Issues in the Phonology-Phonetics Interface in African Languages

Douglas Pulleyblank and Blake Allen
University of British Columbia

1. Introduction

In this paper, we examine a range of ways that the study of phonetics informs our understanding of phonological patterns, with attention paid specifically to African languages. We start by touching on some issues so basic that they seem virtually definitional: for example, the fact that phonological transcriptions are based on phonetically motivated categories. Typical work on both African and non-African languages begins with a transcription of primary data, typically in some version of the International Phonetic Alphabet (IPA). As Ladefoged & Halle (1988) phrase it, "The IPA alphabet is defined in terms of a series of charts. With their column and row names, together with accompanying notes and conventions, the whole series forms a theory of phonetic description." (Ladefoged & Halle, 1988:577-578). In effect, phonetic theory has begun to inform our analyses before we even consider the analysis to have begun.

Ladefoged & Halle also suggest that "the most fundamental insight gained during the last century has been the realization that it is the features rather than the sounds that are the basic building blocks of spoken language" (Ladefoged & Halle, 1988:577-578). Two notions are central here. First, "natural" classes, those sets of sounds which function as units in the phonologies of natural languages, are based on phonetically motivated classifications of speech sounds. Second, theories of distinctive features are based on phonetically motivated classes. As Clements & Hume (1995) describe, features "are defined in terms of specific patterns of acoustic and articulatory realization which provide the crucial link between the cognitive representation of speech and its physical manifestation" (Clements & Hume, 1995:245).

The raw material on which much, perhaps most, linguistic work is conducted is the phonetic content of utterances. Such material consists of observations about the production of sounds, their physical properties, as well as how they are perceived. Since these data constitute the frame over which phonological analyses in particular are constructed, their accuracy is extremely important to phonologists. Crucial in this regard is the technological shift that has taken place over the last few years. Fifty years ago, the primary instruments for recording speech data were the ear and the pen or pencil. Recording equipment became increasingly common, but the initial utility of recordings was to allow the ear multiple opportunities to "get it right." The use of instruments for speech analysis was restricted to specialized speech laboratories, where expensive equipment was employed. Such equipment also imposed severe limits on the kinds of phonetic investigation that was possible. For example, one of the early work horses of acoustic analysis was the Kay Sona-Graph Sound Spectrograph (Fry, 1955; Lindblom, 1962; Hazen, 1973). This technology, available from the 1950's, was widely used for speech analysis[1]. While offering novel possibilities for speech analysis, the device had significant limitations. In particular, most models could only analyze samples of a maximum of 2.4 seconds, they were not terribly portable, required specialized paper, and so on.

This situation has changed drastically in recent years. Acoustic analysis is now possible for anyone with a computer and a connection to the internet. Software such as Praat (Boersma, 2001; Boersma & Weenink, 2012) can be downloaded for free, allowing easy access to a sophisticated suite of speech

* This work was supported by a Standard Research Grant from the Social Sciences & Humanities Research Council of Canada to the first author.

1 For example, it was the basic instrument for acoustic analysis at the University of Ibadan when the first author was a student there in the 1970's.

analysis tools. The modern speech laboratory is therefore readily available to researchers both in a university context and in the field. Instrumental research, instead of being a specialized area for trained and funded university-based researchers, is now a tool that can be incorporated into nearly any linguist's research.

In spite of this ready availability, however, instrumental phonetic research appears to be under-utilized. For example, we conducted an informal survey of instrumental phonetic work on African languages in two journals, *Phonology* and *Journal of African Languages and Linguistics*, the former as a representative of a specialized phonology journal, the second as an Africanist journal. During the seventeen-year period from 1995 to 2011, there were only 20-30 instances of instrumental phonetic research on African languages, less than an average of one per year per journal. By far the most common topic was tone (phonology) involving F_0 analysis (phonetics). There were about half as many articles looking at harmony as tone, employing techniques such as formant analysis (common) and ultrasound (uncommon). Instrumental analyses of voicing were also found, along with a small number of studies involving intensity, articulator position tracking, segment durations, voice onset time, airflow/pressure studies, palatography and fibrescopy. Relevant phonology topics after harmony were, at one or two discussions per topic, syllabification, footing, contrastive length, laryngeal contrasts, voice quality and clicks. All in all, given the ready availability of the technology, we might expect to see more work examining phonological topics instrumentally.

What, then, can be done in this "machine-age" of phonological research? Apart from the simple answer—"more!"—we discuss in the following sections three basic uses of instrumentation: (i) checking what we hear or see; (ii) figuring out what's going on when something is hard to hear or see; and (iii) determining (potential) patterns that cannot be heard or seen.

2. Verifying Audible Distinctions

One important function of instrumentation is to provide an objective way of verifying the sorts of observations made on the basis of auditory perception. We illustrate with an example from tongue root harmony, chosen for its relevance to a somewhat less transparent case to be considered thereafter.

2.1. Ebira: an unsurprising phonology-phonetics mapping

The harmony system of Ebira (Benue-Congo, Nupoid; Nigeria) has been discussed in Orie (2003). Ebira has a nine-vowel inventory, with non-low vowels occurring in advanced/retracted pairs. As illustrated in (1), disregarding low vowels, advanced vowels only co-occur with other advanced vowels, while retracted vowels only co-occur with other retracted vowels. Retracted vowels are indicated by a subdot (e.g. *ẹ, ụ*).

(1) Tongue root harmony in Ebira

Advanced		Retracted	
òzè	*road*	ọhẹ́	*pillar of house*
òbó	*rope*	ọpọ́	*mask*
ècè	*wine*	ẹdọ	*antelope*
isú	*house rat*	ìkụ̀	*sickness*
uye	*meat*	ụ̀nọ	*cow*
ukere	*wooden door*	ụkọ́rọ	*work*
okuku	*imaginary being*	ẹcụ́kụ	*bone*

Auditory analysis and phonological patterning lead us to postulate a vowel system where both high vowels and mid vowels are distinguished by their tongue root values. That is, auditory and phonological patterning lead us to something like Table 1.

High	Advanced	i	u
	Retracted	ɪ	ʊ
Mid	Advanced	e	o
	Retracted	ɛ	ɔ
Low	Retracted	a	

Table 1: Ebira vowels

Reported acoustic data are consistent with this analysis: see Figure 1, adapted from Ladefoged & Maddieson (1996:305).

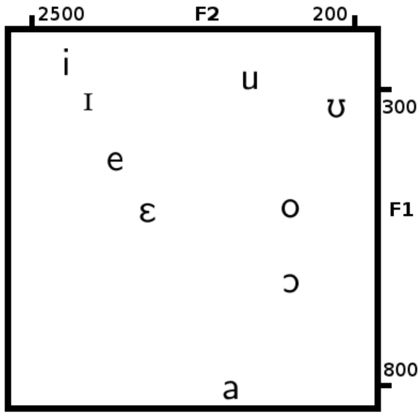

Figure 1: Ebira vowel formants (Hz)

In this case, the use of instrumental data illustrates two related points. First, well-trained linguists can provide excellent language data—a point that is not surprising but worth emphasizing. Second, phonetic experimentation can confirm auditory investigation.

Importantly, however, instrumental work often serves somewhat more interesting functions. Consider, for example, the phonological proposal in Clements (1990) that vowel height involves hierarchically nested values of a single vowel height feature.

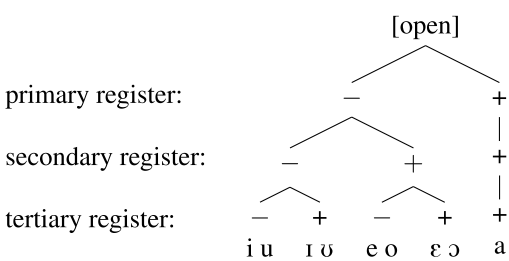

Figure 2: Hierarchical vowel height (Clements, 1990)

According to this proposal, a 'primary' division of the vowel height range is established by binary specifications of the feature [±open], with additional subdivisions of the same feature establishing additional vowel heights. According to this proposal, the height properties of the vowels of a language like Ebira could be captured by [±open] as specified in Figure 2.

This hypothesis contrasts crisply with the *interactive model* of vowel height, in which surface height results from the combined effect of multiple interacting features. According to the interactive model, the vowels of a language like Ebira could result from the interactions of three features: [±high], [±low] and [±ATR].

+high	i	u	+ATR	−low
	ɪ	ʊ	−ATR	
−high	e	o	+ATR	
	ɛ	ɔ	−ATR	
	a		−ATR	+low

Table 2: Interactive vowel height

While both the hierarchical model and the interactive model are consistent with the patterns observed in Ebira, we will see in the next section that the two models differ in their ability to account for other types of data.

2.2. Ịjọ: A phonology-phonetics mapping involving crossed heights

Initially, the harmony patterns of Ịjọ (Niger-Congo; Nigeria), described in Williamson (1969) and Akinlabi (1997), seem remarkably similar to those just seen for Ebira. Like Ebira, Ịjọ has nine vowels; like Ebira, setting aside the behavior of low vowels, we see that advanced vowels co-occur with advanced vowels, and retracted vowels co-occur with retracted vowels.

(2) Tongue root harmony in Ịjọ

Advanced		Retracted	
éré	*female*	ẹ́rẹ́	*name*
póló	*compound*	kọrọ	*rafia palm*
énéme	*oil palm*	ẹ́tẹ́ꜜrẹ́	*mat*
ḅuru	*yam*	ḅụrụ	*become rotten*
kíꜜrí	*ground*	fị́rị́	*work/message*
íkú	*louse*	ị́kụ́	*rock*
kúrúsu	*cannon*	ị́lụ́kụ́	*folk story*
númé	*song*	fẹ́rụ́	*wind*
poku	*bat*	dụ́rọ́	*barracuda*
ókúró	*shoe*	tọ́rụ́mꜜgbọ́	*eye*

Like Ebira, we might expect that an acoustic plot of the vowels of Ịjọ would be similar to Figure 1. In fact, as seen in Williamson (1969) and Ladefoged & Maddieson (1996), the vowel space for Ịjọ is quite different.

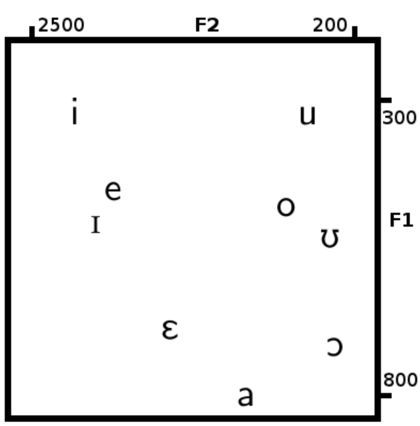

Figure 3: Ịjọ vowels, adapted from (Ladefoged & Maddieson, 1996:305)

Comparing the acoustic plots in Figure 3 with the phonological patterns of (2), we observe a mismatch between the apparent phonetic “height” of certain vowels and their phonological behavior. The vowels [e] and [o] behave phonologically as though they are advanced mid vowels while phonetically, they are higher in some sense than the “high retracted” vowels [ɪ] and [ʊ].

The observed mismatch is a problem for the hierarchical approach to height features but not for the interactive approach. The problem for the hierarchical model is that sub-heights are formally nested. One value of [open] establishes a particular range of heights, with sub-values of [open] defining particular parts of that defined range. Hence all “high” vowels (defined as [−open 1, −open 2]) should be higher (lower F1 values) than corresponding “mid” vowels (defined as [−open 1, +open 2]). That is, for vowels within the same register, any vowel that is [+open n] should not be acoustically higher than a corresponding vowel that is [−open n]. A vowel system with the properties seen in Ịjọ therefore constitutes an apparent counterexample to the hierarchical model of vowel height.

In contrast, the Ịjọ pattern is expected within an interactive model. If two distinct features have effects on F1, there is no *a priori* reason for one feature to have a stronger effect than the other. That is, given two interacting features like [high] and [ATR], one would not expect one feature or the other to have a greater impact on F1. Imagine that [high] has a stronger effect than [ATR]. This would lead to a pattern like Ebira where the [ATR] distinctions are nested with ranges established by [high]. Imagine in contrast that [ATR] has a stronger effect. The expectation now would be that the vowels would “cross.” That is, the tongue root effect is so strong that an advanced vowel like [e] ends up “higher” than a (weakly) high vowel like [ɪ]. The two possibilities are schematized in Table 3.

Strong [high]; weak [ATR] (Ebira)	Weak [high]; strong [ATR] (Ịjọ)
ɪ, ʊ: [+high] ⇑; [−ATR] ↓	ɪ, ʊ: [+high] ↑; [−ATR] ⇓
e, o: [−high] ⇓; [+ATR] ↑	e, o: [−high] ↓; [+ATR] ⇑

Table 3: Interactions between [high] and [ATR]

The interactive model makes an additional prediction that can again be checked instrumentally. If “[high]” and “[ATR]” were just different labels for articulatorily undifferentiated features affecting F1, then it would certainly be suspicious to think of these as distinct features. If, on the other hand, the two features have distinct articulatory correlates, then it is appropriate, indeed arguably required, to establish them as distinct. In this regard, the remarks in Jenewari (1973) concerning the Kalaḅarị and Okrika dialects of Ịjọ are significant: “[Orupabo’s] X-ray study of the vowels shows that vowel harmony in Okrika is conditioned by differences in pharynx width controlled by advancing the tongue root. By analogy, one could therefore suggest that vowel harmony in Kalaḅarị is also conditioned by the same articulatory feature as that in Okrika.” [2]

To conclude this section, we have seen that instrumental investigation can serve to confirm patterns for which we would otherwise depend solely on auditory data. Moreover, instrumental data may serve to test different phonological hypotheses in instances where the phonological patterns *per se* do not provide definitive evidence. In the next section, we turn to cases where auditory investigation runs into difficulties because of distinctions which are difficult to perceive with the naked ear.

3. Perceptually Difficult Distinctions

Instrumental methods can also ascertain the nature of distinctions that are difficult even for a trained field linguist to hear. Even if a particular pattern is not readily perceptible, there may be a sense that *some* distinction exists in the language under study, and at this point phonetic measures become indispensable. In addition, phonological or syntactic analysis may predict that there should be a certain distinction, even if its acoustic correlates have never been noticed. We will consider cases of both types below.

3.1. Perceptual uncertainty: Okpẹ

Okpẹ, an Edoid language of Nigeria, has been analyzed as having more underlying vowels than observed on the surface. According to Hoffmann (1973), Okpẹ has the nine phonemic vowels { i, ị, e, ẹ, a, ọ, o, ụ, u }, where the dotted vowels are “open” counterparts of the corresponding undotted vowels.

[2] Jenewari (1973) is referring to Orupabo (1973), a paper that we have not been able to consult.

Hoffmann presents compelling phonological evidence in favor of the nine-vowel analysis, pulling from observations related to vowel harmony, glide formation and vowel elision. The nine-vowel analysis is assumed in Kenstowicz & Kisseberth (1979), a book that was highly influential in establishing Okpẹ as one of the canonical examples of *absolute neutralization*. The crucial argument for neutralization of the high "retracted" vowels with the mid "advanced" vowels is from Hoffmann's observation that "[a]uditorily, no phonetic difference could be detected, and their phonetic identity was further supported by the native-speaker judgement that the two verbs *só* 'steal!' and *só* 'sing!' spoken in isolation are identical [...]" ((Hoffmann, 1973):101).

Omamor (1973), however, expressed doubt that Okpẹ exhibits surface neutralization. Although her focus is the related language Uvwiẹ, Omamor presents pilot acoustic data on both Uvwiẹ and Okpẹ. Her data suggest that both Uvwiẹ and Okpẹ have vowel systems analogous to that of Ịjọ (Figure 3): the mid advanced vowels are acoustically "higher" (i.e. lower F1) than the high retracted vowels. It seems that the overlap between high and mid vowels in this type of vowel system could lead to the divergence between acoustics and perception that Omamor points to in her conclusion. She suggests that these languages may not have nine vowels perceptually, but that they do have nine vowels both in terms of their phonological patterns and in terms of their acoustics.

Clearly, the Okpẹ case requires additional research, both acoustic and perceptual. Our purpose in citing these papers, however, is to demonstrate that perceptual uncertainty can potentially be resolved by recourse to instrumental examination.

3.2. Predictions of phonology: Kinande

A different kind of case can be observed in Kinande, a Bantu language of the Democratic Republic of the Congo. We begin with background. Kinande raises issues similar to those already discussed for languages like Ịjọ and Okpẹ. The phonology of the language motivates a pattern of vowel harmony ((Valinande, 1984; Mutaka, 1995; Archangeli & Pulleyblank, 2002)). For example, the infinitive prefix is advanced before advanced vowels (e.g. ẹ-rị-lị́b-a [erili:ba] 'to cover', ẹ-rị́-hụ́k-a [erihu:ka] 'to cook') and retracted before retracted vowels (e.g. e-ri-lím-a [ɛrɪlɪ:ma] 'to cultivate', e-ri-húm-a [ɛrɪhʊ:ma] 'to beat'). Kinande contrasts seven vowels lexically, { i, ɪ, ɛ, a, ɔ, ʊ, u }, but expands this set as a result of harmony. The nonlow vowels that occur phonetically are { i, ɪ, e, ɛ, ɔ, o, ʊ, u }. As discussed above, this is the kind of system argued by Clements to exhibit nested vowel aperture values. In Clements (1991), this is precisely the analysis Clements proposes for Kinande. Like Ịjọ, however, phonetic evidence supports an interactive model rather than a hierarchical model. As seen in both Gick et al. (2006) and Kenstowicz (2009), the acoustic vowel space of Kinande is of the 'crossed' type, comparable to the pattern for Ịjọ seen in Figure 3. In addition, Gick et al. (2006) present articulatory evidence derived from ultrasound data that the distinction between { i, e, o, u } and { ɪ, ɛ, ɔ, ʊ } is one of tongue root advancement/retraction.

Of interest here is the behavior of the low vowels. As noted in Gick et al. (2006), there have been differences of opinion as to whether low vowels, intrinsically retracted, have advanced variants in harmonically advanced contexts. Consider a pair of forms such as in (3).

(3) Low vowels in an advanced context

a.	Infinitive	e-ri-kár-a	[ɛrɪka:ra]	'force'
b.	Agentive	ọ-mụ-kạ́r-ị	[omukə:ri]	'forcer'

In the first example, (3a), the low vowel occurs with retracted vowels and is itself retracted. In the second example, (3b), the low vowel occurs with advanced vowels on both sides. Although transcribed in (3b) with an advanced vowel symbol, the low vowel's advancement is precisely the issue in such cases. There are two basic possibilities. On the one hand, the vowel could be phonologically retracted, with any degree of phonetic advancement due to phonetic coarticulation (the 'coarticulation' hypothesis); on the other hand, the vowel could be phonologically advanced in this context (namely, 'to the left of an advanced vowel'), hence fully advanced phonetically (the 'assimilation' hypothesis).

Gick et al. (2006) examine this issue instrumentally. They note that with sequences of low vowels, the coarticulation and assimilation hypotheses make different predictions.

(4) Low vowels: coarticulation or phonological spreading?

a. Coarticulation

+ATR
|
V2 V1 V
LO LO

Coarticulation
V1 > V2

b. Assimilation

+ATR
V2 V1 V
LO LO

Assimilation
V1 = V2

According to the coarticulation hypothesis, we would expect the degree of advancement to vary according to proximity to the phonetic source of advancement. That is, the closer a low vowel is to a following advanced vowel trigger, the more advancement the low vowel should exhibit. As indicated in (4a), we would expect ‘V1’ to be more advanced that ‘V2’. In contrast, according to the assimilation hypothesis, both vowels are advanced phonologically. As a result, we would expect both vowels to be equally advanced phonetically, that is, we would expect that ‘V1’ would be as advanced as ‘V2’ (4b).

These sorts of questions are difficult to answer through auditory data alone, and the answers bear on important theoretical issues. In the Kinande case, the issue at hand is whether low vowels undergo advancement harmony (contra certain claims about possible effects of tongue root advancement on low vowels, e.g. Kaye et al. (1985)) or whether they are skipped over by harmony (raising important issues about transparency in a harmony system). In Kinande, the answer appears to be that low vowels are undergoers. Based on ultrasound imaging data, Gick et al. (2006) show two points specifically relevant for low vowels. First, low vowels in advanced contexts exhibit advancement. Second, sequences of low vowels in an advancement context do not differ in their degree of advancement. Consider (5).

(5) Sequences of low vowels in a harmonic context: /...A...A...[hi atr].../

a. *Surrounded by atr vowels:* [mótwəsə́kíːre] ‘we remained’
b. *To the left of an atr vowel:* [kə́gəsu] (proper name)

In both contexts illustrated in (5), the ultrasound data in Gick et al. (2006) are consistent with the assimilation hypothesis, not the coarticulation hypothesis.

Before finishing this discussion of ‘hard-to-hear’ properties, we turn to a case from Yoruba of a theoretically motivated distinction, one where investigators had not noticed a difference auditorily.

3.3. Predictions of syntax: Yoruba

Yoruba, a Niger-Congo language spoken primarily in Nigeria, exhibits three tones: H(igh), M(id), L(ow). Of interest are various instances of the M tone. The M tone may appear in underived contexts:

(6) Underived mid tones in Yoruba

a. High kọ́ [kɔ́] ‘build’
b. Mid jẹ [ɟɛ] ‘eat’
c. Low mọ̀ [mɔ̀] ‘know’

In addition, the M tone may appear in contexts where it is morphologically or syntactically derived. The example in (7d) shows a morphologically derived case where a HL sequence surfaces as mid.

(7) Morphologically derived mid tones

a. /fìlà-kí-fìlà/ [fìlàkífìlà] ‘any cap’
b. /ọmọ-kí-ọmọ/ [ọmọkọ́mọ] ‘any child’
c. /eré-kí-eré/ [erékéré] ‘any play’

High-Low to Mid

d. /ẹ̀wà-kí-ẹ̀wà/ [ẹ̀wàkẹwà] ‘any beans’

The example in (8c) shows a syntactically derived case where a low tone verb surfaces as mid when followed by an object complement.

(8) Syntactically derived mid tones
Syntactically derived mid tones

a.	High	mo kọ́ ilé	[kɔ́]	'I built a house'
b.	Mid	mo jẹ ẹ̀wà	[ɟɛ]	'I ate beans'
c.	Low to mid	mo mọ̀ súlè	[mɔ̃̀] → [mɔ̃]	'I know Sule'

Although it had generally been assumed that both underlying and derived mid tones were the same, Ajíbóyè et al. (2011) suggest on the basis of syntactic *information-preservation* that the apparent syntactic change of a low to mid could not in fact involve the loss of the underlying low tone. The paper argues that the morphologically and syntactically derived mid tones should be distinct, and tests this hypothesis acoustically. The results are summarized in Figure 4.

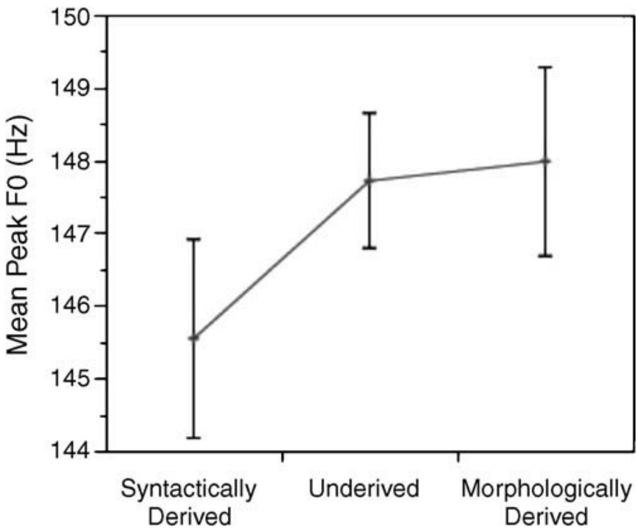

Figure 4: Derived and Underived Mid Tones in Yoruba

As illustrated, the syntactically derived mid tone is significantly different – lower – than the underlying mid tone and the morphologically derived mid tone. This is the result, Ajíbóyè et al. (2011) argue, of the retention of the low tone of the verb in (8c) as a floating tone.

The crucial point for our purposes here is that instrumental investigation of tonal acoustics serves to confirm a hypothesis that was advanced on the basis of a syntactic hypothesis. There had not been auditory evidence for a distinction between the three types of mid tones, but the syntax led us to check the tones acoustically.

4. Inaudible Distinctions

In addition to using instrumental techniques to confirm intuitions about audible patterns or to decide among multiple interpretations of an ambiguous signal, phonetic investigation can elucidate phonological patterns which lack an acoustic correlate entirely. Unlike cases we have discussed previously, these patterns are only accessible via an instrumental investigation of the vocal apparatus.

We focus here on the phonological insights to be gained from one particular type of inaudible articulatory data: the *inter-speech posture*, or ISP. The idea that every language exhibits a particular completely unmarked articulatory position is not new (Honikman, 1964; Laver, 1978; Jenner, 2001), but recently Gick et al. (2004) and Wilson (2006) have provided empirical support for such proposals using ultrasound. They suggest that there is a language-specific default position, the "inter-speech" posture, to which a speaker's articulators return between utterances. Recent studies by Namdaran (2006), Hudu (2010), and Allen & Pulleyblank (2012) have found a strong correlation between the ISP and phonologically-motivated claims about the organization of features in St'át'imcets, Dagbani, and Yoruba, respectively.

To illustrate the phonological relevance of ISP and the utility of ultrasound imaging, we will review the two of these studies that investigate Niger-Congo languages, Dagbani ((Hudu, 2010)) and Yoruba ((Allen & Pulleyblank, 2012)). These papers share two theoretical goals. First, each provides evidence from ultrasound imaging of the tongue that the harmonic classes in their respective languages are articulatorily distinct in terms of tongue root position, not tongue body height. Because a difference between significantly distinct tongue root position and significantly distinct tongue body height is, in principle, audible, these investigations constitute cases like those discussed in Section 3.

Second, both papers go beyond providing evidence that an [ATR] feature defines these harmonic classes: they argue that there is a relation between harmonically initiated articulator movement and the inter-speech posture. Both propose that in cases where the phonology necessarily involves tongue root advancement—phonological use of [+ATR]—the articulatory movement is one of advancement from ISP; in cases where the phonology necessarily involves tongue root retraction—phonological use of [−ATR]—the articulation is retraction from ISP. These predictions follow from the Direct Mapping Hypothesis of Hudu (2010). In essence, he proposes that the articulator position associated with a phonologically active feature value will always be significantly distinct from that articulator's ISP, while non-active values need not be. Significant displacement should also be in the direction associated with that feature, e.g. active [+ATR] should be significantly *advanced* compared to ISP, not retracted.

Dagbani and Yoruba constitute a relevant pair of languages for study. Hudu (2010) presents evidence that vowel harmony in Dagbani makes critical reference specifically to [+ATR], while Archangeli & Pulleyblank (1989, 1994) argue that harmony in Yoruba makes critical reference to [−ATR]. The Direct Mapping Hypothesis therefore predicts that Dagbani will show articulatory advancement from ISP while Yoruba will exhibit retraction from ISP. As we will describe in detail below, the applicability of this hypothesis to active-advancement languages is categorically supported by Dagbani ultrasound data (Hudu, 2010), and its applicability to active-retraction languages is tentatively supported by similar Yoruba data (Allen & Pulleyblank, 2012).

4.1. Dagbani: Advancement From ISP

The vowel system of Dagbani, a Gur language of Ghana, involves the interaction of two important sub-phonologies ((Hudu, 2010)). On the one hand, the realization of a vowel depends on its position in the word, whether it is medial or final in a root, and so on. On the other hand, a vowel's realization depends on harmony with respect to vowels in adjacent syllables. These interactions are intriguing and complex, and are covered in greater detail in Hudu (2010).

4.1.1. Dagbani Vowel Harmony

For our purposes, the crucial point about Dagbani is that an understanding of vowel harmony involves an understanding of the distribution of [+ATR]. Essentially, [+ATR] may be underlyingly present on a single vowel, [i]; [+ATR] may be assigned on the basis of position, for example, in root-final position for non-low vowels; [+ATR] may spread as a result of harmony, subject to various conditions including consonant opacity. For arguments demonstrating the dominance of the [+ATR] value, we refer the reader to Hudu (2010). We illustrate the pattern here with a single example pattern, which involves suffixation of a high retracted vowel.

Dagbani exhibits three high vowels underlyingly, two retracted { ɨ, ʊ } and one advanced { i }; there is a single low vowel contrastively { a }.[3]

[3] We abstract away from the mid vowels here as they present a somewhat more complex pattern.

(9) Dagbani high and low vowels

	Retracted			Advanced	
a.	dɨ̀m	'bite'	d.	ʤɨ̀m	'belch'
	nɨ̀ŋ	'do'		ʒɨ̂m	'blood'
b.	kʊ́ŋ	'empty'			
	tʊ̀m	'work'			
c.	tàm	'forget'			
	dâm	'alcohol'			

When a high retracted vowel is suffixed to such roots, the suffix vowel is retracted after a retracted root (10a-c). When the suffix vowel occurs to the right of an advanced vowel, however, the suffix surfaces as advanced (10d). Note that the vowel [u] does not occur contrastively, emerging only as a result of patterns of advancement.

(10) Dagbani root-triggered advancement

	Retracted			Advanced	
a.	bɨ́n-ɨ̂	'a thing'	d.	pín-î	'a gift'
	dɔ́lɨ́-bʊ̂	'following'		díʔ-í	'a mirror'
b.	dʊ́ʔ-ʊ̂	'a pot'		dí-hí-bû	'feeding'
c.	dá-bʊ̂	'buying'		tí-bû	'vomiting'

In this and numerous other examples that Hudu discusses, a [+ATR] value is phonologically active, or phonologically "dominant". Hudu's argument is that there are no patterns in the grammar of Dagbani that crucially refer to the value [−ATR], while various patterns make crucial reference to [+ATR].

4.1.2. Ultrasound Investigation of Dagbani

Turning to the vowel system's articulatory properties, Hudu (2010) demonstrates that the significant articulatory distinction between Dagbani's harmonic classes is one of tongue root position, not tongue body height. His methodology makes use of ultrasound video recording: target vowels were embedded in frame sentences which were produced by native speakers of Dagbani, and these productions were recorded with an Aloka SSD-900 ultrasound machine. After recording, vowel midpoint image frames were extracted from the video, and tongue root and tongue body positions were measured in pixels from the transducer midpoint to the tongue contour surface at predetermined degrees of displacement from the midline.

Analysis of these displacement values using standard *t*-tests yielded evidence of significant tongue root position differences between harmonic pairs of vowels; we reproduce a sample of his results for a single speaker in Figure 5. A significant distinction between the two classes in the tongue body height dimension was not found.

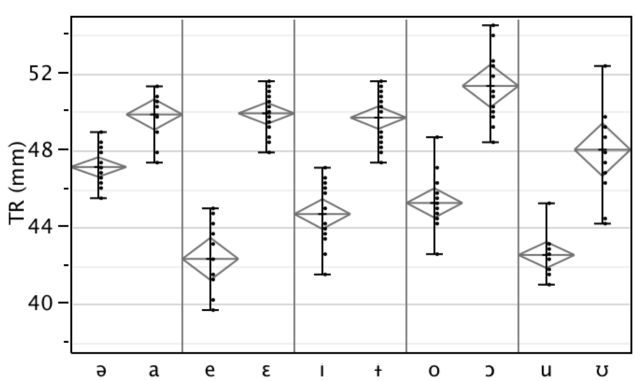

Figure 5: Tongue root distance from ultrasound transducer for both harmonic classes, from subject AB

Hudu (2010) also found support for the strong version of the Direct Mapping Hypothesis, which predicts as outlined above that in a language with phonologically active tongue root advancement (e.g. Dagbani) the [+ATR] tongue root position will be significantly advanced compared to ISP. Data for all five of his subjects corroborate this hypothesis. Four subjects' data are reproduced in Figure 6.

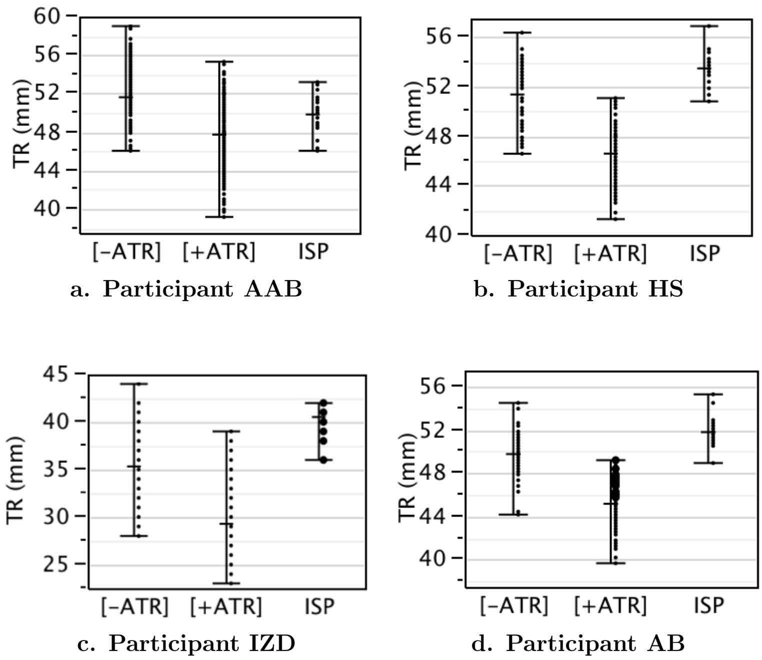

Figure 6: Tongue root distance from ultrasound transducer for both harmonic classes and ISP; four subjects

We summarize Hudu's 2010 results in Table 4. The notation $x > y$ indicates that x was significantly advanced in the tongue root dimension compared to y for a particular speaker.

subject code	*tongue root*
AAB	ATR $>$ ISP $>$ RTR
AB	ATR $>$ RTR $>$ ISP
AIM	ATR $>$ RTR $>$ ISP
HS	ATR $>$ RTR $>$ ISP
IZD	ATR $>$ RTR $>$ ISP

Table 4: Dagbani: ISP relative to grouped vowel classes

While most subjects' ISP was retracted compared to both the ATR and RTR positions, one subject's ISP was significantly advanced compared to the RTR position and significantly retracted compared to the ATR position. Crucially, despite this variation, the phonologically active ATR position is significantly advanced compared to ISP. Hence these data provide initial support for the strong version of Hudu's (2010) Direct Mapping Hypothesis.

4.2. Yoruba: Retraction From ISP

With evidence from only one language, however, it is impossible to determine whether the correlation between ISP location and phonological feature activity is more than a fortuitous coincidence. Seeking a way to further test the Direct Mapping Hypothesis, Allen & Pulleyblank (2012) report on a similar ultrasound analysis on Yoruba, a language claimed to demonstrate active tongue root retraction rather than advancement.

4.2.1. Yoruba Vowel Harmony

Standard Yoruba has seven oral vowels { i, e, ɛ, a, ɔ, o, u } and three nasalized vowels { ĩ, ã, ũ }. In anticipation of the evidence cited below, we refer to the { i, e, o, u } class as “advanced” and to the { ɛ, a, ɔ } class as “retracted.”

High	Advanced	i	u
Mid	Advanced	e	o
	Retracted	ɛ	ɔ
Low	Retracted	a	

Table 5: Standard Yoruba oral vowels

The basic harmony pattern centers around the behavior of the mid vowels. The two vowels { e, o } may appear in sequence with each other but not with { ɛ, ɔ }, and vice versa.

(11) Harmony in Yoruba mid vowels

Advanced			*Retracted*		
olè	[ole]	‘thief’	ẹ̀kọ	[ɛkɔ]	‘pap’
owó	[owo]	‘money’	ẹsẹ̀	[ɛsɛ]	‘foot’

This harmony pattern has been treated in a variety of ways. Fresco (1970) analyzes the relevant distinction between the mid vowels as due to the feature [tense], with high vowels redundantly [+tense] and low vowels redundantly [−tense]. Oyelaran (1973), on the other hand, classes the vowels { ɛ, ɔ } as [+low], along with { a }, while { i, e, o, u } are [−low]. A variant of this approach is taken by Goad (1993) who analyzes the relevant feature for { ɛ, a, ɔ } as monovalent [LOW]. Archangeli & Pulleyblank (1989, 1994) propose that the basic classification is along the lines of Fresco (1970), but that the relevant feature is [ATR].

Of particular interest here is the behavior of mid vowels when adjacent to high and low vowels. When following a high or low vowel, mid vowels may belong to either harmonic class: both { e, o } and { ɛ, ɔ } may follow high and low vowels.

(12) Mid vowels after high and mid vowels

Advanced			*Retracted*		
ilé	[ile]	‘house’	ilẹ̀	[ilɛ]	‘land’
ìgò	[igo]	‘bottle’	itọ́	[itɔ]	‘saliva’
ate	[ate]	‘hat’	àjẹ̀	[aɟɛ]	‘paddle’
àwo	[awo]	‘plate’	aṣọ	[aʃɔ]	‘cloth’

Such examples show that the harmonic value of a mid vowel does not interact with a vowel to its left. With a following high vowel, this pattern is echoed. Both advanced and retracted vowels may occur to the left of a high vowel:

(13) Mid vowels before high vowels

Advanced			*Retracted*		
ebi	[ebi]	‘hunger’	ẹ̀bí	[ɛbi]	‘land’
orí	[ori]	‘head’	ọ̀kín	[ɔkĩ]	‘egret’

Before low vowels, however, the situation is different: only retracted mid vowels may occur to the left of a low vowel.

(14) Mid vowels before low vowels

ẹ̀pà	[ɛkpa]	'groundnut'
ọjà	[ɔɟa]	'market'

Overall, the conclusion argued for by Archangeli & Pulleyblank (1989, 1994) is that tongue root retraction ("[−ATR]") is the phonological active feature, with a retracted vowel requiring retraction on the preceding (non-high) vowel but with an advanced vowel imposing no phonological requirements.

4.2.2. Ultrasound Investigation of Yoruba

Yoruba therefore constitutes a good language on which to further test the predictions of the strong version of Hudu's Direct Mapping Hypothesis. While both Dagbani and Yoruba exhibit tongue root distinctions in their vowel systems, Yoruba exhibits phonologically active retraction compared with Dagbani's active advancement. It is therefore possible to test whether the apparent correlation between ISP and vowel displacement is more than the result of only having analyzed languages with active advancement. Allen & Pulleyblank (2012) use a methodology very similar to that in Hudu (2010). However, whereas the latter relies on *t*-tests of measured distances between tongue contour edges and the transducer midpoint, the former performs calculations of statistical significance using a Smoothing Spline ANOVA algorithm (Davidson, 2006) implemented in R (R Core Team, 2012).

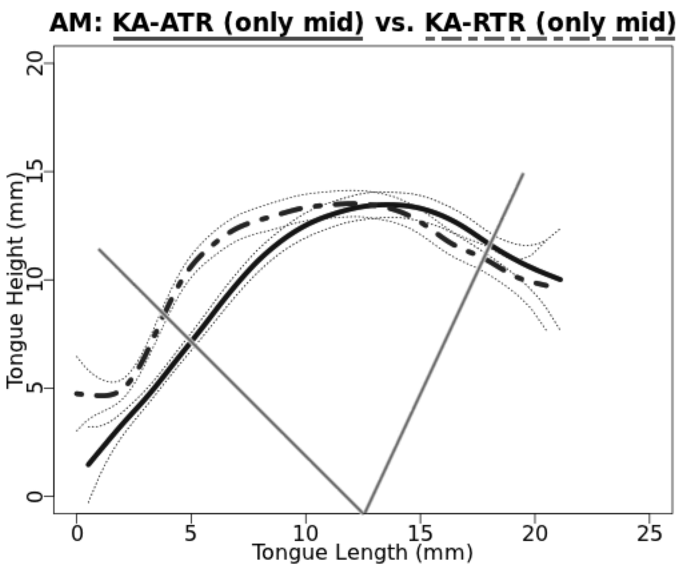

Figure 7: SSANOVA: ISP (solid line) vs. Advanced Mid Vowels (dotted line) for subject KA

The Smoothing Spline ANOVA algorithm produced graphs of averaged tongue contours and their Bayesian confidence intervals (the small dotted lines). See Figure 7 for an example of one speaker, with advanced vs. retracted mid vowels displayed. Overlapping confidence intervals at the tongue root locus (leftmost solid grey line) or tongue height locus (rightmost solid grey line) indicate a lack of statistical significance in that articulatory dimension.

Table 6 uses the same convention as Table 4 to summarize the results from Allen & Pulleyblank (2012), with one typographical addition: $x > y$ again indicates that x was significantly advanced in the tongue root dimension compared to y for a particular speaker; $x = y$ indicates that x was not significantly advanced or retracted compared to y.

subject code	*tongue root*
FD	ATR > RTR = ISP
KA	ATR > RTR > ISP
LB	ISP > ATR > RTR
RA	ISP > ATR > RTR
VA	ATR > ISP > RTR
YB	ISP = ATR > RTR

Table 6: Yoruba: Tongue root ISP relative to grouped vowel classes' tongue root; advanced vowels listed as ATR, and retracted as RTR

Setting aside the ISP for now, it is evident that the displacement between the advanced position and the retracted position was significant for all six subjects. Although not included here in a table, comparisons of these two harmonic classes in the tongue body height dimension demonstrated a complete *lack* of significance for all six subjects. Allen & Pulleyblank (2012) therefore conclude that tongue root position is the relevant articulatory correlate distinguishing the two harmonic vowel classes, lending support to a phonological analysis which invokes a tongue root feature rather than a height feature.

The ISP locations in Table 6 exhibit more variation than those in Dagbani, but still trend toward alignment with predictions made by Hudu's (2010) strong Direct Mapping Hypothesis. Specifically, the RTR position is significantly retracted from ISP for four out of six subjects. For details on the factors that we believe contribute to this lack of homogeneity, see Allen & Pulleyblank (2012).

5. Conclusion

Our primary purpose here has been twofold: to demonstrate several illustrative uses of phonetic investigation in informing phonological analyses, and also to establish a modest typology of such uses. Until recently, acoustic and articulatory studies in phonology have focused either on seeking confirmation of patterns perceived by field workers, or on using quantitative measures to help linguists choose among multiple interpretations of a signal or vocal gesture which would otherwise be left ambiguous. We add one more type of research to this set: studies in the last several years investigating inter-speech positions and their relationships to articulatory positions, which exemplify phonetic examination of an intrinsically inaudible aspect of language. We do not propose this typology to be of theoretical significance, but we hope that it can help guide phonologists interested in integrating phonetic data into their research.

We close, however, with a point that is of theoretical significance. With an increasing emphasis placed on using details of phonetics to supplement phonological theory, there has never been a more pressing need to ascertain the validity and specific nature of theoretical principles at play in the interface between phonology and phonetics. In this paper, we have discussed potential relationships between phonological and phonetic form, especially the Direct Mapping Hypothesis proposed by Hudu (2010). One goal of the research described here has been to test the limits of such principles, so that they can be employed with greater confidence and precision in designing projects and interpreting the results thereof. With the quickening development of software and hardware for phonetic analysis, we expect that the field's understanding of the phonetics-phonology interface will advance remarkably in the coming decades—contingent on further refinement of our understanding of the relationship between phonetic and phonological data.

References

Ajíbóyè, Ọládiípò, Rose-Marie Déchaine, Bryan Gick & Douglas Pulleyblank (2011). Disambiguating Yorùbá tones: At the interface between syntax, morphology, phonology and phonetics. *Lingua* 121, 1631–1648.

Akinlabi, Akinbiyi (1997). Kalaḅaṛị vowel harmony. *The Linguistic Review* 14, 97–138.

Allen, Blake & Douglas Pulleyblank (2012). Articulatory mapping of Yoruba vowels: an ultrasound study. Ms., University of British Columbia.

Archangeli, Diana & Douglas Pulleyblank (1989). Yoruba vowel harmony. *Linguistic Inquiry* 20, 173–217.

Archangeli, Diana & Douglas Pulleyblank (1994). *Grounded Phonology*. MIT Press, Cambridge.

Archangeli, Diana & Douglas Pulleyblank (2002). Kinande vowel harmony: Domains, grounded conditions, and one-sided alignment. *Phonology* 19, 139–188.

Boersma, Paul (2001). Praat, a system for doing phonetics by computer. *Glot International* 5:9/10, 341–345.

Boersma, Paul & David Weenink (2012). Praat: doing phonetics by computer (version 5.3.31). [computer program] Retrieved 2012 October 13, from http://www.praat.org/.

Clements, George N. (1990). The role of the sonority cycle in core syllabification. Kingston, John & Mary Beckman (eds.), *Papers in Laboratory Phonology I: Between the Grammar and Physics of Speech*, Cambridge University Press, Cambridge, 283–333.

Clements, George N. (1991). Vowel height assimilation in Bantu languages. *Working Papers of the Cornell Phonetics Laboratory* 5, 37–76.

Clements, George N. & Elizabeth V. Hume (1995). The internal organization of speech sounds. Goldsmith, John (ed.), *The Handbook of Phonological Theory*, Blackwell, Cambridge Mass., 245–306.

Davidson, Lisa (2006). Comparing tongue shapes from ultrasound imaging using smoothing spline analysis of variance. *Journal of the Acoustical Society of America* 120:1, 407–415.

Fresco, Edward (1970). Topics in Yorùbá dialect phonology. *Studies in African Linguistics, Supplement 1* .

Fry, D.B. (1955). Duration and intensity as physical correlates of linguistic stress. *The Journal of the Acoustical Society of America* 27:4, 765–768.

Gick, Bryan, Ian Wilson, Karsten Koch & Clare Cook (2004). Language-specific articulatory settings: Evidence from inter-utterance rest position. *Phonetica* 61:4, 220–233.

Gick, Bryan, Douglas Pulleyblank, Fiona Campbell & Ngessimo Mutaka (2006). Low vowels and transparency in Kinande vowel harmony. *Phonology* 23:01, 1–20.

Goad, Heather (1993). *On the Configuration of Height Features*. Ph.D. thesis, University of Southern California.

Hazen, Kirk (1973). Effects of differing phonetic contexts on spectrographic speaker identification. *The Journal of the Acoustical Society of America* 54:3, 650–660.

Hoffmann, Carl (1973). The vowel harmony system of the Okpe monosyllabic verb. *Research Notes from the Department of Linguistics and Nigerian Languages* 6, 79–111.

Honikman, Beatrice (1964). Articulatory settings. *In honour of Daniel Jones* 73–84.

Hudu, Fusheini A. (2010). *Dagbani tongue-root harmony: a formal account with ultrasound investigation*. Ph.D. thesis, University of British Columbia.

Jenewari, Charles (1973). Vowel harmony in kalaḅarị Ịjọ. *Research Notes from the Department of Linguistics and Nigerian Languages* 6, 59–78.

Jenner, Bryan (2001). Genealogies of articulatory settings: Genealogies of an idea. *Historiographia Linguistica, 28* 1:2, 121–141.

Kaye, Jonathan, Jean Lowenstamm & Jean-Roger Vergnaud (1985). The internal structure of phonological elements: a theory of charm and government. *Phonology* 2, 305–328.

Kenstowicz, Michael (2009). Two notes on Kinande vowel harmony. *Language Sciences* 31, 248–270.

Kenstowicz, Michael & Charles Kisseberth (1979). *Generative Phonology: Description and Theory*. Academic Press, London.

Ladefoged, Peter & Morris Halle (1988). Some major features of the international phonetic alphabet. *Language* 64:3, 577–582, URL `http://www.jstor.org/stable/414533`.

Ladefoged, Peter & Ian Maddieson (1996). *The Sounds of the World's Languages*. Blackwell, Oxford.

Laver, John (1978). The concept of articulatory settings: an historical survey. *Historiographia Linguistica, 5* 1:2, 1–14.

Lindblom, Björn (1962). Accuracy and limitations of sonagraph measurements. *Proc. 4th Int. Congr. Phonet. Sci.*, Mouton, The Hague.

Mutaka, Ngessimo (1995). Vowel harmony in Kinande. *Journal of West African Languages* 25, 41–55.

Namdaran, Nahal (2006). Retraction in Státimcets: An ultrasonic investigation. M.A. thesis.

Omamor, Augusta Phil (1973). Uvwiẹ - a case of vowels merging. *Research Notes from the Department of Linguistics and Nigerian Languages* 6, 113–143.

Orie, Ọlanikẹ Ọla (2003). Two harmony theories and high vowel patterns in Ebira and Yoruba. *The Linguistic Review* 20, 1–35.

Orupabo, G. J. (1973). A tentative phonology of the Okrika dialect of Ịjọ. Ms., Ibadan.

Oyelaran, Ọlasope O. (1973). Yoruba vowel co-occurrence restrictions. *Studies in African Linguistics* 4:2, 155–182.

R Core Team (2012). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, URL `http://www.R-project.org`. ISBN 3-900051-07-0.

Valinande, Nzama K. (1984). *The structure of Kinande*. Ph.D. thesis, Georgetown University.

Williamson, Kay (1969). Ịjọ. Dunstan, Elizabeth (ed.), *Twelve Nigerian Languages*, Africana Publishing Corporation, New York, 97–114.

Wilson, Ian L. (2006). *Articulatory settings of French and English monolingual and bilingual speakers*. Ph.D. thesis, University of British Columbia.

Selected Proceedings of the 43rd Annual Conference on African Linguistics: Linguistic Interfaces in African Languages

edited by Ọlanikẹ Ọla Orie and Karen W. Sanders

Cascadilla Proceedings Project Somerville, MA 2013

Copyright information

ISBN 978-1-57473-460-7 library binding

Ordering information

Orders for the library binding edition are handled by Cascadilla Press.
To place an order, go to www.lingref.com or contact:

Cascadilla Press, P.O. Box 440355, Somerville, MA 02144, USA
phone: 1-617-776-2370, fax: 1-617-776-2271, sales@cascadilla.com

Web access and citation information