formant change. Separate analyses of variance were performed on each subset. Factors in each analysis were V1, identity of the lexical item (each vowel combination was represented by two different words), position (onset vs. onset + 25 ms, treated as a repeated measure) and condition (control vs. sentence).

A significant effect of condition indicates that the vowels in the sentence condition differ from vowels in the control condition in the appropriate direction, indicating at least some effect of assimilation. Where there was a significant interaction of condition with V1, the effects of condition were tested separately for each vowel, and if the analysis of a given vowel showed an interaction of condition and lexical item, each lexical item was tested separately (see Zsiga 1993 for details.) The effect of condition at onset + 25 ms was signifi-



partial assimilation:

values at onset + 25 ms significantly different from V1 and V2 ($p < .05$)



partial assimilation:

values at onset + 25 ms significantly different from V1 and V2 ($p < .01$)



complete assimilation:

no difference at onset + 25 ms between control V2 and V1#V2 sequence.

cant ($p < .05$) everywhere, at the appropriate level of breakdown, except for those cells shaded light gray in Table 2. That is, for all sequences except those shaded gray (including those shaded medium gray, which couldn't be tested), some effect of assimilation was found.

Clearly, the cases in Table 2 that show no assimilation effect are not randomly distributed. They fall into two classes: words ending with a high vowel, and /a#o/ sequences where the F1 values of V1 and V2 are similar. The only exceptions to this generalization are *eze aṭo* for subject 2, and two cases where the effect of condition approached significance: *asatō aṭo* for subject 2 ($p = .08$) and *edo aṭo* for subject 3 ($p = .06$). Where the formant values of V1 and V2 are close, it is assumed that any change due to partial assimilation might be too small to be statistically significant. The lack of significant effects for the high vowels is consistent with previous assertions (see §3.1) that high vowels are less likely to be heard as assimilated than nonhigh vowels. In this data set, however, there were numerous cases where high vowels did show assimilatory effects. There were, in fact, no cases where an item failed to show an assimilation effect in at least one of the formants (where both were tested). For all vowels, V1 in the sentence condition became more like V2.

At least some assimilation is evident in these tokens: vowels in the assimilation and no-assimilation contexts are different, in the direction predicted. The important question, however, is whether assimilation applies categorically, leaving no evidence of V1 in the acoustic record, or whether, despite the effects of assimilation, a residue of V1 remains. Evidence for a residual V1 can be seen if V1#V2 sequences in the sentence condition differ significantly from underlying /a#a/ and /o#o/ sequences in the same condition. For each item where there was evidence for assimilation (a significant effect of condition), a t-test was used to compare the V1#V2 sequence with the appropriate V2#V2 sequence. The results of the t-tests are shown in Table 2. If the value of the t-statistic was significant ($p < .05$) the cell is marked with an asterisk. Highly significant results ($p < .01$) are marked with a double asterisk. These are the cases that clearly show partial assimilation. The formant values of V1 in a V1#V2 sequence are significantly different from both V1 and V2.

If V1#V2 does not differ significantly from V2#V2 in either F1 or F2, at either onset or onset + 25 ms, assimilation is considered complete. These cells are indicated with a minus sign in Table 2. Subjects 1 and 3 produced items with apparent complete assimilation:¹² subject 1 *asatō aṭo*, *ota ozo*, and *oda ozo*; and subject 3 *eze aṭo*, *eso aṭo*, *aṭo aṭo*, and *asatō aṭo*. For the most part, these are the sequences of vowels whose formants are most similar: /a/ and /o/. Other items showed significant differences in one formant but not the

¹² Note that for subject 2 *eze aṭo* showed no assimilation in F2, for subject 3 this token showed complete assimilation in both formants, and for subject 1 this token showed partial assimilation in both formants. These results are interpreted as idiosyncratic differences between the subjects in the pronunciation of this phrase. The case of *asatō aṭo* for subject 2 looks anomalous, because the statistical analysis indicates complete assimilation in F1 and no assimilation in F2. However, F1 values for /a/ vary considerably and can be similar to the F1 values for /o/, so it is not too surprising that the difference between them should not turn out significant.

other (*mmaḍu ɔtɔ* for S1), or at onset but not at onset + 25 ms (*mmaḍu ɔtɔ* for S3). Of course, assimilation cannot APPLY to one formant but not the other. Depending on the vocal tract shape that results when assimilation applies, and on the differences in the formant patterns of unassimilated V1 and V2, assimilation may be evident in one formant but not the other.

The amount of assimilation varied, from little to apparently complete. Difference or similarity in the articulations of V1 and V2 seem to play a large part in how assimilation applies: assimilation is most likely to appear complete for vowel sequences that are most similar, and least likely to have a significant effect for vowel sequences that are most different. Overall however, the statistics reported in this section show that, in almost every case, V1#V2 sequences in the sentence condition differ both from V2#V2 controls in the sentence condition and from V1#V2 controls in the pair or word condition. Further, the differences lie in the direction expected for assimilation. Assimilation takes place, and it is not complete.

One further experimental finding is of interest here. Consistent with reports in the literature, it was found that the duration of an assimilated sequence was no shorter than a sequence of two vowels. No consistent correlation was found between the measured value of F2 at onset + 25 ms and the duration of a given vowel sequence. In an /e#o/ sequence, for example, those tokens that had low F2 values at onset + 25 ms, indicating a greater degree of assimilation, were just as long as tokens that had high F2 values at the same point.

3.3. A GESTURAL ACCOUNT. The experiment reported above has shown that a complete linguistic description of Igbo vowel assimilation should account for the following findings:

1. Igbo vowel assimilation is gradient, with the formants of V1 falling in between the values expected for V1 and the values expected for V2.
2. Assimilation is variable. Depending (in part) on the differences in the formant patterns of an underlying V1 and V2, assimilation may appear absent, partial, or complete.
3. Vowels are reduced word-finally.
4. The duration of a vowel sequence is not correlated with the amount of assimilation. Sequences showing extensive assimilation are no shorter than those showing little or no assimilation.

I argued elsewhere (Zsiga 1993) that these acoustic patterns can best be accounted for in terms of changes in the temporal relations among overlapping articulatory gestures. After reviewing those arguments I will argue that in Igbo changes in phase relations among the vowel gestures, driven by temporal reduction of the word-final vowel gesture, result in the perception of assimilation. The first vowel is shortened; the second vowel then lengthens, preserving the overall syllable duration and tonal pattern. The combination of a shortened gesture for V1 and a lengthened gesture for V2 leads to the perception of partial assimilation.

As mentioned above, word-final back vowels were found to have an unexpectedly high F2. This acoustic result is consistent with shortening or weaken-

ing of a word-final vowel gesture after a coronal consonant (Ohala 1981, Wood & Pettersson 1988). While it was not possible to accurately measure word-final vowel length, impressionistically, final vowels in the words in isolation sounded shorter and less forcefully articulated than initial vowels. (In fact, sixteen tokens from the pair and isolation conditions had to be eliminated from the analysis because formants could not be measured due to low amplitude or lack of voicing during the vowel. The quality of these vowels was little more than coronal release.) This suggests that the word-final syllable in Igbo is a weak position, in which the acoustic vowel target is not fully reached.¹³ While the tongue body may assume a normal position, other vocal-tract adjustments that would make the vowel quality clearly audible may not be made. For an /a/ vowel, for example, the jaw may not reach a fully lowered position, or the tongue tip may remain raised.

One way to model this word-final weakening is simply to reduce the duration of the vowel gesture, without any modification to vowel target. If the activation period of the vowel gesture is shortened, the articulators will not reach their specified positions. It is not possible, based on the data from this experiment, to prove definitively that vowel shortening, rather than, for example, consonant lengthening, is responsible for the effects on word-final back vowels seen here; data from word-final vowels in other consonantal contexts would be necessary.¹⁴ (Coronals were used here because a complete set of plausible, phonetically controlled phrases using labial or velar consonants could not be found.) Vowel shortening, however, is consistent with all of the effects found here, and it offers a plausible link between the F2 effects found for the back vowels in the control conditions and vowel assimilation in fluent phrases.

When a word-final vowel gesture is shortened, the syllable may or may not be actually reduced in length, depending on the temporal patterning of other gestures. If no other vowel gesture follows, the word-final vowel may be perceived as shortened and reduced. When another vowel gesture does follow, reduction in the word-final vowel may lead to the perception of assimilation. The proposed relationship between reduction and assimilation is illustrated in Figures 5 and 6.

¹³ It is difficult to determine on the basis of the data collected in this experiment whether the reduction is a word-final or phrase-final phenomenon. In the word condition, each word forms its own phonological phrase, so the vowels measured are both word- and phrase-final. In the phrase condition vowels were measured in the middle of a two-word phrase, but as the words were pronounced slowly, it might be argued that each word formed its own phonological phrase in this condition as well. Based on F2 values and the differing amounts of overlap among the vowel distributions in the two conditions, reduction was less extreme in the pair condition. The likely conclusion is that vowels are reduced word finally, and reduced even more phrase finally.

¹⁴ It should be noted as well that word-final vowel shortening in Igbo is not consistent with the word-final vowel lengthening found by Beckman and Edwards (1990) for English. This suggests that the effects of final position might be language-specific. Beckman and Edwards themselves note that the effect of word-final lengthening is small and 'not consistently evident across speakers and rates' even for English (1990:176). While some languages may show word-final lengthening, the fact that apocope is also a common process cross-linguistically suggests that final shortening or weakening is equally possible.

Figure 5 shows a vowel sequence with no assimilation. The top of the figure shows the waveform and formant tracings for the phrase *ede ɛtɔ* 'three cocoyams', spoken in a fluent phrase (one of the tokens that was graphed in Figure 4B). In this case, the formant patterns show two clearly articulated vowels. Below the spectrogram is a hypothetical gestural score, showing the pattern of gestural organization that might have produced this spectrogram. The gestural score diagrams tongue tip gestures for the /d/ and /t/, tongue body gestures for the vowels, and a tongue root constriction gesture for the /a/. Two laryngeal gestures for tone are also shown. Recall that in the gestural scores, the horizontal dimension represents time, and the length of the boxes indicates the period

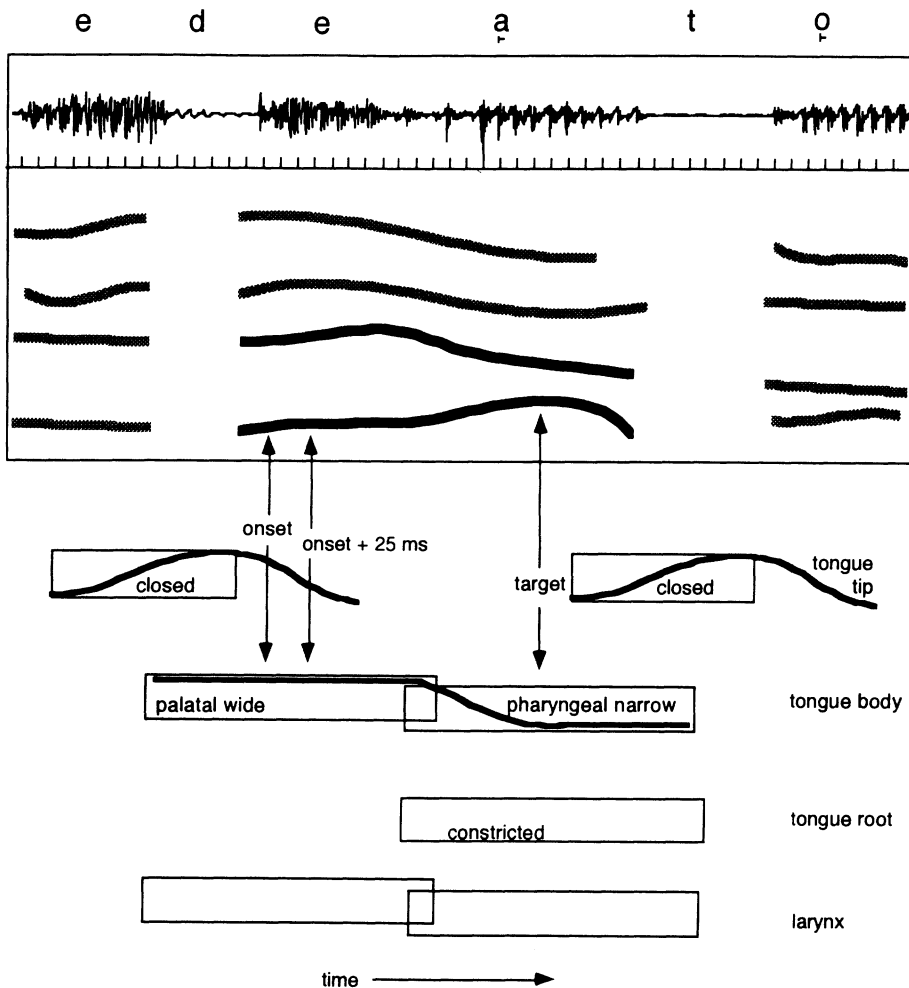


FIGURE 5. A vowel sequence with no assimilation. The top of the figure shows the waveform and formant tracings for the phrase *ede ɛtɔ* 'three cocoyams', spoken in a fluent phrase. Below the spectrogram is a hypothetical gestural score, showing the pattern of gestural organization that might have produced this spectrogram.

of time the articulator set is actively controlled. Hypothetical movement trajectories of the articulators (in the vertical dimension) that would be produced by the gestures are overlaid on the boxes.

The tongue body is high for the /e/, and then moves down for the /a/. The tongue tip, in the meantime, moves up to and then away from the alveolar ridge for the /d/, and then repeats the movement for the /t/. Because there is a full vowel gesture for both vowels, and little overlap between them, target is reached and maintained for both V1 and V2. At onset and onset + 25 ms, the spectrogram shows formants typical of /e/.

When the underlying gestural pattern changes, so does the resulting acoustic pattern. In a gestural approach, the prominence of V2 in an assimilation context can be accounted for by an increase in the duration of the V2 gesture.

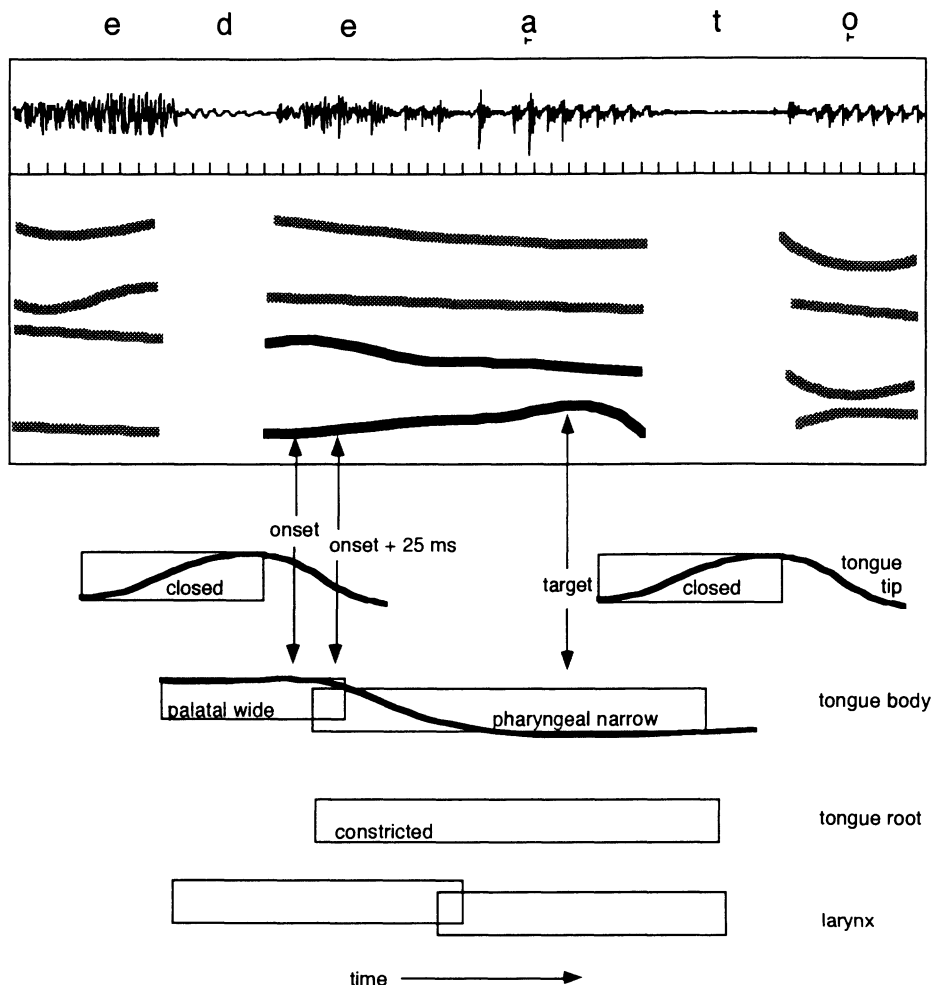


FIGURE 6. Another token of *ede atə*, showing partial assimilation.

an increase can be modeled in different ways, although the fact that the overall duration of the vowel sequence does not vary under assimilation constrains the possibilities here. The least complicated and most explanatory possibility is simple reduction of the V1 gesture.

Figure 6 shows another token of the phrase *ede atọ* pronounced by the same speaker (also from Figure 4B). In this token, the formant pattern starts like that typical of an /e/, but begins to move toward an /a/ pattern by 25 ms after onset. As shown in the hypothetical gestural score, this partial assimilation can be captured by making one change to the gestural score: reducing the duration of the /e/ gesture.

All other timing relations remain the same as in the previous figure—there is the same amount of overlap between V1 and V2, and the same overall duration. (It is assumed in a gestural account that timing relations will remain constant unless an explicit instruction to change is issued. Thus an account that changes only one parameter of gestural organization is the simplest.) When V1 is shortened, if the specified degree of overlap and total vowel duration remain constant, V2 must extend in duration. The result is the perception of assimilation. Assimilation is not complete, however, because a remnant of V1 remains. At onset + 25 ms, articulation is influenced both by the V1 gesture and the V2 gesture, resulting in formants in between V1 and V2.

Linking assimilation to word-final reduction accounts for the fact that assimilation applies only across word boundaries. The shortening approach can also account for the fact that the effects of assimilation were found to be least for the high front vowels, and greatest for /o/ and /a/. The acoustic effects of the high vowels persist because they have the most extreme articulations.¹⁵ The reduction in duration may never be sufficient to produce the perception of complete assimilation for the high front vowels. Because /i/ and /i̥/ are so different from /a/ and /o/, significantly more reduction would have to take place before their acoustic effect was lost. Conversely, the influence of an /a/ on a reduced /o/ is likely to be great enough to be perceived as complete assimilation.

Another possible change in gestural organization would be to increase overlap between the V1 and V2 gestures: lengthening of V2 without any concomitant shortening of V1. Browman and Goldstein (1990) argue that when two gestures that control the same articulator set—the tongue body in the case of vow-

¹⁵ Emenanjo (1978:25) relates assimilation of the high front vowels to tone. He states that /i/ and /i̥/ may become glides, but only if V1 and V2 are on the same tone level. Because of a tonal raising rule that applies in Igbo noun phrases, this prediction could not be directly tested in this experiment: V1 and V2 were, for the most part, pronounced on the same tone level. If, as is hypothesized here, the tongue body and laryngeal gestures can be timed independently, it is not clear why assimilation should be dependent on tone, or why tone should be a factor only for the high vowels. Emenanjo does not explicitly discuss the possibility of partial assimilation, however. It may be that acoustic changes are evident in the high vowels regardless of tone pattern. When the tones of V1 and V2 are the same, the two vowels are not clearly delimited, and the effect may be described as gliding of V1. When there are two different tones, two different vowels are heard distinctly. If, consistent with the acoustic patterns found here, the partially assimilated sequence sounded more like /i#a/ than /a#a/, it would not be classed as an assimilated sequence. Thus assimilation of V1 to V2 would seem to depend on the two vowels being on the same tone level.

els—overlap in time, each gesture exerts its influence on the articulator, which reaches a position intermediate between the two conflicting targets. Acoustically, the result is partial assimilation: formants intermediate between those that would be expected from either articulation in isolation. Again, more extreme articulations will be the most evident in spite of blending, and thus will seem to persist. The blend of an /i/ and /a/ will differ more from an underlying /a/ than will the blend of an /o/ and an /a/.

While blending due to increased overlap might account for the patterns of partial assimilation seen here, it is not the most likely gestural account for Igbo vowel assimilation, because it fails to link assimilation to word-final reduction. Under the account illustrated in Figure 6, vowel assimilation can be seen as a result of the gestural reorganization caused by the independently needed process of vowel shortening. No special rule of vowel assimilation is needed. When V1 is shortened, V2 is lengthened in compensation. Lengthening of V2 must occur in conjunction with V1 reduction, not only to explain the prominence of V2 in perception, but to account for the consistent duration of the vowel sequences.

Determining how much V2 should lengthen in order to maintain a consistent duration overall poses a problem for articulatory phonology. In autosegmental phonology, the presence of a hierarchical syllabic or prosodic structure, which remains constant over changes at the segmental level, controls the overall duration (Hayes 1989, Hyman 1985, McCarthy 1981, Wetzels & Sezer 1986). Articulatory phonology, which does not encode a hierarchical syllable structure, and which does not allow the specification of timing in milliseconds, relies on inter-gestural timing to control duration. That is, if the duration of a sequence remains constant, this must result from the duration and phasing relations of some particular gestures remaining constant. Which gestures remain constant in Igbo vowel assimilation? One possibility is that the duration and phasing of the laryngeal gestures for tone do not change. This would link the fact that the duration of two vowels is retained to the fact that two tones must be articulated. This tone-based approach works for Igbo because tones and syllables are usually coextensive; unlike other related languages, Igbo (for the most part) allows no contour tones. In many languages, however, tonal patterns expand and contract freely to fit the available tone bearing units (Leben 1978, Hoffman 1963). Thus it seems unlikely that laryngeal gestures for tone could in general serve to maintain an invariant duration. If no hierarchical structure is assumed, however, the tonal gestures are the best candidates for maintaining, by their fixed relations, a constant overall duration.¹⁶

In contrast to articulatory phonology, the theory being developed here assumes that hierarchical structures such as syllables and larger prosodic domains are part of the phonological representation. Such units could be instantiated gesturally, however, as determinants of specific temporal relations among ges-

¹⁶ Another candidate would be a tongue root gesture for [ATR]. No attempt was made in this experiment, however, to tease apart the effects of tongue root and tongue body gestures, so this approach is not pursued.

tures. Given the distributions of tones and the principles of syllable structure in Igbo, such an instantiation could include specification of an invariant duration for tonal gestures in this language. Articulatory phonology would specify these gestural relations directly, without reference to higher level domains. In the diagrams in Figures 5 and 6, the consistent duration of the vowel sequence is represented by two invariant laryngeal gestures. Whether the duration of those gestures is in turn determined by a higher level structure does not directly affect the rest of the argument.¹⁷

Vowel assimilation and vowel reduction can thus be seen as the result of the same process—shortening of a word-final vowel. When a word-final vowel is also phrase final, its acoustics are influenced by the preceding consonant, in this case a coronal, leading to a high F2. When another vowel follows, the gesture for the second vowel is lengthened, filling in the duration required for the articulation of two tones. The result is a gradient assimilation: a reduced V1 followed by a lengthened V2. As described by Welmers (1973:41), ‘the quality of the first vowel merges very quickly into the quality of the second.’

The gestural account thus describes the acoustic patterns seen in Igbo vowel assimilation, and links assimilation to word-final reduction. Both processes can be described in terms of a simple change in temporal patterning among gestures. Other phonetic approaches do not capture the facts as well.

Coarticulation and simple reduction are among the most plausible phonetic approaches to the data reported here. (Symbolic phonological approaches are considered in §3.4.) These approaches either fail to account for the facts of Igbo vowel assimilation, or require an analysis that is more complicated and less well-motivated than the gestural account.

A plausible phonetic approach would be to invoke simple coarticulation between vowels at word boundaries. In a sense, Igbo vowel assimilation is the result of coarticulation, at least in part. The two vowel articulations do influence one another. The noticeable changes in vowel quality found here, however, are clearly greater than could be accounted for by the minimal amounts of articulatory overlap and acoustic influence found in vowel sequences cross-linguistically (Öhman 1966, Manuel & Krakow 1984, Choi & Keating 1991).

Vowel reduction alone might be invoked without any other change in articulation. The discussion above outlines an approach that accounts for Igbo vowel assimilation in terms of reduction of V1, compensated for by an increase in the duration of V2. The facts cannot be accounted for by reduction in V1 without an increase in the influence of V2. That is, V1 in this data was not just reduced to a schwa-like vowel, but became more like V2 in its acoustic characteristics.

Keating’s (1990b) approach to vowel reduction, the WINDOW MODEL OF ARTICULATION, does seem to make predictions very close to those made by the gestural account. In this theory, a WINDOW of possible values is specified for the phonetic

¹⁷ The actual temporal patterning of the tonal gestures could of course be tested empirically. By tracking the pitch contours over sequences of vowels with different tones, one could determine whether the tone pattern varied with differing amounts of assimilation. In this experiment, however, the vowels in the assimilation context were for the most part on the same tone level.

parameters (such as tongue position, velum position, formant values, and so on) needed for the implementation of a given segment. Thus vowels may have a window for tongue body position (or F2 value) within which the vowel would have to be realized. A small window requires precise articulation (as for /s/). A larger window, however, would allow for greater coarticulation with neighboring segments, since an articulator will follow the shortest path that falls within the values allowed. That is, if a preceding or following vowel requires the tongue body to be high, the tongue body will remain at the upper end of the window specified for the current vowel. The larger the window, the more freedom for the current vowel to move toward the position of the vowel following. Thus word-final vowel reduction in Igbo might be modelled as window-enlargement. The predicted path through a fairly wide window from previous vowel target to following vowel target would account for some of the findings here. Formants measured in the word-final vowel would not be as extreme as in a carefully articulated (small-window) vowel, but would be more influenced by the following vowel, resulting in apparent partial assimilation.

The gestural and windows approaches in fact seem to make very similar claims about Igbo vowel assimilation. In both approaches, the demands on the articulator are relaxed, so that the target position for V1 is only approximated, not fully reached, and a greater influence of the following vowel is seen. The gestural approach, as argued above, could produce this effect by shortening the time window allowed for the articulation. Keating's windows approach might produce this effect by enlarging the place-of-articulation window. The details of the trajectories seen in this experiment, however, argue for the gestural interpretation.

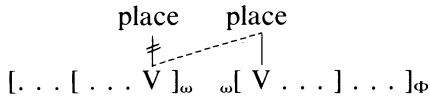
In many cases, the target for V2 is reached very early in the vowel sequence. (Recall Welmers' description of a 'rapid transition' to V2.) The window model predicts a long, smooth transition rather than an early realization. To express this steep trajectory, the V1 window would have to be shortened as well as widened, and the V2 window would have to lengthen in compensation. It is an advantage of the gestural approach that the tongue body position (and thus the acoustic changes) follow from the reduction in time, and do not have to be separately stipulated.

Overall, the gestural approach gives the best account of the acoustic results of Igbo vowel assimilation. Intermediate formant patterns result from the temporal organization of the two tongue body gestures. Reduction of a word-final vowel gesture links vowel shortening and vowel assimilation. Difference or similarity in the articulations of V1 and V2 coupled with variability in the amount of reduction, leads to variability in assimilation (i.e. is it absent, partial, or complete). Consistent duration is maintained by an unchanged temporal pattern for the other gestures. Other phonetic analyses fail to account for at least some of these acoustic results.

Section 3.4 turns to featural accounts of Igbo vowel assimilation. As will be shown, any analysis couched in terms of associations between features and hierarchical nodes will fall short on both descriptive and theoretical grounds.

3.4. EVIDENCE AGAINST A FEATURAL ACCOUNT. In previous accounts, Igbo vowel assimilation has been described as a rule spreading the place features of one vowel to another. The formalization from Zsiga 1992 is given in 21.

(21) Vowel assimilation as feature spreading



In this formalization, the place features of V2 spread onto the place features of V1, and the original place features of V1 are delinked. Spreading takes place only at the boundary between phonological words (ω), within a phonological phrase (Φ). Such a rule falls short on descriptive grounds because it indicates a categorical assimilation. It presents theoretical difficulties as a DOMAIN JUNCTURE RULE, a rule type that has proven problematic in phonological theory. These two points will be discussed in turn.

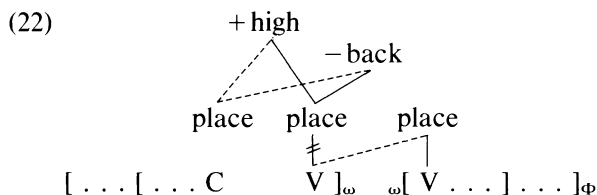
First, the rule in 21 is empirically inadequate. The output of this rule, two vowels linked to a single place node, is not distinct from the representation of a sequence of underlyingly identical vowels. As has been shown in §3.2, however, the output of assimilation is distinct from a sequence of identical vowels. A feature-geometrical representation provides no straightforward mechanism for indicating partial or gradient assimilations. A feature may be linked to one root node, or two (as in an assimilated [mp] sequence), or several (as in a single [round] feature linked to all the vowels in a word by principles of vowel harmony), but there is no way for a feature to link part-way to a root node, indicating a partial assimilation. Because of its categorical nature, needed to express lexical alternations, spreading and delinking will always be inadequate for representing noncategorical processes, such as different degrees of vowel nasalization, partial devoicing, or partial assimilation.

Note that, for the purposes of this paper, it is not crucial whether categorical phonological alternations are expressed as autosegmental spreading rules (as proposed by Goldsmith 1976, Clements 1985, Sagey 1986, and others) or as the product of interacting ranked constraints (as in Prince & Smolensky 1993). In both kinds of representation, the temporal relations between features are expressed through linkings to abstract hierarchical nodes, and so are necessarily categorical. If all phonological alternations are expressed in terms of the addition or deletion of association lines, then all phonological alternations will be categorical, whether the particular pattern of association is arrived at through derivational rules or the choice of an optimal parse of a given underlying string. Either a feature is linked to a higher level node or it is not.

The empirical problem with spreading the place features of V2 and delinking the place features of V1 to account for Igbo vowel assimilation is that a remnant of V1 clearly persists in assimilation contexts. In order to preserve a remnant of V1, one might argue that the assimilated vowels become glides. Emenanjo (1978) describes partial assimilation of the high front vowels as gliding. In one sense, the term GLIDING might be used to describe the acoustic effect of high

vowel reduction. However, this term is usually used to mean a change in the status of a high vowel from syllable nucleus to syllable onset, and this will not account for the facts of assimilation found here. Assimilation in Igbo applies to all vowels, and nonhigh vowels have not been found to become glides. A different process would have to be proposed to account for the assimilation, for example, of /*o*/ to /*a*/. In addition, the assimilated vowels continue to bear tone.

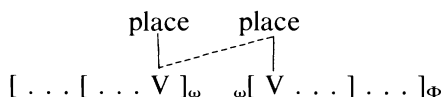
Clark (1990), in order to account for apparent partial assimilation of the high front vowels, proposes spreading the delinked vowel features onto the preceding consonant. A rule similar to Clark's is shown in 22. (Her actual proposal assumes a different feature geometry). A palatalized or labialized consonant preceding a V2#V2 sequence might produce some of the changes in the vowel transitions measured in this experiment.



However, while palatalization and labialization are common processes, it would be unusual for all of the vowel features to spread onto the consonant, as would be required to express Igbo vowel assimilation. In a partially-assimilated /*a*#*o*/ sequence, for example, higher formants are found than in an /*o*#*o*/ sequence. To produce these higher formants, unrounding and lowness would have to spread onto the consonant. These consonants are not pharyngealized in the sense in which the term is usually used.

Finally, in perhaps the most plausible feature-spreading approach to Igbo vowel assimilation, a phonologist might try to salvage the vowel features of V1 by formulating assimilation as the sharing of place features between V2 and V1 without delinking the features of V1. Such a rule is shown in 23.

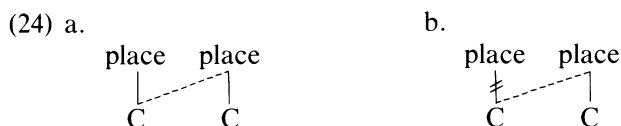
(23) Spreading without delinking



One significant drawback to this kind of representation is that it requires some specific and otherwise unmotivated phonetic interpretation rules. The problems inherent in using spreading without delinking to formalize partial assimilation are directly addressed by Nolan (1992).

Nolan 1992 presents articulatory evidence that assimilation in English of word-final alveolars to a following stop is a gradient process: a remnant of the alveolar gesture may persist. He discusses a possible formalization as in 24a, where the place node of the second consonant spreads to the first, without

causing delinking of its original place node, as in the complete assimilation shown in 24b. The distinction is the same as that between 21 and 23 above.



Nolan rejects the rule in 24a, however, on the grounds that it would require an unmotivated implementation: in order to account for the preception of assimilation, the place node of the alveolar would have to be interpreted as weaker than the place node of the following consonant. He argues that 'association lines do not come in different strengths' and that 'there is no independently justified principle' that would give priority to one place node over another (1992:277). Hayes (1992:282, commenting on Nolan's paper) concurs. Noting that 'standard phonological assumptions' would interpret the representation in 24a as a contour, Hayes argues that this is a 'qualitatively incorrect' representation for gradient assimilation.

Further, Nolan argues that there is no reason to force a rule such as 24a into the phonology. A simpler and more constrained explanation of the assimilation of word-final alveolars follows from the dynamics of articulation. Experimental data on the movements of articulators in English show a significant amount of overlap between the periods of closure for adjacent word-final and word-initial stops, which may increase in casual or fast speech (Browman & Goldstein 1990; see also Zsiga & Byrd 1990, Byrd 1992, 1994, Zsiga 1994, 1995). The acoustic consequences of overlap (hiding of the alveolar gesture),¹⁸ together with the tendency demonstrated by Nolan for word-final coronal stops to be partially reduced, can explain why word-final coronals are so often perceived as deleted or assimilated. The cases of coronal assimilation might show no greater temporal overlap than other consonant sequences: the perception of assimilation could result from the particular acoustic consequences of the vocal tract configuration. Given that a statement of the timing relations between adjacent stop consonants in casual speech will be an independently needed principle of English articulation, no additional phonological rule of assimilation is necessary.

The same reasoning applies to Igbo vowel assimilation. The rule in 23 requires a phonetic implementation different from the usual interpretation of a phonological contour. While 23 expresses the fact that the V2 articulation imposes itself on V1, acoustically, the result of assimilation is not a contour. Assimilation does not result in V1 followed by V2 within a single timing slot, as a contour tone consists of a high tone followed by a low, but in formants intermediate between V1 and V2. Given that the place features are on the same tier, they are presumed to be ordered and thus would be implemented sequentially. Even if one were to invoke spreading of individual articulator features, a feature-

¹⁸ Byrd (1992), using a gestural model of speech synthesis, has shown that, given the same degree of overlap, a /d#b/ sequence will be perceived as assimilated, while a /b#d/ sequence will not. Byrd (1994) finds the same asymmetry for /d#k/ and /k#d/.

spreading approach to Igbo vowel assimilation cannot avoid creating a contour, because the same features are involved for both vowels.

The rule in 23 complicates not only the interpretation of phonological structures, but also phonological theory itself, because of the way its domain of application must be specified. The formalization in 23, like those in 21 and 22, indicates a domain juncture rule: a rule that applies across a boundary between constituents, within the domain of a larger constituent (Selkirk 1980, Nespor & Vogel 1982, 1986). These domains must be specified in order to capture the fact that vowel assimilation applies only at word boundaries, within smaller syntactic constituents, as described above in §3.1.

Processes described as JUNCTURE RULES have proven problematic in phonological theory. Such rules require reference to two levels of structure at the same time, but few juncture rules have been proposed, and those that have (such as vowel degemination in Greek, Sanskrit final voicing, American English final-t deletion, and Italian raddoppiamento sintattico, see Vogel 1984, 1987), apply only at the boundaries of phonological words (as does Igbo vowel assimilation), never at the boundaries of any other phonological domain.¹⁹ Nolan (1992) reanalyzes final-t deletion as a phonetic process, not a phonological rule. Rice (1990) and Selkirk (1986) argue that all juncture rules can be reanalyzed phonologically and this rule type eliminated. Selkirk argues that the 'phonological effects [of juncture rules] derive from the status of words in the word phonology, and not from their juncture in the sentence phonology' (1986:403). Rice (1990) suggests that juncture rules can be eliminated by extending the strict cycle condition (Kiparsky 1985) into the post-lexical phonology, and she offers a reanalysis of Sanskrit final voicing along those lines. For either theory, Igbo vowel assimilation would pose a problem if it were to be analyzed as in the autosegmental rules above (21, 22, or 23), making reference to both the phonological word and phonological phrase domains.

Linking vowel assimilation to word-final reduction in the gestural approach, however, solves the problem of describing vowel assimilation as a domain juncture rule, applying only across word boundaries. The problem is solved not by eliminating reference to domain boundaries, but by eliminating the need for a rule of assimilation at all. The amount of overlap between vowels or consonant at word edges is a language-specific property (see Smith 1992, Zsiga & Kaufmann 1996), and a statement of timing relations between segments at word boundaries is independently needed by the phonetics in order to describe patterns of articulation in any language. There is no need to restate them in the phonology. Under the account proposed here, if any rule applies in Igbo vowel sequences, it is vowel reduction, a process that occurs at a domain edge, the end of a phonological word. Assimilation is not an independent process, but it

¹⁹ Again, the problem of juncture rules does not go away if the alternation is recast in terms of constraints. A NONALIGNMENT constraint could be imposed here, requiring that a [place] feature associated to a vowel NOT be aligned with the right edge of a word. A higher-ranked constraint, aligning a place feature with a phonological-phrase edge, might account for the apparent sensitivity to two different domains. It does not, however, solve the problem of why constraints sensitive to only these two domains should interact in this way.

follows from reduction, as originally specified temporal relations are maintained among the gestures, even though one gesture has been reduced in time. Assimilation then requires no formal statement that involves the juncture of two domains.²⁰

It has been shown in this section that featural representations for Igbo vowel assimilation are empirically inadequate, and that they lead to complications both in the phonology and in the phonetic interpretation of phonological representations. The gestural account, however, captures the gradient nature of the assimilation, links vowel assimilation to word-final reduction, and avoids formulating vowel assimilation as a juncture rule. Just as Nolan (1992) argued that the assimilation of final alveolars follows from general principles of consonant timing at word boundaries in English, I argue here that vowel assimilation results from general principles of vowel timing at word boundaries in Igbo. In neither case is it necessary to force a featural rule into the phonology. When a straightforward and principled phonetic explanation exists, proposing phonological rules for gradient processes unnecessarily complicates the phonology. The phonological component becomes more powerful and less constrained when the phonological repertoire must be expanded to account for processes that properly lie outside its domain. In contrast, reserving autosegmental representations for categorical alternations 'removes from the phonology the burden of representing those sound patterns for which its natural descriptive mechanisms are inappropriate' (Lieberman & Pierrehumbert 1984:229).

It is important to note, however, that even though a feature spreading representation fails for Igbo vowel assimilation, in general a featural representation of post-lexical vowel assimilation is not ruled out as a possible phonological rule. The mapping from features to gestures does not necessarily take place at the end of the lexical phonology. Some rules of the phrasal phonology may well be categorical, in which case they would best be represented in terms of spreading or delinking of autosegmental features, even across word boundaries. Other African languages, such as Yoruba (Bamgbose 1965), Kikuyu (Armstrong 1967), and Edo (Welmers 1973), have rules of vowel assimilation that have been reported as categorical. Should their categorical application hold up (and the findings here suggest that it would be worthwhile to reexamine such processes) a featural representation would be required. (If such rules applied across the board, as is apparently the case in the languages named above, or only within a domain or at a domain edge, they would not post difficulties as juncture rules, either.) Careful investigation is required to determine whether a phrasal process is categorical or gradient, and thus whether it should be assigned to the featural or gestural component.

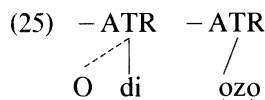
The range of processes properly captured by either a featural description

²⁰ Clearly, eliminating the need to formalize Igbo vowel assimilation as a domain juncture rule does not completely solve the problem of juncture rules for all of phonological theory. Clark (1990), for example, argues that several tonal rules in Igbo must be considered juncture rules. The analysis presented here does, however, eliminate one case from this problematic category, and it suggests that it would be worth investigating other juncture rules from a gestural point of view.

or a gestural description is limited. Gradient processes such as Igbo vowel assimilation require reference to the specific temporal information that the gestural approach provides. Categorical alternations such as Igbo vowel harmony neither need nor benefit from detailed timing information. Thus two distinct representations are required. Their similarities are tantalizing, but they can not be collapsed. Section 4 takes up the question of how these two distinct representations may be related.

4. FEATURES TO GESTURES: *ede ət̩*. In this section I present a complete derivation of a sequence of Igbo vowels, taking the word pair *ede ət̩* as an example, from the underlying featural representation to the articulatory instantiation. As discussed in §1, featural and gestural representations are in many ways homologous. Both representations are based on articulators; the crucial difference between them is that gestures have inherent quantitative specifications. If underspecification or derivation is assumed, then temporal specification can be seen as the final step in the filling in of noncontrastive information. If no underspecification or derivation of phonological features is assumed, the mapping process from features to gestures would begin with a fully specified phonological representation.

In an approach that assumes underspecification, harmonic affixes would have no [ATR] value in the underlying representation, but would receive the appropriate value either by spreading from the root or by the application of default rules. While a noun phrase like *ede ət̩* will not exhibit harmonic alternations, in the case of a verb phrase like *əd̩ ət̩* 'there is another', [-ATR] spreads from the verb *d̩* to the underspecified prefix *ə/ə*, as in 25.²¹



Other predictable feature values, such as those for [+/- back], would be supplied by default rules. In a nonderivational approach, the fully specified representation with appropriate values for [ATR] and [back] would be chosen as the form that best satisfies the constraints of the language.

Once all the vowels have been fully specified, the vowel features must be related directly to articulatory gestures and the parameters that characterize them. Although the best phonological representation for vowels remains a matter of debate (Clements 1991, Ní Chiosáin 1991, Hume 1990, 1992), the mapping from features to gestures is most straightforward for those feature geometries (Sagey 1986, McCarthy 1988, Padgett 1991, Keyser & Stevens 1994) that most closely approximate the anatomy of the vocal tract. Taking these geometries

²¹ Clark (1990) argues that the initial vowels in nouns are also prefixes. If that is the case, spreading may be invoked for nouns as well. Both phonological and phonetic evidence point to [-ATR] as the marked value in [ATR] harmony systems (Pulleyblank 1988, Welmers 1973). [-ATR] vowels have an additional constriction in the pharynx, [+ATR] vowels have no additional gesture.

as a starting point, there is a nearly transparent mapping between terminal features and gestures.

The relation between features and articulators can be argued to be universal. Thus, [+round] is always interpreted as a labial gesture, and [−ATR] is always interpreted as a tongue root constriction gesture (constricted pharynx). The features dependent from the dorsal node specify constriction location and degree for the tongue body gesture. The feature [+low] means pharyngeal constriction location, and [−back] means palatal constriction location. The combination [+back, +high] translates into velar constriction location and narrow constriction degree, while [+back, −high] translates into uvular constriction location and narrow constriction degree. Wood (1982), on the basis of x-ray evidence from a number of languages, argues that velar and uvular best describe the place of articulation of /u/ and /o/, respectively. The exact specification of place and degree of constriction will be language specific.

The gestural specification of an /e#a/ sequence, as in *ede atə*, is shown in Figures 7 and 8. Fig. 7A illustrates a fully specified autosegmental representation. (For clarity, laryngeal features for voicing and tone are not shown, although they are assumed to be present). Fig. 7B shows how these features can be interpreted as gestures, with specified constriction location, constriction degree, and with specified internal durations and temporal relations to other gestures.²² Fig. 8A shows the gestural pattern without the phonological structure of root and class nodes and association lines. Gestural phasing is shown in Fig. 8B.

The feature-to-gesture correspondence is specified by the universal and language-specific principles discussed above. The duration of the gestures shown in Fig. 8B and their temporal overlap relations are specified by language-specific rules or constraints. (It would be interesting to consider gestural phasing as a kind of phonetic alignment constraint.) The movements of the articulators unfold over time as determined by the variables in the task-dynamic equations that describe each gesture.

Temporal relations among gestures are described by language-specific phasing principles (see Browman & Goldstein 1990). These principles are stated in terms of which points in the trajectories of two different gestures occur simultaneously, as shown in 26.

- (26) a. C target = V onset (begin a vowel gesture when the preceding consonant has reached its target)
- b. V1 release = V2 onset (begin the second vowel gesture in a sequence as the first begins moving away from its target position)

Points such as onset, target, and release are specified as numerical values in the 360-degree cycle. Onset, for example, equals 0°, and release of target approximately 330°. Not every point will be available for manipulation, but only certain landmarks, such as onset of movement or the achievement of target (Browman & Goldstein 1992).

²² This diagram, and others that follow, is meant to be illustrative, and not a precise gestural description.

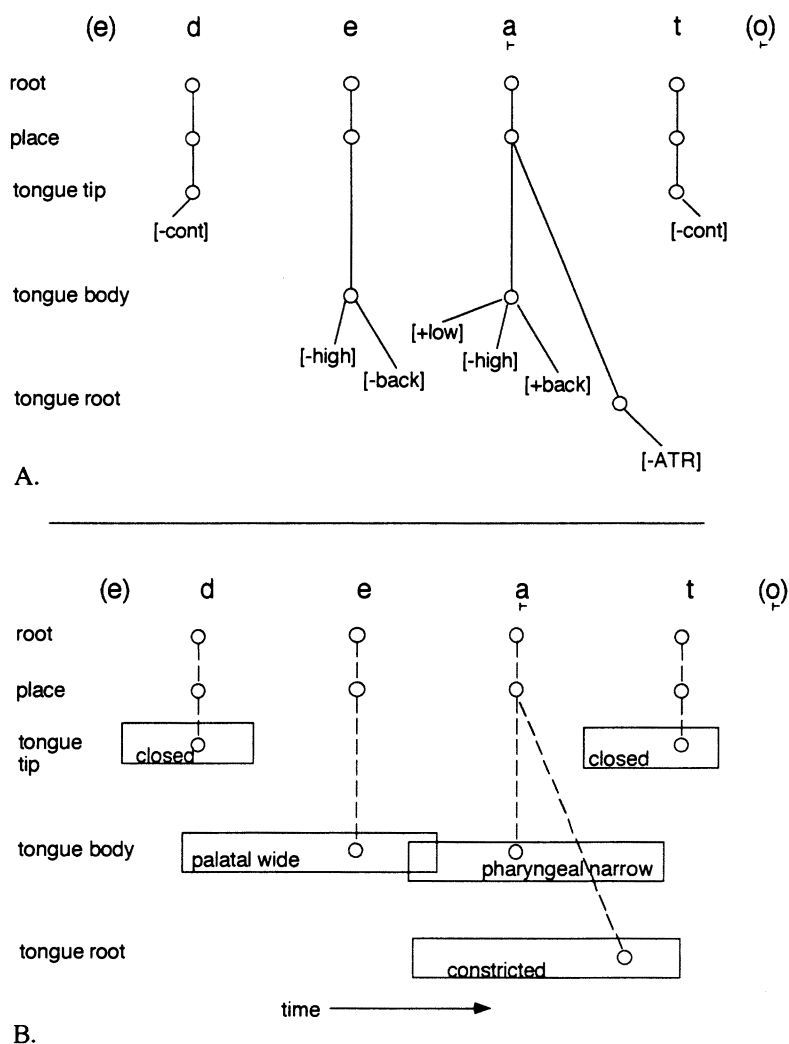


FIGURE 7. Representation of an /e#a/ sequence. A. Fully-specified featural representation. B. Correspondence of features to gestures.

It will be assumed here that phasing among articulators works outward from the smallest domain to the largest. First, a given oral gesture is phased with respect to its associated laryngeal or velic gestures. In the simplest case, the group of features associated to a single root node (such as [cor] and [nasal] for /n/, or [lab] and [asp] for /p^h/) corresponds to a CONSTELLATION of gestures in the sense of Browman and Goldstein (1986), a set of gestures that are phased to one another and that form a unit. The several oral gestures that make up a vowel may also form a constellation. If root nodes match gestural constella-

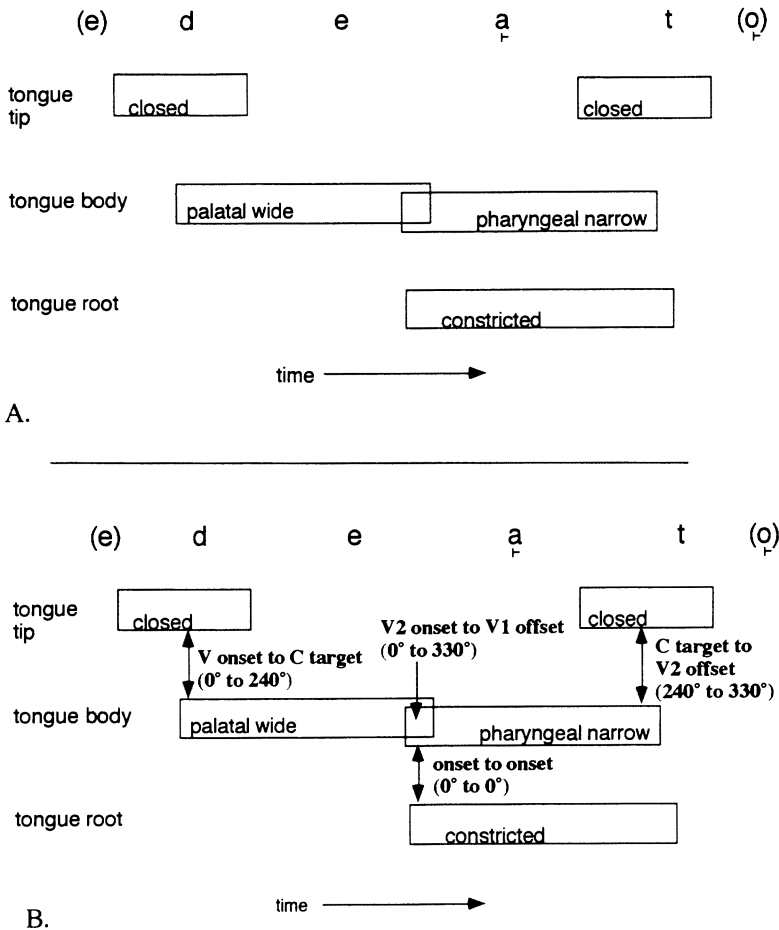


FIGURE 8. Representation of an /e#a/ sequence. A. Gestural representation. B. Gestural representation with phasing indicated.

tions, phasing rules can be seen as supplying a specific temporal relation for each association line that links an articulator feature to a root node.²³

In more complicated cases, such as strings of consonants in an onset, the differences between autosegmental and articulatory phonology over the usefulness of assuming a segmental organization come into play, and the root nodes posited in the former theory will not match the gestural constellations posited

²³ Actually, given the absence of place and root nodes in articulatory phonology, an extra step would be required. The group of features linked to a root node must first be interpreted as a constellation, where gestures are associated directly to one another rather than to a shared root node. It is the associations in the constellation that are then supplied with specific phasings. Note that the contrast between unspecified association lines, and specified temporal relations can still be maintained without root nodes, by assuming that articulators are associated directly to one another at all points in the derivation.

in the latter theory. (Browman & Goldstein 1990 states that gestural constellations will sometimes correspond to segments, sometimes to syllables, sometimes to intermediate constituents like onsets.) It is assumed here that segments (root nodes) are relevant for phonological alternations, but cease to be relevant when features are mapped into gestures. While the debate over segmental organization cannot be resolved here, in order to make the mapping explicit, it will be assumed that the phasing in more complicated cases operates in the same way as in the simpler cases, forming gestural units for each root node. (These will not necessarily correspond to the constellations of articulatory phonology.) Later phasing specifications may then alter this segment-based organization, so that segmental boundaries are blurred. In the case of multiply linked features, as in harmony systems, the segmental boundaries are already blurred in the autosegmental representation. Here, the single feature can be interpreted as a single underlying gesture that persists throughout the domain. Its onset is phased to the first oral gesture in the domain, its offset to the last. The assumption that root nodes map into gestural units is intended here as no more than a working hypothesis, awaiting further study of the gestural mechanics of harmony systems and syllable structure. It is likely that other hierarchical structures also serve as a locus for the specification of phasing relations.

Once the simpler gestural constellations are formed, timing of oral gestures within a syllable would follow, according to language-specific principles such as 26a above. Vowel and consonant gestures within syllables are tightly coordinated, and their coordination is predictable (Browman & Goldstein 1990). Finally, timing across syllables and across words is specified. A vowel in one syllable may be coordinated with either the vowel or consonant of another syllable, according to language-specific principles (Smith 1992). It seems that the timing of gestures across larger phonological domains (such as between words) is freer and more variable than the timing found within syllables. This variable timing then leads to the many casual speech processes that are found to apply at word boundaries. It may be that for domains larger than the phonological word, there can be no overlap between gestures at the domain boundary. This would account for the lack of juncture rules referring to domains larger than the phonological word. The gestural phasing may specify that no interaction is possible, because no gestural overlap is allowed. Thus casual speech variation due to gestural overlap may be limited to only certain domains, such as between phonological words, where overlap is allowed, but not exactly specified.

The hypothesized phasings among gestures in the sequence *ede atɔ* (based on Browman & Goldstein 1990, 1992) are shown in Fig. 8B. This phasing corresponds to that in the gestural score in Fig. 5, which illustrates a gestural organization that would result in no assimilation. A word-level rule in Igbo reduces the duration of the word-final vowel, as diagrammed in Fig. 6. As all other temporal relations remain fixed (crucially, V2 onset to V1 offset), the V2 gesture expands to fill the time span required.

Thus the gestural pattern, consistent with the acoustic effects of Igbo vowel assimilation, can be derived directly from autosegmental features. Principles of correspondence between features and gestures, along with the language-

specific specification of overlap relations, map features into a gestural pattern. Each gesture is specified for constriction location, constriction degree, stiffness, and phasing. The gestural organization then controls the movements of articulators to produce an utterance.

In conclusion, I have demonstrated the need for both qualitative (featural) and quantitative (gestural) representations. Categorical alternations, such as Igbo vowel harmony, are best represented featurally. Gradient processes, such as Igbo vowel assimilation, are best represented gesturally. Neither representation can account for the full range of data. Representations that have access to specific quantitative information cannot constrain categorical changes. Representations that do not have access to specific quantitative information cannot capture gradient changes. The close correspondence between features and gestures, however, provides a straightforward articulatory implementation of the phonological representation.

REFERENCES

- ARCHANGELI, DIANA, and DOUGLAS PULLEYBLANK. 1994. *Grounded phonology*. Cambridge, MA: MIT Press.
- ARMSTRONG, LILIAS E. 1967. *The phonetic and tonal structure of Kikuyu*. London: International African Institute.
- BAMGBOSE, AYO. 1965. Assimilation and contraction in Yoruba. *Journal of West African Languages* 2.21–27.
- BECKMAN, MARY, and JAN EDWARDS. 1990. Lengthenings and shortenings and the nature of prosodic constituency. In Kingston & Beckman, 152–78.
- BROWMAN, CATHERINE P., and LOUIS GOLDSTEIN. 1986. Towards an articulatory phonology. *Phonology Yearbook* 3.219–52.
- , —. 1989a. Articulatory gestures as phonological units. *Phonology* 6.201–51.
- , —. 1989b. Some notes on syllable structure in articulatory phonology. *Phonetica* 45.140–55.
- , —. 1990. Tiers in articulatory phonology, with some implications for casual speech. In Kingston & Beckman, 341–76.
- , —. 1991. Gestural structures: Distinctiveness, phonological processes, and historical change. *Modularity and the motor theory of speech perception*, ed. by Ignatius Mattingly and Michael Studdert-Kennedy, 313–38. Hillsdale, NJ: Lawrence Erlbaum.
- , —. 1992. Articulatory phonology: An overview. *Phonetica* 49.55–180.
- BYRD, DANI. 1992. Perception of assimilation in consonant clusters: A gestural model. *Phonetica* 49.1–24.
- . 1994. Articulatory timing in English consonant sequences. (UCLA Working Papers in Phonetics 86.) Los Angeles: UCLA Dept. of Linguistics.
- CHOI, JOHN, and PATRICIA A. KEATING. 1991. Vowel-to-vowel coarticulation in Slavic languages. *UCLA Working Papers in Phonetics* 78.78–86.
- CHOMSKY, NOAM, and MORRIS HALLE. 1968. *The sound pattern of English*. New York: Harper and Row.
- CLARK, MARY M. 1990. *The tonal system of Igbo*. Dordrecht: Foris.
- CLEMENTS, GEORGE N. 1985. The geometry of phonological features. *Phonology Yearbook* 2.225–52.
- . 1991. Place of articulation in consonants and vowels: A unified theory. *Working papers of the Cornell Phonetics Laboratory* 5.77–123.
- . 1992. Phonological primes: Gestures or features? *Phonetica* 49.181–93.

- COHN, ABIGAIL C. 1990. Phonetic and phonological rules of nasalization. (UCLA Working Papers in Phonetics 76.) Los Angeles: UCLA Dept. of Linguistics.
- COLE, JENNIFER, and CHARLES KISSEBERTH. 1994. Paradoxical strength conditions in harmony systems. Paper presented at the Twenty-fifth annual meeting of the North Eastern Linguistic Society, University of Pennsylvania.
- DOCHERTY, GERARD and D. ROBERT LADD (eds.) 1992. *Papers in laboratory phonology 2: Gesture, segment, prosody*. London: Cambridge University Press.
- EMENANJO E. NOLUE. 1978. *Elements of modern Igbo grammar*. Ibadan, Nigeria: Oxford University Press.
- GIMSON, A. C. 1962. *An introduction to the pronunciation of English*. London: Edward Arnold.
- GOLDSMITH, JOHN. 1976. *Autosegmental phonology*. Cambridge, MA: MIT dissertation.
- GRACCO, VINCENT L., and JAMES H. ABBS. 1988. Central patterning of speech movements. *Experimental Brain Research* 71.515–26.
- GREEN, M. M., and G. E. IGWE. 1963. *A descriptive grammar of Igbo*. London: Oxford University Press.
- HAWKINS, SARAH. 1992. An introduction to task dynamics. In Docherty & Ladd, 9–25.
- HAYES, BRUCE. 1989. Compensatory lengthening in moraic phonology. *Linguistic Inquiry* 20.253–306.
- . 1992. Comments on chapter 10. In Docherty & Ladd, 280–86.
- HOFFMAN, CARL. 1963. *A grammar of the Margi language*. London: Oxford University Press.
- HUME, ELIZABETH. 1990. Front vowels, palatal consonants, and the rule of umlaut in Korean. *North East Linguistics Society* 20.230–43.
- . 1992. *Front vowels, coronal consonants, and their interaction in non-linear phonology*. Ithaca, NY: Cornell University dissertation.
- HYMAN, LARRY M. 1985. *A theory of phonological weight*. Dordrecht: Foris.
- JAKOBSON, ROMAN; GUNNAR M. FANT; and MORRIS HALLE. 1951. *Preliminaries to speech analysis: The distinctive features and their correlates*. Cambridge, MA: MIT Press.
- KEATING, PATRICIA A. 1988. The phonology-phonetics interface. *Linguistics: The Cambridge survey*, vol. 1, *Linguistic Theory: Foundations*, ed. by F. J. Newmeyer, 281–302. Cambridge: Cambridge University Press.
- . 1990a. Phonetic representations in a generative grammar. *Journal of Phonetics* 18.321–34.
- . 1990b. The window model of coarticulation: Articulatory evidence. In Kingston & Beckman, 451–75.
- KELSO, J. A. S.; BETTY TULLER; ERIC BATESON; and CAROL FOWLER. 1984. Functionally specific articulatory cooperation following jaw perturbations during speech: Evidence for coordinative structures. *Journal of Experimental Psychology: Human Perception and Performance* 10.812–32.
- KEYSER, SAMUEL J., and KENNETH N. STEVENS. 1994. Feature geometry and the vocal tract. *Phonology* 11.207–36.
- KINGSTON, JOHN, and MARY BECKMAN (eds.) 1990. *Papers in laboratory phonology 1: Between the grammar and physics of speech*. Cambridge: Cambridge University Press.
- , and AVIS H. COHEN. 1992. Extending Articulatory Phonology. *Phonetica* 49.194–204.
- , and RANDY L. DIEHL. 1994. Phonetic knowledge. *Language* 70.419–54.
- KIPARSKY, PAUL. 1985. Some consequences of Lexical Phonology. *Phonology Yearbook* 2.85–138.
- KOLLIA, H. BETTY; VINCENT L. GRACCO; and KATHERINE S. HARRIS. 1992. Functional organization of velar movements following jaw perturbation. *Journal of the Acoustical Society of America* 91.2474 (A).

- KRAKOW, RENA A. 1989. The articulatory organization of syllables: A kinematic analysis of labial and velic gestures. New Haven, CT: Yale University dissertation.
- LADEFOGED, PETER. 1968. A phonetic study of West African languages. Cambridge: Cambridge University Press.
- , IAN MADDIESON. 1990. Vowels of the world's languages. *Journal of Phonetics* 18.93–122.
- , ———. 1996. Sounds of the world's language, chapter 10: Multiple articulatory gestures. Oxford: Blackwell.
- ; KAY WILLIAMSON; BEN ELUGBE; and SISTER ANN ANGELA UWALAKA. 1976. The stops of Owerri Igbo. *Studies in African Linguistics*, supplement 6. 147–63.
- LEBEN, WILLIAM. 1978. The representation of tone. *Tone: A linguistic survey*, ed. by Victoria Fromkin, 177–219. New York: Academic Press.
- LIBERMAN, MARK, and JANET B. PIERREHUMBERT. 1984. Intonational invariance under changes in pitch range and length. *Language sound structure*, ed. by Mark Aronoff and Richard T. Oehrle, 157–233. Cambridge, MA: MIT Press.
- ; J. M. SCHULTZ; S. HONG; and V. OKEKE. 1993. The phonetic implementation of tone in Igbo. *Phonetica* 50.147–60.
- MANUEL, SHARON Y., and RENA A. KRAKOW. 1984. Universal and language-particular aspects of vowel-to-vowel coarticulation. *Haskins Laboratories Status Report on Speech Research SR77/78.69–78*.
- MCCARTHY, JOHN J. 1981. A prosodic theory of non-concatenative morphology. *Linguistic Inquiry* 12.373–418.
- . 1988. Feature geometry and dependency: A review. *Phonetica* 45.84–108.
- MCMAHON, APRIL; PAUL FOULKES; and LAURA TOLLFREE. 1994. Gestural representations and Lexical Phonology. *Phonology* 11.277–316.
- MOHANAN, K. P. 1986. The theory of lexical phonology. Dordrecht: Reidel.
- NESPOR, MARINA, and IRENE VOGEL. 1982. Prosodic domains for external sandhi rules. *The structure of phonological representations*, ed. by Harry van der Hulst and Norval Smith, 225–55. Dordrecht: Foris.
- , ———. 1986. Prosodic phonology. Dordrecht: Foris.
- NÍ CHIOSÁIN, MAIRE. 1991. Topics in the phonology of Irish. Amherst, MA: University of Massachusetts dissertation.
- NOLAN, FRANCIS. 1992. The descriptive role of segments: Evidence from assimilation. In Docherty & Ladd, 261–80.
- ; TARA HOLST; and BARBARA KÜHNERT. 1996. Modelling [s] to [ʃ] assimilation in English. *Journal of Phonetics* 24.113–37.
- OHALA, JOHN J. 1981. The listener as a source of sound change. *Proceedings of the Chicago Linguistics Society, Papers from the Parasession on Language Behavior* 17.178–203.
- ÖHMAN, S. E. G. 1966. Coarticulation in VCV utterances: Spectrographic measurements. *Journal of the Acoustical Society of America* 30.149–54.
- PADGETT, JAYE. 1991. Stricture in feature geometry. Amherst, MA: University of Massachusetts dissertation.
- . 1995. Partial class behavior and nasal place assimilation. Santa Cruz, CA: University of California, MS.
- PIERREHUMBERT, JANET B. 1990. Phonological and phonetic representation. *Journal of Phonetics* 18.375–94.
- PRINCE, ALAN, and PAUL SMOLENSKY. 1993. Optimality theory: Constraint interaction in generative grammar. New Brunswick, NJ: Rutgers University, MS.
- PULLEYBLANK, DOUGLAS. 1988. Vocalic underspecification in Yoruba. *Linguistic Inquiry* 19.233–70.
- RICE, KEREN. 1990. Predicting rule domains in the phrasal phonology. *The phonology-syntax connection*, ed. by Sharon Inkelas and Draga Zec, 289–312. Chicago: University of Chicago Press.

- SAGEY, ELIZABETH C. 1986. The representation of features and relations in nonlinear phonology. Cambridge, MA: MIT dissertation.
- . 1988. On the ill-formedness of crossing association lines. *Linguistic Inquiry* 19.109–18.
- SALTZMAN ELLIOT. 1986. Task dynamic coordination of the speech articulators: A preliminary model. *Experimental brain research series* 15, ed. by H. Heuer and C. Fromm, 129–44. New York: Springer Verlag.
- , and J. A. S. KELSO. 1987. Skilled actions: A task dynamic approach. *Psychological Review* 94.84–106.
- , and KEVIN G. MUNHALL. 1989. A dynamical approach to gestural patterning in speech production. *Ecological Psychology* 1.333–82.
- SELKIRK, ELISABETH. 1980. Prosodic domains in phonology: Sanskrit revisited. *Juncture*, ed. by Mark Aronoff and Mary-Louise Kean, 107–29. Saratoga, CA: Anma Libri.
- . 1986. On derived domains in sentence phonology. *Phonology Yearbook* 3:371–405.
- SHAIMAN, SUSAN. 1989. Kinematic and electromyographic responses to perturbation of the jaw. *Journal of the Acoustical Society of America* 86.78–88.
- SMITH, CAROLINE L. 1992. The timing of vowel and consonant gestures. New Haven, CT: Yale University dissertation.
- STERIADE, DONCA. 1990. Gestures and autosegments: Comments on Browman and Goldstein's paper. In Kingston & Beckman, 382–97.
- VOGEL, IRENE. 1984. On constraining prosodic rules. *Advances in nonlinear phonology*, ed. by Harry van der Hulst and Norval Smith, 217–33. Dordrecht: Foris.
- . 1987. Vowel degemination and fast speech rules. *Phonology Yearbook* 4.61–85.
- WELMERS, WILLIAM E. 1973. *African language structures*. Los Angeles: University of California Press.
- WETZELS, LEO, and ENGIN SEZER. 1986. *Studies in compensatory lengthening*. Dordrecht: Foris.
- WILLIAMSON, KAY. 1972. *Igbo-English dictionary*. Benin City, Nigeria: Ethiopia Publishing.
- WOOD, SIDNEY. 1982. X-ray and model studies of vowel articulation. *Lund University Department of Linguistics and Phonetics Working Papers* 23.1–192.
- , and THORE PETTERSSON. 1988. Vowel reduction in Bulgarian: The phonetic data and model experiments. *Folia Linguistica* 22.239–67.
- ZSIGA, ELIZABETH C. 1992. A mismatch between morphological and prosodic domains: Evidence from two Igbo rules. *Phonology* 9.101–35.
- . 1993. Features, gestures, and the temporal aspects of phonological organization. New Haven, CT: Yale University dissertation.
- . 1994. Acoustic evidence for gestural overlap in consonant sequences. *Journal of Phonetics* 22.121–40.
- . 1995. An acoustic and electropalatographic study of lexical and post-lexical palatalization in American English. *Papers in laboratory phonology 4: Phonology and phonetic evidence*, ed. by Bruce Connell and Amalia Arvaniti, 282–302. Cambridge: Cambridge University Press.
- , and DANI BYRD. 1990. Acoustic evidence for gestural overlap in consonant sequences. *Journal of the Acoustical Society of America* 88.S82.
- , and STEFAN KAUFMANN. 1996. Palatalization and gestural overlap in Russian and English. Poster presented at the Fifth Conference on Laboratory Phonology, Chicago.