VOWEL HARMONY AND COARTICULATION IN THREE DIALECTS OF YORUBA: PHONETICS DETERMINING PHONOLOGY

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Marek A. Przezdziecki

January 2005

VOWEL HARMONY AND COARTICULATION
IN THREE DIALECTS OF YORUBA:
PHONETICS DETERMINING PHONOLOGY

Marek A. Przezdziecki, Ph. D.
Cornell University 2005

This dissertation examines the phonology and acoustic phonetics of vowels in three dialects of Yoruba—Standard Yorùbá, Mọ̀bà, and Àkùrẹ́ Yorùbá—to investigate the role of coarticulation in the phonologization of vowel harmony (Ohala 1994). The phonological vowel patterns of the three dialects are presented. Àkùrẹ́ Yorùbá exhibits Advanced Tongue Root (ATR) vowel harmony in mid and high vowels, while harmony in Mọ̀bà and Standard Yorùbá does not extend to high vowels. In order to investigate this relationship, recordings of VCV nonsense words from speakers of each dialect were analyzed. Following Hess (1992), the first formant (F1) was determined to be the acoustic measurement best correlated to the ±ATR vowel sets. Other measurements—F2, F1 bandwidth, fundamental frequency, vowel duration, and spectral measures—were not found to correlate with ATR. Using F1 as a measure, vowel to vowel coarticulation in high vowels in Mọ̀bà and Standard Yorùbá was found to resemble high vowel harmony in Àkùrẹ́ in the target vowels, the context, and the phonetic effect. This was particularly true for /i/; however the coarticulatory effects on /u/ were weaker and not statistically significant. As expected, the effect of vowel to vowel coarticulation in Mọ̀bà and Standard Yorùbá was smaller and less robust than for vowel harmony in Àkùrẹ́. A decision tree model is proposed that is able to generate the high vowel harmony pattern from the Àkùrẹ́ acoustic data. More

interestingly, the model succeeds at extracting—to a large degree—the high vowel harmony pattern from Mọ̀bà and Standard Yorùbá, the dialects without high vowel harmony. The model does not require any reference to features or natural classes, suggesting that it is not necessary to posit features as a prerequisite to learning a phonological pattern, nor as an explanation for universal patterns. The study argues that the acoustic patterns found in vowel to vowel coarticulation are sufficient to result in vowel harmony. The findings are consistent with the view that proto-Yorùbá did not have harmony in its high vowels (Fresco 1970, Oyelaran 1973, and Capo 1985), and that high vowel harmony developed in Àkùrẹ́ and related dialects.

BIOGRAPHICAL SKETCH

Marek Przezdziecki was born in London, England on March 11, 1962. His family moved to Rochester, New York in 1967. He received a Bachelor of Science degree in Computer Science and Applied Mathematics from the State University of New York at Albany in 1983. In 1991, he received a Master of Science in Education from the University of Rochester. He worked as a text to speech scientist at Eloquent Technology and Speechworks International from 1999 to 2003.

For Vivian, Sam, and Ned

ACKNOWLEDGEMENTS

I am grateful to all members of my committee, past and present. Vicki Carstens encouraged an interest in Yorùbá and first suggested the obvious, that I go to Nigeria. Frederick Parkinson's classroom was open to new ideas and fostered fruitful discussions. Amanda Miller-Ockhuisen arrived at Cornell later on, and contributed greatly to the experimental side. Akinbiyi Akinlabi has provided me with invaluable support on Yorùbá phonology. Draga Zec has always been eager to debate and to make sure I support my claims. I am especially indebted to my chair, Abby Cohn, who devoted so much time and energy into helping me turn an epiphany into a dissertation.

This project relied on the generosity of many Yorùbá teachers and consultants, people who gave their time with little reason simply because I was interested in their language. Akinloye Ojo, Rolake Akinola, Mayowa Adefolaju, and Kayode Kekere made trying to learn Yorùbá fun. I enjoyed meeting and working with many Yorùbá consultants in Toronto, Rochester, and the Washington, D.C. area, in particular Mrs. Esther Opadiran, Mr. Isaac Ige, and Prince Elvis. I regret that I was not able to visit and record people from all of their hometowns, like Ilesa, Ekiti, and Ondo. Oladiipo Ajiboye was kind to help me with Mọ̀bà questions via email. My trip to Nigeria was made so much nicer by the generosity of Remi Ṣonaiya and her family who provided a home away from home and invaluable assistance in my travels. I was fortunate to make the acquaintance of Chief James Dele Popoola at Cornell University, who provided me with Standard Yorùbá and Mọ̀bà data and helped make my visit to Ọ̀tùn very memorable. Once in Ọ̀tùn, Dejo and Kayode helped me in every possible way. I was directed to Ṣebi and Samson in Àkùrẹ́ who were great contacts. In Ìbàdàn, Professor Francis Egbokare was kind enough to help me find speakers and show me around campus. I would also like to thank the people who agreed to be recorded for this study.

My linguistic development has been greatly influenced by many linguists. I have benefited from discussions with Lee Bickmore, Nick Clements, Mark Hale, Morris Halle, Larry Hyman, Scott Myers, Jay Padgett, Charles Reiss, Daniel Silverman, and many others. At Cornell, I would like to thank my linguist friends and colleagues over the years, David Parkinson, Aneta Pavlenko, Antony Dubach Green, Tobey Doeleman, Lisa Lavoie, Bill Ham, Teresa Galloway, Bruce Cain, Eungyeong Kang, Justin Spence, Catalin Kaser, Arthur Bell, Anastasia Riehl, and many others. For all things administrative, Sheila Haddad and Angie Tinti have made life much easier for me. I would also like to thank Eric Evans for extensive help in the Phonetics Laboratory.

This work was partly supported by a National Science Foundation Graduate Research Fellowship and a Cornell University Sage Fellowship. Travel to Nigeria was partly funded by a travel grant from the Mario Einaudi Center for International Studies International Research at Cornell.

I would like to acknowledge the support that friends and family have given me in so many ways over these years. My parents and my in-laws have gently encouraged me and asked “When are you going to be finished?” much less than they could have. I would especially like to thank my wife, Vivian Lorenzo, and my children, Sam and Ned, for allowing me the time to work, time and time again, and giving me a reason not to. Vivian, I owe you 723 loads of laundry, 2311 meals, 27 summer camp applications, plus a whole lot more that cannot be measured or counted.

In spite of all of the help I have received, some errors surely remain, for which I have only myself to blame, and arguably my parents.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

ANOVA	analysis of variance
ATR	advanced tongue root
BW	(formant) bandwidth
C	consonant
F0	fundamental frequency
F1, F2, F3	first formant, second formant, third formant
FG	feature geometry
gen.	genitive
Hz	hertz
IPA	International Phonetic Alphabet
nom.	nominative
OT	Optimality Theory
pl.	plural
sg.	singular
SPE	Sound Pattern of English
SY	Standard Yorùbá
UG	universal grammar
V1, V2	first vowel, second vowel
VCV	vowel-consonant-vowel

CHAPTER ONE

PHONETICS TO PHONOLOGY

It is central to phonology to identify and characterize sound patterns of individual languages, and recurring sound patterns across languages. Phonologists attempt not only to describe the patterns but also to account for their existence, or more specifically, to explain why certain patterns occur in languages while other conceivable patterns do not. While linguists do not agree on the origin of these phonological patterns, the existence of frequently occurring phonological patterns is uncontroversial. The most prevalent explanations for the existence of phonological patterns fall into two general classes. In the first class, the phonological patterns are a reflex of the inherent linguistic properties in the mind. This approach may be called nativist: patterns reflect innate constraints (in a general sense of the word) which directly impact the kinds of rules, constraints, or representations that are available to speakers and/or learners of a language. Another way of saying this, is that our innate linguistic cognitive machinery can only handle certain patterns, and those are the ones that emerge as language universals. A second class of explanation attributes the phonological patterns to a wide variety of phonetic explanations, including auditory, perceptual, and articulatory explanations. (While some of these domains, namely audition and perception, are cognitive in nature, they are not purely linguistic.) In this class of explanation, there are many variants. In one such view, phonological patterns are remnants of earlier phonetic patterns. These phonetic patterns, in turn, are rooted in the physical world, as it applies to the human articulatory apparatus and perceptual system used in speech. It is this view that is supported by the evidence presented in this dissertation.

I examine the vowels of three Yorùbá dialects to show how phonological patterns emerge from phonetic patterns. Àkùrẹ́ Yorùbá exhibits a phonological pattern—ATR high vowel harmony—that is absent from Standard Yorùbá (SY) and Mọ̀bà Yorùbá. However, an examination of the phonetic details of vowel-to-vowel coarticulation in SY and Mọ̀bà shows a phonetic pattern that resembles Àkùrẹ́'s harmony in several ways. The similarities suggest that the high vowel harmony of Àkùrẹ́ emerged diachronically from coarticulation in an earlier stage of the language, supporting the view held by Ohala (1994:491) that "Vowel Harmony [...] is a fossilized remnant of an earlier phonetic process involving vowel-to-vowel assimilation." More generally, this is the type of evidence we expect to find if phonology emerges from phonetics.

The parallels between phonology and phonetics have long been recognized. While many phonologists may accept that phonology emerges from the phonetics, the dominant phonological theories are not consistent with the implications that follow from such a view. For if phonetics plays a role in determining patterns of phonology—and in this way the universal patterns—then phonological theories should not have a distinct, and therefore redundant, explanation for these same universal patterns. Knowing where phonological universals come from plays a role in determining the nature of the phonological model. If patterns are a more or less direct result of the constraints imposed by our mental language faculty, then a model of the speaker's phonological knowledge must make reference to these constraints, or more precisely, the model of allowable human phonologies has to make reference to them. On the other hand, if the patterns come from phonetic constraints, then the model must not make reference to the constraints if the model is intended to model phonological knowledge.

In the following sections of this chapter, I discuss recent phonological theories and how they attribute phonological patterns, not to phonetic origins, but instead to innate properties of the phonological component of the mind.

1.1 Universal phonological patterning

A fundamental observation in phonology is that phonological patterns recur throughout the languages of the world. With most phonological patterns, if they are found in one language, they can almost always be found in other languages. The examples I give here are coda consonant devoicing, nasal assimilation, and vowel harmony. It is common for languages allowing coda consonants to prefer voiceless codas over voiced codas. A well known example of this is Polish, where word-final consonants are devoiced, as shown in (1). Obstruents can be voiced or voiceless when followed by a vowel as in the plural forms, but when the obstruents are word final, as in the singular, they are always voiced.

(1) Polish Consonant Devoicing (Kenstowicz, 1994)

singular	plural	gloss
klup	klubi	‘club’
trup	trupi	‘corpse’
trut	trudi	‘labor’
kot	koti	‘cat’

When nasal consonants precede stops, the nasals frequently assimilate to the place of the stop. This is the case in Bassar, a Gur (Niger-Congo) language from Togo and Ghana, as shown in (2).

(2) Bassar Nasal Assimilation (data from Lare 1990)

[b,p]:	búgbúmbu, ‘kapok tree’	binìcàmbi, ‘foreigners’
	mbúmi, ‘anger’	udumpu, ‘house’
	mpɔ̀mpɔ̀limi, ‘mold’	
[d,t]:	ndààmi, ‘drink’	digòóǹdi, ‘kola nut’
	tɔ̀ntɔ̀, ‘cat’	ulantan, ‘monkey’

[k,t]: ŋkaami, ‘bile’ kúkúŋkoou, ‘feather’

ŋgèèmi, ‘fatigue’ gàŋgààŋkú, ‘centipede’

Finally, let’s look at vowel harmony in Turkish. Vowel harmonies of various types are found in numerous languages. In each case, the harmony may be characterized as an assimilation in which a vowel exhibits some characteristic (i.e., assimilates a feature) of an adjacent vowel usually without regard to intervening consonants. This general pattern of assimilation between vowels is common across different types of harmonies. However, the harmony types vary considerably in the feature or features assimilated, the domain, the direction of the assimilation, and the sets of target vowels undergoing and triggering the assimilation. In Turkish vowel harmony, the feature [back] spreads from the initial vowel rightward. The vowels of Turkish are listed in Table 1.1.

Table 1.1. Turkish Vowels (from Beddor and Yavuz 1995)

	Front	Back
High	i, ü (=IPA y)	ï (=IPA ɯ),u
Low	e, ö (=IPA ø)	a,o

The spreading of [back] is shown in (3) where the nominative plural ending surfaces with a back vowel in [-lar] after a back vowel in [somun-], but it surfaces with a front vowel in [-ler] after a front vowel in [kemik-]. The harmony extends to the second vowel in the suffix in the genitive plural, so the back vowels in [-larɯ] occur after a back vowel, while the front vowels in [-leri] occur after a front vowel. [1]

[1] In most words, adjacent vowels agree in backness, though some words are **disharmonic**, for example [biɾa], ‘beer’; [misal], ‘example’ (Beddor and Yavuz, 1995).

(3) (from Beddor and Yavuz, 1995)

	BACK	FRONT
	‘loaf’	‘bone’
nom. sg.	somun	kemik
nom. pl.	somun-l**a**r	kemik-l**e**r
gen. pl.	somun-l**arɯ**	kemik-l**eri**

In addition to backness spreading, another aspect of this harmony is that the feature [round] spreads rightward to high vowels, but not to other vowels.

For each of these three patterns—coda devoicing, nasal assimilation, and vowel harmony—the relatively common occurrence of the pattern is underscored by the absence or near absence of patterns that seem to do the opposite of what these patterns do. For example, while many languages allow both voiced and voiceless obstruents in coda positions, we do not find languages where a coda obstruent must be voiced as in (4).

(4) Unattested coda obstruent voicing pattern (compare with (1))

singular	plural
klu$\boxed{b}$	klu$\boxed{b}$i
tru$\boxed{b}$	tru$\boxed{p}$i
tru$\boxed{d}$	tru$\boxed{d}$i
ko$\boxed{d}$	ko$\boxed{t}$i

Likewise, languages exist in which nasal stops may be dissimilar with respect to place of articulation to following consonants. However, we do not find languages where nasals consonants always differ from the place of the following stop as in (5).

(5) Unattested nasal dissimilation pattern (compare with (2))

singular	plural
ba	a$\boxed{\text{n}}$-ba
ta	a$\boxed{\text{m}}$-ta
ka	a$\boxed{\text{m}}$-ka

Finally, while most languages do not have vowel harmonies, we do not find languages with an anti-harmony vowel system such as shown in (6), where adjacent vowels cannot share the same value of front and back. If such systems do exist, they are at least very uncommon.

(6) Unattested vowel disharmony pattern (compare with (3))

stem	somyn	kemɯk
stem+suffix	somyn-l**a**r	kemɯk-l**e**r
stem+suffix+suffix	somyn-l**ari**	kemɯk-l**erɯ**

Returning to the attested patterns, for each phonological pattern above we observe a phonetic phenomenon exhibiting parallel characteristics. In the case of devoicing of final obstruents, the phonetic phenomenon of devoicing occurs in languages that do not have phonological devoicing. For nasal assimilation, perception seems to be most important; pre-stop nasals have weak place cues (Malécot 1956, Ohala 1990, Ohala and Ohala 1993). Finally, vowel to vowel coarticulation exhibits parallels to vowel harmony (Beddor and Yavuz 1995, Majors 1998). In fact, commonly occurring phonological rules are usually phonetically natural, in the sense that they are often associated with common phonetic phenomena. These pairings, called phonetic and phonological doublets by Cohn (1998), have been observed by Chomsky and Halle (1968), Hyman (1976), Ohala (1995), and many others. Myers (1997) lists several of these doublets: nasalization of vowels adjacent to nasal consonants (Cohn 1993); closed syllable shortening of vowels (Maddieson 1985); palatalization (Zsiga 1995); and vowel reduction (Lindblom 1963). Other examples include lenition (Lavoie 1999) and tonogenesis (Hombert 1978). In fact, for most commonly occurring phonological patterns, we find similar patterns in the phonetic domain. This observation is the central point of departure for this dissertation, in which I present evidence to support the claim that phonology and phonetics have many parallels simply because phonological phenomena emerge from phonetic ones.

In the next section, I examine how recent phonological theories, starting with Chomsky and Halle (1968), have accounted for universal patterns and the parallels between phonology and phonetics.

1.2 SPE explanations for phonological universals

Of the explanations for the universal patterns, the most widely held is based on Chomsky and Halle's (1968) *The Sound Pattern of English* (SPE). Although phonological theory has changed considerably, many of the assumptions set forth in SPE still hold today. The SPE view holds that grammars are constrained by innate linguistic structure, which is then reflected in the patterns of language. Chomsky and Halle propose a set of universal features to account for all phonological contrasts in natural spoken languages. They explicitly state that these features as well as other linguistic properties are universal and innate. "The significant linguistic universals are those that must be assumed to be available to the child learning a language as an *a priori*, innate endowment (p. 4)." In calling this the *nativist* position, I recognize that there is not one unified nativist position, but rather a continuum, or even several continua, from nativist to non-nativist positions. The hypothesized innate component, known as *Universal Grammar*, or *UG*, is specifically linguistic and mental in its nature, as opposed to being derived from general non-domain-specific cognitive structures, from non-linguistic auditory structures, or from physical properties of the vocal tract. While these other domains—or components of them—are certainly innate, they are not typically considered part of UG, and I do not consider them part of UG here. Universal Grammar, by the definition I adhere to here, is the set of linguistic rules, constraints, parameters, and other mental structures taken to form the innate linguistic component of our cognitive system. By Universal Grammar, I do not mean the universal properties, characteristics, or tendencies of language in themselves. My defining UG in this way is intended to be a strictly terminological issue, rather than a theoretical one. Furthermore, by stating what kinds of properties I consider UG, I am not claiming that these properties exist, but rather, to the extent that they do exist, they would make up UG. An empirical question I address here is the extent to which the universal properties are reflexes of UG.

Chomsky and Halle's model of universality as it relates to universal grammar can be schematized as I have done in Figure 1.1.

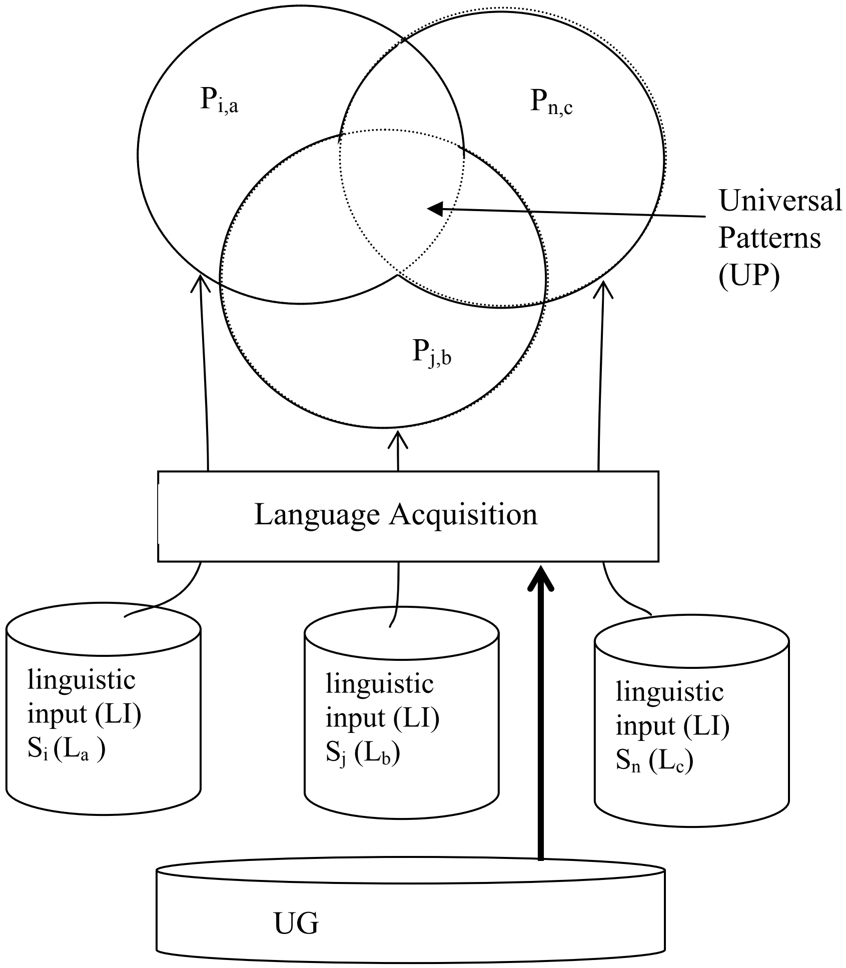


Figure 1.1. Schematization of SPE model of universality as it relates to Universal Grammar (UG). L_a, L_b, L_c are languages; and $S_{i,}$ S_j, S_k are speakers.

For each speaker (S), the acquired phonology (P) is a function of the linguistic input (LI) and universal grammar (UG). Linguistic input, of course, varies depending on the language environment, as well as on non-linguistic factors. Looking at a large number of speakers and languages (more than the three I show here), one can find the universal patterns. The language acquisition process is at least partly constrained by UG. From these universal patterns (UP), one can then posit what kind of structure in the UG could project those patterns. The primary arguments in favor of this position reside in two domains: (1) learnability and (2) universal patterns.

The learnability argument says that the linguistic input (LI) that learners are exposed to is insufficient to arrive at the grammar (P). In SPE, Chomsky and Halle claim that children are able to acquire language only because of innate linguistic properties. These *a priori* linguistic structures reduce the grammars available for consideration to the learner, thus rendering the learning task simpler, and accounting for the ease with which children learn language in the face of limited input.

> That there must be a rich system of a priori properties—of essential linguistic universals—is fairly obvious from the following empirical observations. Every normal child acquires an extremely intricate and abstract grammar, the properties of which are much underdetermined by the available data. This takes place with great speed, under conditions that are far from ideal, and there is little significant variation among children who may differ greatly in intelligence and experience (SPE p. 4).

Another way of looking at it is that children have much of the language learned—in the structure of UG—before they even hear their first phone. If it can indeed be shown that certain linguistic structures must be present for language acquisition to take place, then this evidence alone would suffice to prove the existence of these structures. Of course, this case is very difficult to make, considering the complexity of the environment to which children are exposed. The issue of learnability is beyond the scope of this work. To my knowledge there is no conclusive evidence that linguistic knowledge is a necessary precondition to language acquisition, and I proceed under the assumption that researchers have not, by this approach, established the existence of innate linguistic structures in the phonological domain.

The universal patterns argument is the one I address here, specifically the explanation of universal phonological patterns. In SPE, the properties of UG that are posited to account for language learning also account for the recurrence of phonological patterns. The grammars allowed by a human linguistic-cognitive system are limited by UG, the result of which is that the grammars existing in the world's languages are similarly limited. The intention of the SPE grammatical model is to readily generate common patterns, or more exactly, to readily generate grammars that in turn generate common (natural) patterns. At the same time, UG and its subsequent grammars should not readily generate unnatural or unattested patterns, such as those shown in (4)–(6) above. It is a common theoretical assumption, if sometimes an unstated one, that the posited Universal Grammar should, in this way, tightly fit or generate the attested data (that is, UP from Figure 1.1, above). This assumption is explicitly stated in Archangeli and Pulleyblank (1994:3):

> Phonological theory must proved an explicit characterization of the full range of attested representations and rules, but at the same time should be incapable of characterizing many of the imaginable patterns that are not attested in natural language. The success of a particular theory can then be measured by the tightness of the fit between the theory and the range of attested phonological patterns.

I call this an *articulated* [2] UG, in which universal patterns are by their very nature a reflex of UG. It is important to keep in mind that the SPE phonological model is intended to represent speaker knowledge, with a strong component coming from innate linguistic structure. The same is true for many subsequent grammatical models, such as feature geometry (Clements, 1985; Sagey, 1986; McCarthy, 1988; Odden, 1991; Halle, 1995; and Clements and Hume, 1995) and optimality theory (Prince and Smolensky, 1993; McCarthy and Prince, 1993).

While it may seem straightforward and uncontroversial to seek a model that directly accounts for the attested patterns, I argue here that those patterns whose origins can be attributed to factors outside the grammar, should not be represented in a model of UG. If we intend to model the patterns alone, then indeed we want a model that generates the patterns that exist or could exist and only those patterns. With a

---

[2] I use *articulated* in the sense of 'detailed.'

descriptive model[3] such as this, we are not concerned with the origin of the patterns, in fact, we are making no claims about the origin or explanation of the patterns. We are only concerned about modeling the patterns. Descriptive modeling is valuable, and a necessary step in seeking the explanation for phonological patterning. However, a descriptive model is not in itself an explanatory one, nor should it claim to be. A descriptive model is successful if it elegantly captures the generalizations found in the data.

Most theoretical phonologists today are not satisfied with restricting their studies to the modeling of the patterns. Instead, striving for explanatory models, phonologists seek not only to model the patterns, but also to provide an account for the source of the patterns. For phonologists who advocate an articulated UG, the explanatory model will be equally as articulated, because the speaker's *a priori*, or innate, knowledge will be represented in the model. In fact, for any explanatory model, the innate knowledge—to the extent it is posited—must necessarily be represented. The issue here is not whether innate knowledge should be represented in an explanatory model, but rather what the content of the innate knowledge is. In assessing the adequacy of a descriptive model, the criteria are satisfied if the possible outputs of the grammatical model are the same as (or close to) the variation of output of actual speakers. The criteria for assessing the adequacy of an explanatory model are more strict; the model must not only get the output right, but must get it right for the right reasons. If we believe that the attested universal patterns are largely due to the constraints imposed by UG, then a model based on those patterns may be a good approximation of a speaker's knowledge. However, if we believe that at least some of the universal patterns are due to extragrammatical causes, then we do not want to posit a mechanism in our model that arrives at the same patterns without reference to the

[3] The terms *descriptive* and *explanatory model* come from Chomsky (1965).

extragrammatical causes. By Occam's Razor, we should not posit two sources of the same phenomenon.

In SPE (as in Feature Geometry and Optimality Theory), the criteria for a descriptive model are—to a certain degree—satisfied. To the extent that a feature geometry model, for example, elegantly captures the patterns found in languages, then it is a successful descriptive model. From this descriptive model, one could then hypothesize that the success of the model is because the feature geometry posited is somehow present in the speaker or learner's mind. In order to show this, however, it is not sufficient to show additional patterns captured by the model. If it can be shown that the existing patterns are due to the manner in which languages change over time, then, depending on the nature of the diachronic influences, we may not want the resulting patterns to be attributed to a model of the speaker's mind.

In SPE, universal patterns are accounted for by positing models capable of representing only those patterns found in languages. The general reasoning is that universal principles restrict the types of allowable phonologies. This view is illustrated in Figure 1.2.

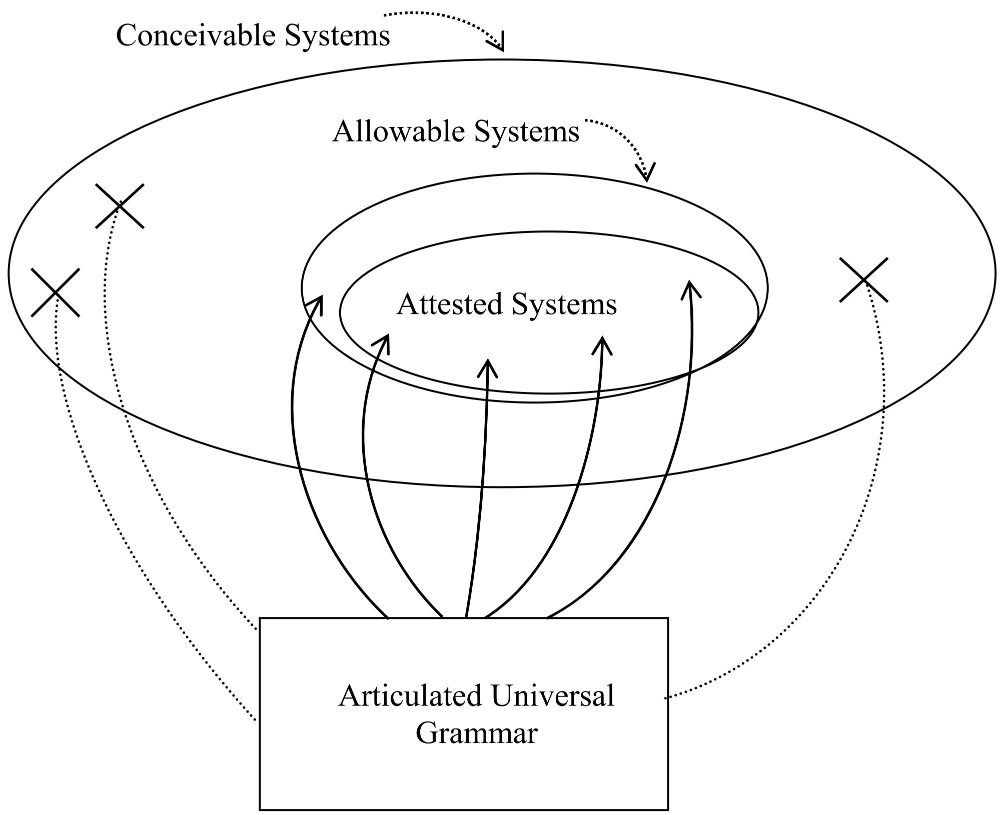


Figure 1.2. Articulated Universal Grammar model, as is found in SPE, and other nativist models.

Universal Grammar generates a number of grammars, represented by the arcs with arrows. Certain conceivable grammars cannot be generated by UG; these are represented by the dashed arcs with Xs at the endpoints. Here, we may include grammars with unattested patterns such as coda consonant voicing (4), nasal dissimilation (5), and vowel disharmony (6) in this group. As we see, only a small subset of the conceivable phonological systems may be generated. A further subset of the allowable systems is attested; since some of the allowable systems have not been attested, we see a difference in size of the smaller ellipses.

In SPE's epilogue Chomsky and Halle (1968:400) provide a revealing caveat that their model is not as tightly fit as they would like. That is, their theory, until that point, generates more grammars than are attested:

> The entire discussion of phonology in this book suffers from a fundamental theoretical inadequacy. Although we do not know how to remedy it fully, we feel that the outlines of a solution can be sketched, at least in part. The problem is that our approach to features, to rules, and to evaluation has been overly formal. Suppose, for example, that we were systematically to interchange features or to replace [αF] by [$-\alpha$F] (where $\alpha = +$, and F is a feature) throughout our description of English structure. There is nothing in our account of linguistics theory to indicate that the result would be the description of a system that violates certain principles governing human languages. To the extent that this is true, we have failed to formulate the principles of linguistic theory, of universal grammar, in a satisfactory manner. In particular, we have not made any use of the fact that the features have intrinsic content.

They identify a number of ways in which the framework does not always favor (or predict) natural over unnatural phenomena, such as:

- A common phonological process may require a more complex rule than an uncommon phonological process would.
- Similarly, classes that are observed to be natural in most phonological systems (such as voiced obstruents) are sometimes more complex to represent than other less natural classes (such as voiced segments); that is, an unnatural class might require fewer features to characterize than a natural class.
- Within a language, certain groups of rules that appear to be doing the same thing (such as strengthening or lenition) may have to be represented as different rules.
- Symmetric systems are not more easily represented than asymmetric ones.

- Certain feature combinations are more common than others, e.g. +voc implies +son.

The fact that Chomsky and Halle consider these observations as shortcomings of their framework shows that they considered the articulated UG presented to this point as not articulated enough to account for the attested data. That is, the model they posit does not sufficiently constrain the set of possible output grammars, allowing many types of unattested grammars and not favoring more commonly occurring phenomena.

An example of the first observation is presented here. Chomsky and Halle give examples of pairs of rules in which one rule is more common, yet is either just as or more complicated to represent, for example (7).

(7) (from SPE, p. 401)

(i) k → č / ____ [−cons, −back]

(ii) č → k / ____ [−cons, +back]

While palatalization, as in (7i), is a common phonological process, the alternation in (7ii) is not. However, in the SPE framework the representations are equally complex, and so they should be equally natural and likely. They clearly assume that the unequal occurrence of the phenomena described by (7i) versus (7ii) should be the result of properties of the grammar, as opposed to the result of extragrammatical factors. "All of these examples [shown in SPE pp. 401-2], and many others like them, point to the need for an extension of the theory to accommodate the effects of the intrinsic content of features, to distinguish 'expected' or 'natural' cases of rules and symbol configurations from others that are unexpected and unnatural (p. 402)." Because of the perceived shortcomings of their framework to that point, Chomsky and Halle suggest changes in their formal system so that natural phenomena are more simply represented than unnatural phenomena. They achieve this by positing several universal stipulations:

- In the lexicon, segments are not fully specified by their +, − values. Instead, a segment's features may be specified by *u* (unmarked), *m* (marked), +, and −. Features are specified as unmarked, except those necessary to identify the quality of the segment. If the value of some feature F is predictable from the value of feature G, then only G is specified. As a result, phonemes will have different complexities depending on their feature specifications, so an /a/ will be less complex than an /ü/, for example.
- Universal marking conventions will fill in the feature specifications.[4]
- If a feature value can be predicted from an adjacent segment or from a different feature in the same segment, then its value will be unmarked.

Let's look at a specific example. Vowels are typically sonorant, [+son], but the grammatical system of SPE up to this point does not reflect this fact. The model could just as easily generate a phonology where all vowels are [−son]. Because this would be a very unnatural phonological grammar, Chomsky and Halle perceive this as a problem. As a remedy, they integrate the idea of markedness into the theory so that the intrinsic content is encoded into the grammar. In the case of vowels, the universal marking convention in (8) stipulates that all vowels be sonorant:

(8) [+voc][5] → [+son] (from SPE, p. 404)

---

[4] Chomsky and Halle state that these universal conventions are not part of a language's grammar: "Being universal, these rules [the replacing of *u* and *m* by + and −] are not part of a grammar, but rather conventions for the interpretation of a grammar; they do not affect the complexity of a grammar as determined by the evaluation measure, any more than the rules for interpreting → or {}." The discussion clearly suggests that the conventions are part of UG.

[5] [+voc] stands for *vocalic*, which includes laterals in addition to vowels and glides.

In this way, each language does not have to specify that its vowels are sonorant, since it is universally specified a single time as a built-in part of UG. This marking convention allows for the uncommon occurrence of non-sonorant vowels, as in the case of devoiced vowels, for example. In that case, however, speakers of languages with devoiced vowels would have this stipulated explicitly in the grammar, which is not problematic since devoiced vowels are the exception. The use of marking conventions permits some kinds of non-occurring phonological patterns to be absent from possible grammars, thus creating a tighter fit between model and attested systems. In addition to these marking conventions, Chomsky and Halle suggest that rules be linked with marking conventions, thus positing additional built-in mechanisms to ensure that certain kinds of rules are favored by the model.

As we have seen, Chomsky and Halle propose changes in the universal component of their theory by positing a more articulated UG, so that the intrinsic content of the features, as well as the more natural phonological rules, are reflected in the grammar. However, in what sense do the features have intrinsic content, and in what sense are the rules more natural?

Let's first discuss the features. By assuming that the intrinsic content is stipulated in UG, we do not expect to find an account for this elsewhere. Yet for many of the examples Chomsky and Halle give, the intrinsic content of a feature is arguably due to its phonetic realization or due to perceptual factors. Take, for example, the case of sonorants being voiced by default while obstruents are not. A physical account of why sonorants are voiced is discussed in SPE (pp. 300-301) as follows (see also Westbury and Keating 1985, and Ohala 1995). Voicing requires, in addition to a particular vocal cord positioning, an air pressure difference across the larynx. A higher pressure in the lungs creates airflow past the vocal cords, which causes them to vibrate due to the Bernoulli effect and the elasticity of the vocal cords. Sonorants have an unobstructed vocal tract, so that a sufficient pressure difference across the vocal cords causes spontaneous voicing. Obstruents on the other hand have by definition an obstruction such that a pressure difference across the vocal cords is harder to maintain. This is because as the air passes across the vocal cords, either it has nowhere to go in the case of a total obstruction in stops, or it escapes slowly due to the partial obstruction in the case of fricatives. Either way, the pressure difference across the vocal cords becomes equalized, causing the air to cease flowing, and the vocal cord vibration—the voicing—to stop. What has been said for sonorants, is of course true for (voiced) vowels, a subset of sonorants. One may say that vowels have intrinsic content, but we have seen that at least one component of the intrinsic content is in fact a product of the physical layout of the vocal tract and of the laws of physics (Ohala 1995). This evidence makes it clear that vowels (and sonorants in general) are spontaneously voiced for reasons having to do with the physical world including the layout of the vocal tract. Chomsky and Halle recognize this, yet they nonetheless posit the markedness convention in (9), which essentially states that a vowel is voiced unless otherwise specified, as well as having the other features values listed on the right.

(9) A marking convention for vowels (from SPE, p. 405)

[**+voc** −cons] → [−ant −strid +cont **+voice** −lateral …]

The convention attributes the default voicing of vowels to the universal grammar. This creates a redundancy—the default voicing of vowels is now an inherent property of the universal grammar by stipulation, as well as a property of the physical world—and thus is a violation of Occam's Razor. Such a model gets the fact right—that vowels are typically voiced—but for the wrong reason; that is, it does not meet the criteria for an explanatory model. If the generalization that vowels are voiced by default is a general physical property of vowels, then we need not posit structure to support the same ends unless further evidence requires such structure. Instead it is sufficient to posit a UG that allows for grammars with all voiceless vowels. Then the absence of such grammars results from the physical reasons already mentioned, and not as a property of UG. If this is the case, then a more appropriate phonological model would be one that neither favored nor disfavored voiced vowels. Such a model would leave this to extra-linguistic factors such as those that favor spontaneous voicing. This model deliberately does not have a tight fit between the attested phonological data and the possible grammars it could generate.

1.3 Phonetics and phonology

Having discussed the parallels between phonology and phonetics, one may be inclined to suggest that these are in fact the same thing. The intent of this dissertation is not to reduce phonological patterns to phonetics, but rather to seek explanations for phonological patterns in the diachronic residue of phonetic phenomena. In order to do this, I first discuss the distinction between phonological and phonetic representations. Keating (1990), Cohn (1990), and Pierrehumbert (1990) argue that phonological representations are abstract, discrete, and timeless, whereas phonetic representations refer to physical dimensions and are temporal. As a result of these two different representations, the rules applying to each of the representations are also different in nature. While phonological rules are categorical, phonetic rules are partial, gradient, and variable (Keating 1990, Cohn 1993, Zsiga 1997, Myers 2000a). A number of recent studies have applied this view to show that processes previously considered phonological are actually phonetic processes. Cohn (1993) found that anticipatory vowel nasalization in English, often considered to be phonological, was actually a phonetic phenomenon, since it was a gradient rather than a categorical process. Likewise, Zsiga (1997) found that vowel assimilation in Igbo displayed the characteristics of a phonetic process. Sproat and Fujimura (1993) found that differences in [l] in English in different environments are phonetic, not phonological. Most researchers agree with the distinction between phonology and phonetics (Cohn 1998, Myers 2000a); however, Browman and Goldstein (1992), and in a different manner, Kirchner (1997, 1998), Steriade (2000), and Flemming (2001) consider phonology and phonetics to be the same domain.

In order to account for the differences between phonological and phonetic rules, Zsiga (1997) proposes that phonology and phonetics have different representations. In her view, the phonological representation, with articulatorily-based features and no time specifications, maps onto a gesture-based phonetic system with timing, based on the theory of Articulatory Phonology (see Browman and Goldstein (1992) and references therein). Note that Zsiga's approach departs from the standard theory of Articulatory Phonology in which the articulatory model handles both the phonetics and, to some degree, the phonology. In this manner, the phonology is able to represent categorical phonological alternations such as ATR vowel harmony in Igbo, while the gesture-based phonetic system can better represent gradient, partial, and variable effects such as vowel assimilation in Igbo. We want to be able to capture the categorical abstract representations in phonology, but also the gradient, variable, and non-universal properties of the output.

An example of the difference between phonological and phonetic variation comes from a study of coarticulation by Beddor and Yavuz (1995). They found more anticipatory (right-to-left) than carryover (left-to-right) coarticulation in Turkish (measured on disharmonic words), even though the direction of the vowel harmony is left-to-right. That is, the *phonetic* effect of coarticulation was more prominent in the direction opposite from the direction of the *phonological* assimilation. If the position put forth here is correct, carryover coarticulation developed into a left-to-right vowel harmony. At this time, the carryover coarticulation is no longer prominent, though the harmony that it induced remains.

In SPE, the linguistic sound system is split into two distinct parts: the phonology and the phonetics. In this view, the application of *language-specific* rules on underlying forms creates surface forms, the output of phonology. Although the feature *values* of the underlying and surface forms of phonology could change as a result of these rules, the essential structure of the two is assumed to be identical; that is, both the surface form and the underlying form consist of a bundle of all of the universal features, completely specified. The surface form—the input into the phonetic component of the language—is mapped by universal rules of implementation into the phonetic output. The mapping of phonology to phonetics consisted of some variation mediated by language-specific phonetic values assigned to a feature value. For example, a given feature value, say [+round], could have a phonetic value indicating extreme rounding for one language, and a different phonetic value indicating weak rounding for another language. In this way, Chomsky and Halle account for cross-linguistic differences in *phonetic* implementation by using phonetic (multivalued) features, while accounting for *phonological* differences by using less specific (binary) phonological feature values. It is important to note that this aspect of their interface from phonological surface form to phonetic form is a simple one—for any language, each phonological feature value of a surface form is mapped to one phonetic feature value. In addition to this mechanism, Chomsky and Halle posit that languages have an articulation base, which describes the characteristic qualities of the sound of a language that distinguish it from other languages. Bradlow (1993:24) found, for example, that certain vowels of Spanish and English are located differently within the acoustic vowel space. Finally, Chomsky and Halle refer to the "universal rules" that account for coarticulation effects, among other things. As the focus of SPE is on phonology and not phonetics, it does not seem worthwhile to determine what aspects of cross-linguistic phonetic variation can be attributed to the base of articulation, to universal rules, or to phonetic feature values.

Thus the mechanisms of phonology and phonetics differ in the SPE model. Consequently, the extent to which the pattern of phonology and phonetics have similarities (such as those mentioned in §1.1) is not predicted or accounted for by the model. Stepping back, it is clear that a central motivation behind positing the structure of the SPE phonological model is to eliminate or reduce redundancy in the theory. If a phenomenon is seen as universal then it is encoded into the model so that its specification is not required in the grammar of each language (and likewise is not required in the grammar of each individual speaker.) This is the case for the positing of features, marking conventions, and linking rules. However, SPE does not account for the redundancy implicit in the parallels between phonology and phonetic phenomena, not a trivial exception, and so the parallelism is an unexplained accident.

In summary, the universal phonological patterns in SPE are attributable to UG via marking conventions, while the universal phonetic patterns are attributable to the separate universal phonetic rules mentioned above. While formal linguistic properties apply to phonological phenomena, they do not apply to phonetic ones. SPE is weakened because it does not account for this parallel. In summary, by incorporating common phonological characteristics and rules into the structure of their model, Chomsky and Halle create an articulated UG. They outlined some problematic areas for future investigation. While they succeeded in attaining a relatively tight fit between model and attested grammars, they have not explained why their model relies entirely on formal mechanisms for constraining the possible grammars, instead of allowing extragrammatical factors, such as phonetic ones, to play a role, factors they themselves recognize and discuss. In addition, SPE does not account for the closely related problem of the unexplained parallels between phonological and phonetic phenomena.

1.4 Resolving the problem of redundancy between phonology and phonetics

The apparent redundancy is resolved if, instead of positing mental constraints on the possible phonologies, we focus on the historical origin of the phonological patterns: the interaction between phonetics, perception, and other cognitive processes. This view has been advanced in many studies, most notably by Ohala (1990, 1993, 1994, 1995, and elsewhere), but also by Myers (1997), Hale (1999), Hale and Reiss (2000), Silverman (2000), and Buckley (2000). Ohala (1995: 58) states:

> The existence of phonetically natural processes in the sound patterns of languages needs no special or extravagant explanation. Universal, physical phonetic factors lead to a speech signal which obscures the speaker's intended pronunciation; listeners may misinterpret ambiguous phonetic elements in the signal and arrive at a pronunciation norm that differs from the speaker's. This is how sound change works and how natural sound patterns arise. Such changes will reflect phonetic constraints without speaker or listener having to know about them.

Examples of phonological alternations with proposed origins in phonetic processes are found in Hombert (1978), Hyman (1976), Beckman, De Jong, Jun, and Lee (1992), Ohala (1993) and Ohala (1995). In these cases, a perceivable "automatic" phonetic effect, after time, is reinterpreted as an intended effect and thus becomes part of the phonology of the language, termed "phonologization" by Hyman. Under this view, the patterns commonly attributed to inherent aspects of phonology are due to the grammaticalization of phonetic-perceptual effects.

Hyman (1976) discusses how phonetic effects become phonologized in the creation of a new tone from a high tone in Pekinese (his data come from Haudricourt, 1961 and Matisoff, 1973). The history of this development is shown in Table 1.2.

Table 1.2. Pekinese Tone example (adapted from Hyman 1976)

Stage I	Stage II	Stage III
pá	pá	pá
bá	bǎ	pǎ

In an earlier stage in the language, Stage I, both voiced and voiceless stops may be followed by a high toned vowel. The voicing variation between [pá] and [bá] is a distinctive phonological difference. It is not predictable from any other part of the utterance; that is, the speaker's brain must make a different command (however this is transmitted) for voicing or voicelessness, hence Hyman calls the voicing difference intended. The high tone on the vowel contrasts with two other tones (not shown here). However, the fundamental frequency of the vowel in [bá] is slightly lower than that of [pá], as is universally the case following voiced consonants compared with voiceless (see Hombert, 1978). The intrinsic pitch difference is a result of the voicing of the stop and is not part of the phonology of the language. It does not arise from a specific command from the speaker but rather is a by-product of the state of the vocal cords of the preceding consonant. The difference in pitch is small but perceivable to the listener, and may provide a cue about the voice quality of the consonant. At Stage II, a later historical stage in the language, the pitch difference has become phonologized. At this point the difference between the pitch of the vowels following the different consonants can no longer be attributed solely to the voice quality of the consonant. The pitch difference has become an extrinsic variation; that is, the speaker now delivers a different intended command for each type of vowel. The voicing distinction remains, so that this phonological contrast is realized by both voicing and tone. Finally, in Stage III, the voicing differences of the initial consonant have been neutralized by a separate process, and the tone differences have become the sole distinctive element. What started out as a phonetic effect with a physical origin, became a phonological effect whose physical origin ultimately became obscured. According to Hombert (1978), the loss of voicing leading to a tonal distinction is the most commonly documented type of tonogenesis.

The mind of the speaker plays a role in determining what kinds of phonetic effects become phonologized, since clearly properties of the mind determine whether a phonetic effect is perceivable. For example, if the pitch difference between vowels following voiced and voiceless consonants mentioned above were not perceivable, then speakers could never generalize the pitch difference as being a distinguishing cue, and so phonologization of the pitch difference would never occur. Clearly, an intrinsic cue must be salient, at some level, if it is to have a chance of becoming phonologized. However, this saliency alone does not assure phonologization; indeed, most salient intrinsic cues do not become phonologized. That is, of the many perceivable intrinsic effects present in a language, most are not becoming phonologized, at any given time, and clearly some kinds of intrinsic effects are never phonologized. While some phonological alternations, such as intervocalic voicing of stops, are quite common, other types of intrinsic cues rarely if ever become contrastive. For example, in most or all languages, high vowels have a higher f0 than low vowels (Peterson and Barney 1952, Lehiste and Peterson 1961, and Whalen and Levitt 1995). Even though the pitch difference is perceivable, tonal contrasts attributed to vowel quality differences are not widely attested. Explanations for why certain phonetic effects become phonologized are not discussed here.

If we accept this account of phonologization—that phonological alternations start out as phonetic changes which eventually get phonologized—then by Occam's Razor we should not posit constraints in an explanatory phonological model to explain the patterns of phonology, rather the explanation lies in the patterns found in phonetics and perception. Under this account, phonology is not concerned about the naturalness of phonological alternations; the apparent naturalness can be attributed to the phonetic origins of the alternations. And so the models of phonological (mental) grammar will necessarily be less constrained than what we have seen, because the tightness of the fit between possible systems generated by a model and attested grammars will not reside wholly in the model. Certain grammars that could exist do not, not because UG does not allow them, but rather because these type of grammars could not occur historically (or more precisely, are much less likely to occur). Figure 1.3 schematizes the space of possible grammars under this view (c.f. Figure 1.2). Universal Grammar has a smaller role in constraining the grammar space. The ellipse representing the attested systems is constrained outside of the model of grammars.

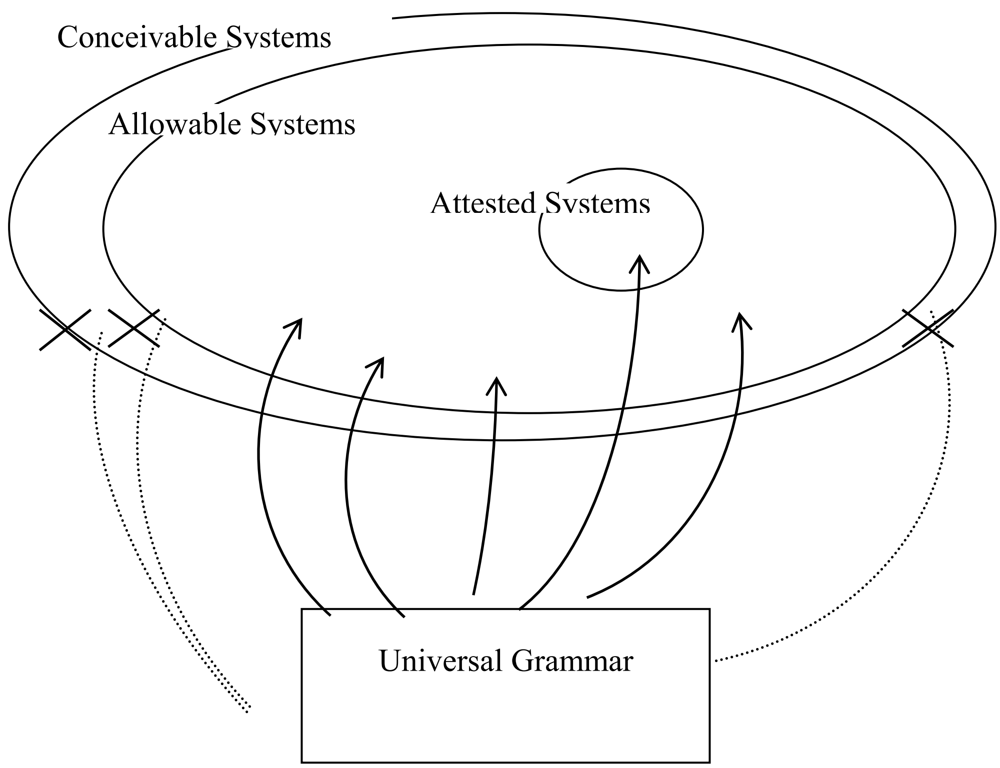


Figure 1.3. Updated model of Universal Grammar with weaker constraints.

Returning to SPE, we can see that the objection raised by Chomsky and Halle (see §1.2 above)—that the SPE framework does not favor natural phenomena—is in fact not problematic under this view. The naturalness of phonological systems should remain outside of the formal system, such that unnatural systems could be represented by the phonological formalism, but these unnatural systems are unlikely to occur. The same reasoning can be applied to proposals in feature geometry (Clements, 1985; Sagey, 1986; McCarthy, 1988; Odden, 1991; Halle, 1995; and Clements and Hume, 1995). Since these proposals typically posit a mental apparatus that effectively constrains the allowable phonological grammars, then these structures become redundant if these patterns can be attributed to phonetic factors. The success of Feature Geometry (FG) to model some phonological phenomena is then not because FG is knowledge, but rather because FG can—to some degree—model phonetic phenomena, the source of phonological phenomena; that is, Feature Geometry plays no role in a speaker's phonology (Ohala 1995).

In Optimality Theory (OT), as proposed by Prince and Smolensky (1993), phonological patterns emerge as an interaction of constraint rankings on an input representation. The constraints are universal; it is the ranking of the constraints that varies from language to language:

> The basic idea we will explore is that Universal Grammar consists largely of a set of constraints on representational well-formedness, out of which individual grammars are constructed. […] The conception we pursue can be stated, in its purest form, as follows: Universal Grammar provides a set of highly general well-formedness constraints. These often conflicting constraints are all operative in individual languages. Languages differ primarily in how they resolve the conflicts: in the way they rank these universal constraints in strict dominance hierarchies that determine the circumstances under which constraints are violated (Prince and Smolensky, 1993: pp. 2-3).

The constraints are formal, part of the formal grammar, and hence not part of the physical world. Alternatively, if we accept that phonology is grammaticalized phonetics, then this would suggest that the constraints are neither innate nor universal, but instead are learned solely by exposure to speech data from the learner's language or languages. In this case, only those constraints (or patterns) evident in the language would be learned. If, for example, the language had no codas at all, then the learner would be not be exposed to coda devoicing, so it would not be necessary for the learner to posit this constraint. In this view, if a phonological pattern were present in the language, then the learner would deduce the necessary constraint from the data she is exposed to (in the most general way that fits the data). If the phonological pattern were not present, then the learner would have no evidence to infer a constraint. Of course, determining whether the pattern was present or not would reside in the mind of the learner. If she thought the pattern were present because of some confusion due to phonetic effects, then the learner might confuse the accidental phonetic effects as being intentional, thus inferring a phonological pattern not present in the grammar of the speaker. For example, if word final obstruents were phonetically devoiced by speakers of a language, the learner may incorrectly infer that voiced stops are devoiced word-finally, thus positing a NoVoicedCoda constraint. In this view, the parallel between the phonological and phonetic phenomena is straightforwardly explained—phonological patterns emerge from phonetic ones. Under this view, many constraints would have apparent phonetic motivation, but only because of the phonetic origin of the phenomena on which the generalizations were made.

1.5 Other ways to solve the phonology/phonetics problem

Another approach has been suggested by Hayes (1999) in which he argues that OT constraints could be learned through induction by infant learners experimenting with making sounds. In this sense, the constraints are not strictly an innate component of Universal Grammar as I have defined it above. The universality of the constraints is then due to the (relative) universality of the vocal tract's physical structure, the perceptual system from the use of which the learner posits the constraints, and the laws of physics. So a constraint will be posited when a learner has evidence from her own experimentation, not necessarily evidence from the language data presented to her. Hayes's view is similar to the view expressed in Donegan and Stampe (1979), who also see much of phonology coming from the learner:

> The mysterious perfection of this childhood learning remains a mystery, but we can hope to make the task seem slightly less awesome by pointing out that most phonological alternations and restrictions are motivated by the nature of the learner rather than the language […]. The German child does not have to learn to devoice all and only the class of word-final obstruents, nor does the Vietnamese child have to learn to avoid coining words that end in voiced obstruents: these are natural restrictions. For a minority of languages, including English, children must learn to pronounce words with voiced final obstruents. This is obviously not easy, but is something which obviously can be accomplished by children (p. 140).

Hale (1999) argues against such a view. He asks us to imagine a child born with a mutation in her vocal tract such that it is not difficult for her to voice obstruents in codas. This child then would not infer the constraint NoVoiceCoda by experimentation. Instead we might assume that if this child were exposed to German, in which obstruents are devoiced word-finally, the child would learn precisely the language she was exposed to, obstruent devoicing not excepted.

1.6 UG is an argument of last resort

Ever since SPE, the predominant assumption is that universal patterns exist as a result of an articulated Universal Grammar (UG). The reasoning is certainly logical—lacking any tenable alternative explanation for universal patterns, and assuming their existence is not accidental, linguists posit one structure, albeit complex, shared by all humans, that could account for all patterns (see Anderson, 1981). Determining the structure of UG has been a continuous project. Linguists continue to find generalizations regarding the universality of phonological patterns. But while they have attained a tighter fit between attested data and the output of their models, the arguments in favor of an articulated UG have not changed. That is, generalizations regarding patterns have been refined, but UG itself has not been corroborated. This lack of corroboration is not, in itself, an argument against UG. However, it should be clear that this line of reasoning in favor of UG is an argument of last resort, only convincing if we *lack any tenable alternative explanation for universal patterns*. We find support of this point in Anderson (1981). He argues in favor of an articulated UG, and against a reductionist approach (in particular that of Donegan and Stampe (1979), in which the patterns of phonology are largely reduced to domains outside of the grammar); nonetheless, he states that "[…] in considering areas of phonological structure […], we can only determine that some property is to be attributed to the essential nature of language if it does not seem to have an account in more general terms. On this view, it is still very much part of the business of phonologists to look for 'phonetic explanations' of phonological phenomena, but not in order to justify the traditional hope that all phenomena of interest can be exhaustively reduced in this way (1981: 497)." To the extent that we can find explanations for universal patterns in domains outside of linguistics, then our conception of UG must be weakened. By this view, it would be unjustified for a phonologist to believe that we should not bother looking for extragrammatical explanations because we have already proven the existence of an articulated UG. By the same token, if no explanation were found in extragrammatical domains, then, as an argument of last resort, we would be forced to posit an articulated UG.

1.7 Summary

Parallels between phonological and phonetic phenomena support the view that phonology emerges from phonetics. In my study, I examine high vowel harmony in Àkùrẹ́ Yorùbá, and compare it with two dialects that lack this phonological pattern. In an examination of both phonological and phonetic interaction of adjacent vowels, I show that although the two phenomena are qualitatively different, they share certain characteristics. I propose that the characteristics found in coarticulation provide the basis for the eventual generalization of this pattern into vowel harmony, and present a hypothesis for how this occurred in Yorùbá. As discussed above, most nativist theories of phonology have no explanation for the similarities of phonology and phonetics.

1.8 Overview of dissertation

In Chapter 2, I discuss the phonology of Àkùrẹ́, Mọ̀bà, and Standard Yorùbá. Since an important characteristic of Yorùbá is the presence of Advanced Tongue Root [ATR] harmony, I give an overview of ATR harmony. Being relatively closely related dialects, Àkùrẹ́, Mọ̀bà, and Standard Yorùbá are quite similar in phonological structure, with one notable exception being the high vowel harmony found in Àkùrẹ́, but not in the other two dialects. I also discuss historical studies that are consistent with the hypothesis that Àkùrẹ́ high vowel harmony emerged from vowel to vowel coarticulation. In Chapter 3, I discuss the methodology of the primary experiment in which I investigate the acoustic characteristics of vowels in the three dialects. Several acoustic measures are evaluated for use in measuring the harmony and vowel coarticulation. F1 is determined to be the best measure for determining ATR contribution. Chapter 4 looks at the results of the experiment in each dialect, as well as a detailed examination of the vowels. The phonological realization of vowel harmony in all three dialects is examined. In addition, the high vowel harmony of Àkùrẹ́ and the coarticulation in Mọ̀bà and SY are examined in detail, with a comparison showing that for high vowels coarticulation resembles vowel harmony, especially for the vowel /i/. In chapter 5, I propose a model that can induce the high vowel harmony pattern from this study's Àkùrẹ́ data. More interestingly, the model succeeds at learning—to a large degree—the high vowel harmony pattern from Mọ̀bà and SY, the dialects without high vowel harmony. The decision tree model used does not require any reference to features or natural classes. I argue that it is not necessary to posit features as a prerequisite to learning a phonological pattern, not as an explanation for universal patterns.

CHAPTER TWO

YORÙBÁ PHONOLOGY

2.1 Introduction

In this chapter, I present a description of the phonologies of Àkùrẹ́, Mọ̀bà, and Standard Yorùbá as they relate to the study of vowel to vowel interactions. This serves as a background for the phonetic experiment discussed in the following chapters. At the same time, the facts of the vowel systems of Àkùrẹ́ and Mọ̀bà Yorùbá are interesting in their own right since they have not been extensively documented. I also present an historical picture to support the claim that proto-Yorùbá did not have high vowel harmony, and that it emerged at some point in the branch of Yorùbá that includes Àkùrẹ́.

This chapter is laid out as follows. Section 2.2 provides some background into the Yorùbá dialects of this study. Section 2.3 is a discussion of [ADVANCED TONGUE ROOT]—the feature and its phonetic attributes. Section 2.4 sketches the vowel and consonant inventory for the three Yorùbá dialects. In the subsequent three sections (§2.5-2.7), the relevant phonologies of Àkùrẹ́, Mọ̀bà, and Standard Yorùbá are laid out. Section 2.8 discusses evidence supporting a seven surface vowel system in Proto-Yorùbá. The summary of the vowel harmony patterns of the three dialects is presented in section 2.9.

2.2 Language background

Linguists are familiar with the use and misuse of the term *dialect*. From a linguistic standpoint, Standard Yorùbá, Àkùrẹ́ Yorùbá, and Mọ̀bà Yorùbá are considered dialects with respect to each other because they are to a large degree mutually intelligible. At the same time, each dialect is a language in its own right, a system with all the linguistic attributes of a language. Linguistically, there is nothing inherent to these systems that makes one a language and the others dialects of that language. However, politically and culturally, a dialect is considered a language by virtue of its status in politics and society. For this reason, Standard Yorùbá is referred to as a language by virtue of its status in Yorùbá society, while Àkùrẹ́ and Mọ̀bà (and for that matter Standard Yorùbá itself) are considered dialects of Yorùbá.

The many dialects of Yorùbá are spoken in southwestern Nigeria, as well as in neighboring countries, Benin and Togo. Standard Yorùbá (SY) is closely related to the Ọ̀yọ́ dialect and is spoken in Ìbàdàn and in much of Yorùbáland in Nigeria. In many of the Yorùbá areas where SY is not a first language, it is spoken as a lingua franca (Capo 1989). Àkùrẹ́ Yorùbá is spoken in the city of Àkùrẹ́, Òndó state, Nigeria; and Mọ̀bà is spoken in and around the village of Ọ̀tụ̀n-Èkìtì, Èkìtì state, Nigeria. Figure 2.1 shows how these and some other Yorùbá dialects are related to each other.

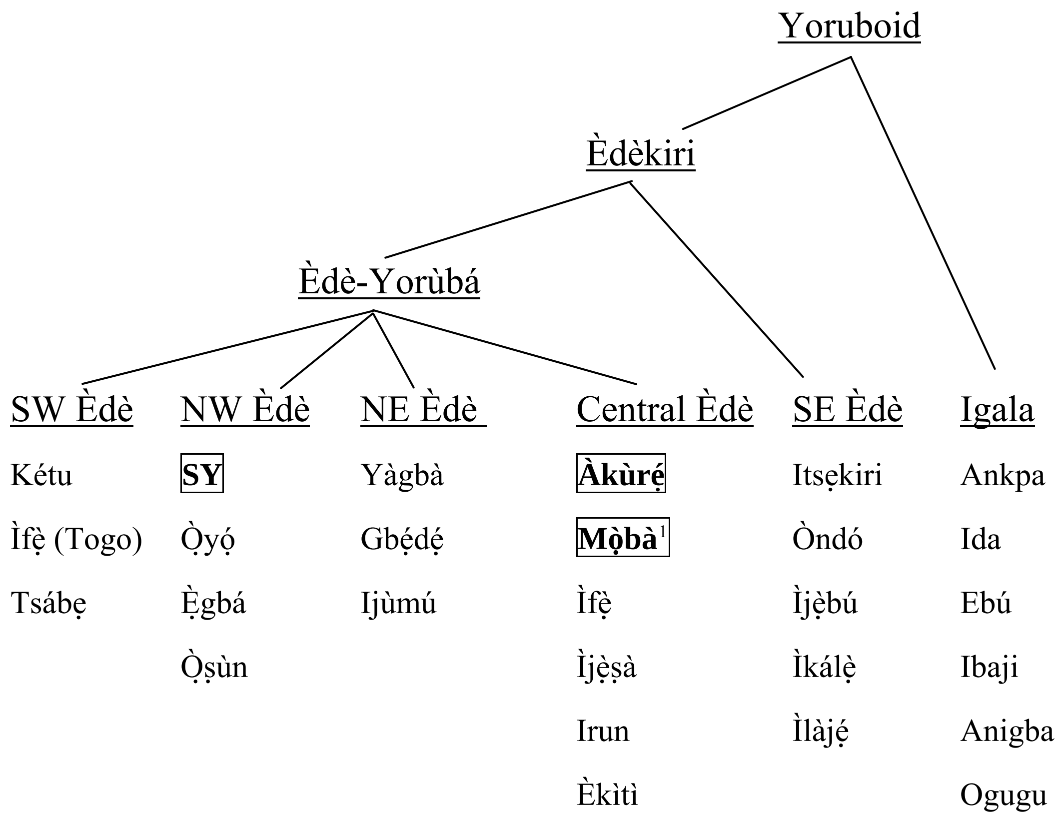


Figure 2.1. Classification of Yorùbá dialects based on Capo (1989), which in turn is based on Akinkugbe (1978), Williamson (1982), and other references therein.

Èdè-Yorùbá is divided into four branches labeled South West Èdè, North West Èdè, North East Èdè and Central Èdè—*èdè* being the word for 'language' in many Yorùbá dialects. Another branch, South East Èdè, is more distant than the others, with a higher label Èdèkiri uniting the five branches. The Èdèkiri branch and the smaller Igala branch together form the Yoruboid group. The Yoruboid group is part of the Benue-Congo[2] group of the Niger-Congo family (Bennett and Sterk 1977, Williamson 1989, Capo 1989). This study focuses on the dialects Standard Yorùbá, Àkùrẹ́, and Mọ̀bà for reasons that are laid out in §3.1. Standard Yorùbá is in the North West Èdè branch of

[1] Mọ̀bà Yorùbá should not be confused with the Moba language (Gur, Niger-Congo) of Togo and Burkina Faso.
[2] Yoruboid languages were formerly considered to be in the Kwa group of the Niger-Congo family.

Èdè-Yorùbá. (While SY was formerly referred to as Central Yorùbá (CY), in Oyelaran (1973) for example, there is no particular connection between SY and Central Èdè). Àkùrẹ́ Yorùbá is in the Central Èdè branch. I have place Mọ̀bà Yorùbá in the Central Èdè branch, following a brief mention in Capo (1989) in which Mọ̀bà is said to be a sub-dialect of Èkìtì due to a reference from Bamiṣilẹ (1987). However, I was not able to find Bamiṣilẹ's (1987) manuscript. Additional discussion of the placement of Mọ̀bà within Central Èdè is found in §2.8.

2.3 Advanced Tongue Root

These three Yorùbá dialects, as well as many or all Yorùbá dialects, exhibit Advanced Tongue Root harmony to some degree or another. The term Advanced Tongue Root, or ATR, was coined by Stewart (1967) to describe a distinguishing feature of sets of vowels, mostly in languages of West Africa, exhibiting certain vowel co-occurrence patterns. ATR languages have two sets of vowels: one set [+ATR] is characterized articulatorily by a wide pharyngeal volume, and the other set [−ATR] by a narrow pharyngeal volume. In these languages, vowels within a certain domain generally come from either the [+ATR] set or the [−ATR] set.

An example from Cahill (1996) is the vowel system of Kɔnni, spoken in Ghana, shown in (1).

(1) Kɔnni (Gur, Niger-Congo) from Cahill (1996). The [+ATR] vowels are [e, o, i, u] and the [−ATR] vowels are [a, ẹ, ọ, ị, ụ]. A $\boxed{\text{box}}$ indicates the root.

[+ATR]		[−ATR]	
$\boxed{\text{tígí}}$-rí	'the house'	$\boxed{\text{kụ̀ụ̀}}$-rị́	'the hoe'
$\boxed{\text{sìè}}$-kú	'the path'	$\boxed{\text{nị̀ị̀}}$-kụ́	'the rain'
$\boxed{\text{kùrì}}$-yé	'has pounded'	$\boxed{\text{pàsị̀}}$-yá	'has peeled'
bè $\boxed{\text{yè}}$-yé	'they have seen.'	bà $\boxed{\text{yị̀}}$-yá	'they have given.'

In each case, the [ATR] value of the root spreads in either direction, so that, for example, the definite article suffix is realized as [+ATR] [rí] after a [+ATR] root, *tígí*-, ‘house’, but as [−ATR] [rị́] after a [−ATR] root, *kụ̀ụ̀*-, ‘hoe’. The [ATR] value can also spread to subject clitics, as *bè* ‘(3rd person singular)’ is found before a [+ATR] root, but *bà* is found before a [−ATR] root.

The symbol from the International Phonetic Alphabet (revised to 1993) for [+ATR] vowels is ‘ ̘ ’ as in [o̘] and [e̘], and for [−ATR] vowels is ‘ ̙ ’ as in [o̙] and [e̙]. Because these diacritics are often difficult to differentiate, I follow the often used convention of marking [−ATR] vowels with an underdot as in /ẹ/, /ọ/, [ị], and [ụ], while leaving [+ATR] vowels unmarked. This is consistent with the orthographic convention used for Yorùbá where [−ATR] mid vowels are marked with an underdot, as in ‘ẹ’, ‘ọ’. The symbols ‘ị’ and ‘ụ’ are not found in Yorùbá orthography since Standard Yorùbá does not have [−ATR] high vowels. Mọ̀bà also does not have [−ATR] high vowels. Although Àkùrẹ́ Yorùbá does not have underlying [−ATR] high vowels, it does have [−ATR] high vowels allophones, for which I use the symbols [ị], and [ụ]. The [+ATR] vowels are left unmarked in Yorùbá orthography, and here. The low vowel /a/ is [−ATR] in all three dialects, and it is left unmarked. See §2.4.1 below for a summarized chart of oral vowels in these Yorùbá dialects.

A characteristic often found in ATR harmony systems is cross-height harmony. It refers to a vowel system in which a harmony works both within and across vowel heights. The target and trigger vowels may have the same height: for example, a [−ATR] high vowel such as [ị] may lower a [+ATR] high vowel such as [u] to the [−ATR] high vowel [ụ]. Or, the target and trigger vowels may have different heights, such that a [−ATR] high vowel such as [ụ] may also lower a [+ATR] mid vowel such as /o/ to a phonetically lower mid vowel [ọ]. In this case, a high vowel acts as a trigger to lower a vowel that is already lower than itself. This is shown in (2) from the Asante dialect of Akan (Kwa, Niger-Congo) where the mid vowel /o/ of the initial word ‘he’, becomes [−ATR] [ọ] before the [−ATR] high vowel [ị].

(2) Asante dialect of Akan from Clements (1981). A $\boxed{\text{box}}$ indicates the root.

o-$\boxed{\text{fiti}}$-i	‘he pierced (it).’
ọ-$\boxed{\text{cịrẹ}}$-ị	‘he showed (it).’

What is not common in an ATR harmony system is for a target vowel to change from one height to another (for example, mid to high) due to the influence of an adjacent vowel (Parkinson 1996).

Most typically, the term ATR is used to divide vowels into two classes based on phonological characteristics, as above. However, in addition to the phonological meaning of [ATR], there is an articulatory one. We may speak of the articulation of the participating vowels as being either [+ATR] or [−ATR]. Articulatorily, [+ATR] vowels are characterized by a large pharyngeal cavity compared to [−ATR] vowels. The pharyngeal volume is enlarged due to a combination of the tongue being moved forward in the mouth (and hence, the tongue root being advanced), the larynx being lowered, and the back pharyngeal wall being moved back and outside (Lindau 1975, 1978, 1979). Figure 2.2 shows x-ray tracings following Lindau (1978) of eight vowels from a single male speaker of the Akyem dialect of Akan. For each ±ATR vowel pair, the [+ATR] vowel, when compared with its [−ATR] counterpart, has the tongue root more forward, the tongue blade higher, and the larynx lowered.

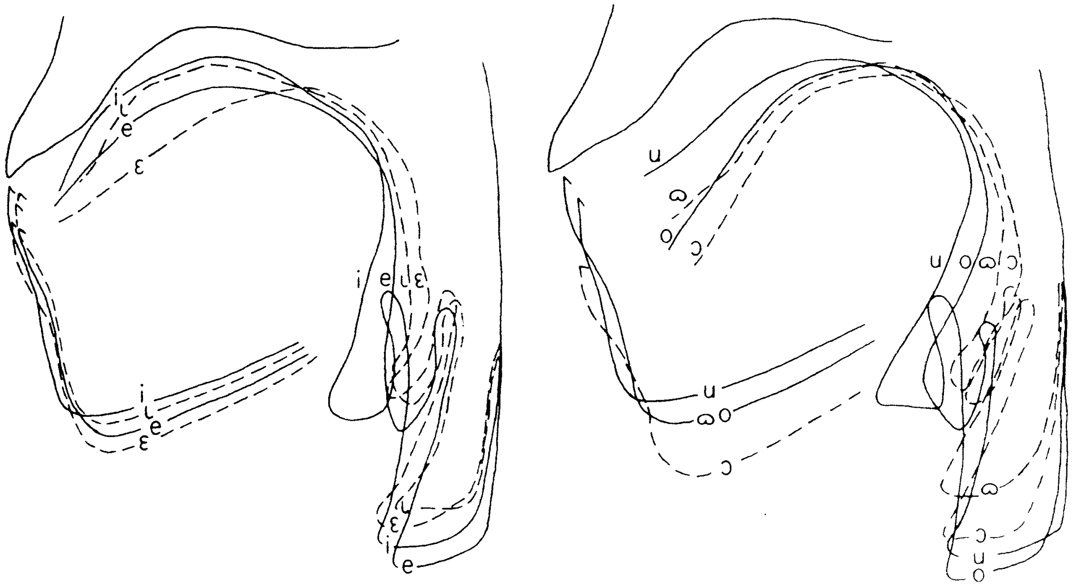


Figure 2.2. Tracings of four front vowels (left) and four back vowels for male subject of the Akyem dialect of Akan (from Lindau 1978). The vowels [i], [e], [u], and [o] are [+ATR]. The [−ATR] counterparts are [ɩ], [ɛ], [ɷ], and [ɔ], corresponding to [ị], [ẹ], [ụ], and [ọ].

Similar results were found in x-ray tracings of an Igbo speaker by Ladefoged (1964) and in an MRI study of an Akan speaker by Tiede (1996).

In both the phonological and articulatory connotations, there is a [+ATR] and a [−ATR] class of vowels. Phonologically, the term ATR refers to a type of vowel harmony, as well as the sets and features of the vowels involved in the harmony. A language may be said to exhibit ATR harmony, as in (1) for Kɔnni, when it exhibits the phonological attributes of an ATR harmony system, even though the vowels of that language have not been instrumentally determined to have the phonetic qualities of ±ATR vowels. This is not to say that labeling a language as having an ATR harmony based solely on the phonology is incorrect. However, it is important to keep in mind that the term ATR can be used with a phonological intent or a phonetic one. Although the phonological and articulatory characteristics are known to co-occur, this is not necessarily always the case. In Igbo and Akan, where articulatory studies have been carried out, the phonological and phonetic characteristics of ATR discussed above pattern together. Acoustic studies of other languages with ATR type harmony indicate that ATR vowels share certain acoustic characteristics. In general, the use of the term ATR, suggests both meanings of the term; however, it is not certain that this is the case for all languages exhibiting ATR-like harmony patterns. That is, we may find ATR-like harmony patterns with vowels that do not share the articulatory characteristics described above; in the same way, it is likely that languages exist whose vowels articulatorily resemble those of Igbo or Akan, but which lack ATR-like harmony patterns.

As pointed out by Clements (1991), the feature [ATR] has also been used in phonological analyses as a cover feature, as in analyses of Esimbi (Hyman 1988) and Sesotho (Harris 1987). In these cases, the phonological patterns are not typical ATR-type harmonies. Nor in these cases is there apparent evidence of ATR articulation. Rather the [ATR] feature is used as an additional phonological vowel feature which permits an elegant analysis. It is not clear why the feature [ATR] is used for these analyses as opposed to a new vowel feature, except that an assumption underlying these analyses, rooted in SPE, is that a limited number of features are universally available to the language learner. One might argue that since [ATR] has already been motivated elsewhere for other languages, it can be freely used in these cases also. If the need to limit the number of phonological features is not essential, then, barring further evidence, there is no need to consider as ATR systems the vowel systems discussed in Hyman (1988) and Harris (1987).

Even before imaging techniques, linguists were able to observe that pharyngeal volume was involved in what we now call [+ATR] and [−ATR] vowels. Pike (1947, cited by Stewart 1967) suggested that [+ATR] vowels were produced by lowering the larynx, advancing the tongue root, and spreading the faucal pillars. Trubetzkoy (1939) cites Dr. A. N. Tucker who studied Nilotic languages: "'squeezed' vowels: the faucal pillars are compressed and the velum is lowered... the 'breathy' vowels the velum is raised, the fauces retracted, and the larynx clearly lowered, so that quite a large cavity is formed behind the oral cavity proper. The glottis appears to be in the position of whispering." These observations were later verified instrumentally, as discussed above, though on a small number of speakers from a small number of languages. The phonological patterns of ATR languages were observed before imaging technology allowed for measurement of articulatory positions. Christaller (1875, in Westermann and Ward 1933) observed the vowel co-occurrence patterning in Twi and Fante. Westermann and Ward (1933) described vowel patterns in Igbo, though not with a harmony metaphor: "This curious usage is possibly due to the fact that the vowels are extremely difficult to differentiate from each other. [...] In the same way the fact that two vowels with neighbouring tongue positions do not occur in one word is probably an unconscious means of preventing vowels from falling together." The two sets have been referred to as [non-covered]/[covered] (Chomsky and Halle 1968), [Advanced Tongue Root]/[Retracted Tongue Root] (Stewart 1967), and [Expanded]/[Constricted] (Lindau 1979), among many other names. Instead of [ATR], Lindau's term [EXPANDED] is a more accurate label for the feature, because the two sets of vowels differ not solely in the relative advancement of the tongue root, but in the expansion of the laryngeal volume more generally. That is, in addition to an advanced tongue root, the [+ATR] vowels may also have a lowered larynx and expanded pharyngeal walls. However, because the term [ATR] is now in general practice, I use it here.

Acoustically, the two sets of vowels differ primarily in F1, as is seen for Akan in Figure 2.3 from Lindau (1979): the F1 of the [+ATR] vowels are lower than their [−ATR] counterparts. In addition, the F2 of the [+ATR] vowels is also typically higher than the F2 of the [−ATR] vowels, though the back mid-vowels occasionally do not show this effect.

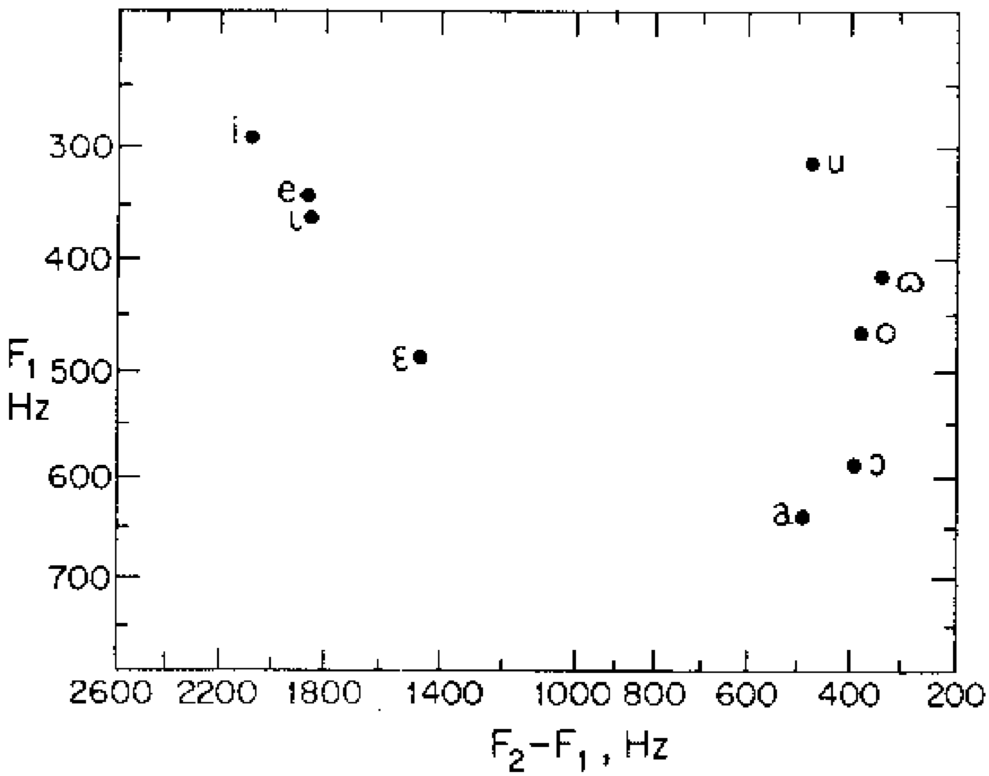


Figure 2.3. Formant chart of Akan vowels from Lindau (1978), an average from spectrograms of four speakers with five tokens for each vowel. The [−ATR] vowels [ị, ẹ, ụ, ọ, and a] are represented here as [ι, ε, ω, ɔ, and a], respectively; the [+ATR] vowels are [i, e, u, and o].

Hess (1992) finds that the first formant (F1) and first formant bandwidth are the best acoustic correlates of ATR in one speaker of Akan; that is, the ±ATR vowels pairs {i/ị, u/ụ, e/ẹ, o/ọ} are best distinguished acoustically by these two measures, with [+ATR] vowels having lower F1s and narrower F1 bandwidths than their [−ATR] counterparts.

Halle and Stevens (1969) and Kenstowicz (1994) argue that the tense-lax distinction found in English, German, and other languages is primarily an ATR distinction, as opposed to height or length. Lindau (1979), Stewart (1967), Ladefoged and Maddieson (1990), and others argue that the ATR feature is not involved in the tense/lax distinction found in other languages. Lindau finds that the two phenomena do not behave the same way articulatorily nor acoustically. Citing evidence from Perkell (1971), she points out that American English tense/lax vowels do not have a difference in larynx position, rather the main difference appears to be height. In addition, in English, the lax vowels are more centralized than their tense counterparts, while in Akan, the corresponding pairs, particularly the back vowels, differ mostly in F1. Impressionistically, mid vowels in ATR systems may resemble mid vowels in tense/lax systems, however, [−ATR] high vowels, in Yorùbá and Igbo at least, in no way resemble lax high vowels in English and German.

For detailed discussions of ATR vowels, see Ladefoged (1964), Stewart (1967), Halle and Stevens (1969), Lindau (1979), Hess (1992), Fulop (1996), and Casali (2000).

In the case of Yorùbá, phonological patterns of ATR harmony are found to a differing degree in each of the three dialects examined here. These are discussed in detail below. Cross-height vowel harmony is found in Àkùrẹ́ (see §2.5 below) and related dialects, though not in closely-related Mọ̀bà, and not in Standard Yorùbá. Mọ̀bà exhibits a more limited form of harmony. Standard Yorùbá exhibits the least amount of harmony. In fact, although for the past few decades SY has been described as an ATR language, it is not clear if the phonological use of the terms ±ATR is merited, since there are arguably no productive alternations in SY (see §2.7). However, the presence of SY word internal distributional patterns consistent with harmony indicates either an active harmony within the domain of the word, or an historical remnant of a harmony in an earlier time of the language, or perhaps both. The phonologies of surrounding NW Èdè and SW Èdè dialects, which have productive harmony evident in the subject clitics, strongly suggest that an earlier language once exhibited a productive ATR harmony. While the harmony may no longer be productive, it may be that the SY vowels still exhibit the articulatory characteristics of ATR vowels. There is no clear acoustic test for determining whether a vowel system is an ATR system. However, Standard Yorùbá mid vowels (from Disner 1983, and this study) are consistent with an ATR system in that the [−ATR] vowels [ẹ, ọ] are more back (lower F2) and lower (higher F1) than their [+ATR] counterparts [e, o].

2.4 Yorùbá phonemic inventory

In this section, I provide an overview of the Yorùbá phonemic inventory. This includes a discussion of vowels, which is relevant to the study, as well as a perfunctory overview of nasal vowels, syllabic nasals, consonants, and tone.

2.4.1 Vowel inventory

In all three dialects of this study there are seven underlying oral vowels, shown in Table 2.1. Vowels in Yorùbá are either [+ATR] or [−ATR].

Table 2.1. Underlying oral vowels for Àkùrẹ́, Mọ̀bà, and Standard Yorùbá

high	**i**	**u**	[+ATR]
mid	**e**	**o**	
	ẹ	**ọ**	[−ATR]
low	**a**		

For SY and Mọ̀bà, the surface vowel inventory is the same as the underlying inventory, this is shown in Table 2.2.

Table 2.2. Surface oral vowels for Mọ̀bà, and Standard Yorùbá

Height			ATR
high	**i**	**u**	[+ATR]
mid	**e**	**o**	[+ATR]
mid	**ẹ**	**ọ**	[−ATR]
low	**a**		[−ATR]

Àkùrẹ́ has two additional surface phones, as shown in Table 2.3. Both high vowels have [+ATR] and [−ATR] allophones: so the /i/ surfaces as either [+ATR] [i] or [−ATR] [ị] and /u/ surfaces as [+ATR] [u] and [−ATR] [ụ].

Table 2.3. Surface oral vowels for Àkùrẹ́

Height			ATR
high	**i**	**u**	[+ATR]
high	**ị**	**ụ**	[−ATR]
mid	**e**	**o**	[+ATR]
mid	**ẹ**	**ọ**	[−ATR]
low	**a**		[−ATR]

2.4.2 Nasal vowels and syllabic nasals

Standard Yorùbá has three nasal vowels /ĩ, ũ, ọ̃ / (Akinlabi, in preparation). These are written as ‘in’, ‘un’, and ‘ọn’ respectively, but the latter vowel may also be written as ‘an’, depending on the lexical item. Examples are found in (3).

(3) Standard Yorùbá nasal vowels

ĩ	pín	'divide, share'	ũ	fún	'give'
	òfin	'law'		ìkún	'knee'
			ọ̃	wọ́n	'be expensive'
				ìtàn	'story'

In addition, vowels are often nasalized after nasal consonants. Nasal vowels for Àkùrẹ́ and Mọ̀bà were not examined, although some discussion is found in Bamgbose (1967) and Capo (1989). Nasal vowels do not play a prominent role in this study, largely because they introduce an extra level of complexity.

Syllabic nasals are found in some lexical items and function words, as shown in (4). In Yorùbá orthography, the syllabic nasal is written as 'n', with an associated tone marking, unless it has a mid tone, which is unmarked. Any 'n' without a preceding or following vowel is a syllabic nasal.

(4) Standard Yorùbá syllabic nasals (Abraham 1958)

ńlá	'big'
nǹkan	'thing'
ń	progressive marker, as in *ó ń sùn*, 'He is sleeping.'
n	first person singular nominative pronoun in negative tenses, as in *n kò rí ọ*, 'I did not see you.'

2.4.3 Consonant inventory

The consonant inventory of the three Yorùbá dialects of this study is shown in Table 2.4. As the syllable structure is strictly (C)V, consonants may only occupy the onset position. The one exception is the syllabic nasal, which cannot have an onset or a coda. Consonant clusters are not allowed, and there are no geminate consonants. Within each table cell, the top row is voiceless, the bottom voiced.

Table 2.4. Consonant Inventory for Àkùrẹ́, Mọ̀bà, and Standard Yorùbá.

	Labial	Alveolar	Alveo-Palatal	Velar	Labio-velar	Glottal
Stops	 b	t d		k g	k͡p g͡b	
Affricates			 d͡ʒ			
Fricatives	f	s	ʃ			(h)[3]
Approximates		r, l	y		w	
Nasals	m	n				

2.4.4 Tone

SY has three phonemic tones, high, mid, and low, written as ‘ ́’, unmarked, and ‘ ̀’, respectively. Vowels and syllabic nasals carry tones. The surface tones may be level or contoured depending on the tonal combinations. Discussion of tone in Àkùrẹ́ and Mọ̀bà Yorùbá is not found in the literature, however, at least in general terms, it behaves the same as tone in Standard Yorùbá. Tone is not a focus of this study; however, tone was controlled in the experimental section of this study. For further discussion of Yorùbá tone, see Abraham (1958), Stahlke (1972),

[3] I found no examples of /h/ in Mọ̀bà Yorùbá, nor in Fresco’s (1970) word list from the related Èkìtì dialect spoken in Ifaki, Nigeria. For Àkùrẹ́, one example is found in Fresco (1970): *ẹ̀hà*, ‘ribs’. SY has several words containing /h/: for example, *ihò*, ‘hole’; *ehoro igbó*, ‘rabbit’; *ìhà*, ‘side’; *ahọ́n*, ‘tongue’; and *ehín*, ‘tooth’. In some SY words, [h] and [y] may alternate, as in *ẹ̀yìn/ẹ̀hìn*, ‘back’.

Akinkugbe (1978), Connell (1990), Akinlabi (1992), Laniran (1993), Akinlabi and Liberman (1999), and Akinlabi (in preparation).

2.5 Àkùrẹ́ vowel harmony

We now look at the vowel systems of the three dialects, starting with Àkùrẹ́, which has the most extensive harmony system. The observations I make here for Àkùrẹ́ are based on studies of Central Yorùbá dialects by Bamgbose (1967), Adetugbo (1967), Fresco (1970), and Omisore (1989). Àkùrẹ́ examples are from Fresco (1970) unless noted. Examples were checked with an Àkùrẹ́ consultant, from whom additional examples were also obtained and noted. ATR vowel harmony in Àkùrẹ́ is manifested in two ways: lexical co-occurrence restrictions and the productive alternation of some clitics. Although these two patterns can be treated as resulting from the same phenomenon occurring in different domains, for expository purposes I present them separately. I look first at co-occurrence within lexical words.

2.5.1 Àkùrẹ́ vowel harmony within words

For words containing only mid vowels, as in (5), the vowels must have the same ATR value; that is, [+ATR] mid vowels cannot occur in the same word as [−ATR] mid vowels.

(5) Àkùrẹ́ words with only mid vowels

ètè	‘lip’	ẹ̀jẹ̀	‘blood’	*eCẹ	*ẹCe
oko	‘farm’	ọkọ	‘husband’	*oCọ	*ọCo
ekpo	‘oil’	ẹ̀dọ̀	‘liver’	*eCọ	*ẹCo
òkè	‘mountain’	ọbẹ̀	‘soup’	*oCẹ	*ọCe

In words that are made up of only high vowels, only [+ATR] vowels are found, as shown in (6). In these cases, high vowels have the quality of [+ATR] high vowels, both impressionistically and acoustically, as we see in Chapter 4.

(6) Àkùrẹ́ words with only high vowels

[igi], 'tree'	*ịCị	*iCị	*ịCi
[ukù], 'belly'	*ụCụ	*uCụ	*ụCu
?iCu[4]	*ịCụ	*iCụ	*ịCu
[ùji], 'shade'	*ụCị	*uCị	*ụCi

The lack of [−ATR] vowels in words with only high vowels suggests that [−ATR] high vowels are not phonemic in Àkùrẹ́. This is corroborated by the absence of monosyllabic words with [−ATR] high vowels, while we find monosyllabic words containing all other vowels, as seen in (7).

(7) Monosyllabic Àkùrẹ́ words

kí	'greet'	jù	'throw'
*Cị		*Cụ	
dé	'arrive'	jó	'dance'
jẹ	'eat'	lọ	'go'
á	'come'		

In disyllabic words made up of high vowels followed by mid vowels, we see the same pattern as for mid vowels alone, that is, [+ATR] and [−ATR] vowels do not co-occur, as seen in (8).

[4] The '?' indicates that no words with the stated form were found in the literature. The apparent gap may or may not be accidental, but it is not investigated here. In the case of iCu, closely related Central Èdè dialects have words of this form as in Ijẹṣa, *ibú*, 'thick forest, stream' (Ijesa consultant); and Ìfàkì and Irun, *iú*, 'grey hair' (Akinkugbe 1978).

(8) Disyllabic Àkùrẹ́ words with high vowels followed by mid vowels

ìwé	‘book’	?ịCẹ	(no examples found)	*iCẹ	*ịCe
ìo	‘horn’	ịjọ́	‘day’	*iCọ	*ịCo
ulé	‘house’	ụsẹ́	‘work’	*uCẹ	*ụCe
uto	‘type of cane’	ụ̀jọ	‘congregation’	*uCọ	*ụCo

In this case, the ATR value of the high vowel is perceivable impressionistically, although the difference between [+ATR] and [−ATR] variants may be subtle even to native speakers. Acoustic measurements, discussed in chapter 4, confirm the differences in the ±ATR high vowel phones. Finally, the vowels differ in their phonological behavior, acting as triggers for harmony to their left, as we see shortly.

In disyllabic words with mid vowels followed by high vowels, as in (9), we again see that [+ATR] and [−ATR] vowels do not co-occur.

(9) Àkùrẹ́ words with [+ATR] mid vowels followed by high vowels

ebí	‘relations’	*eCị
etu	‘Maxwell’s duiker (type of antelope)’	*eCụ
obì	‘kola nut’	*oCị
ojú	‘eye, face’	*oCụ

More specifically only [+ATR] sequences are found—no disyllabic words start with [−ATR] mid vowels followed by high vowels, as schematized in (10). The absence of Àkùrẹ́ words with [−ATR] mid vowels followed by [+ATR] high vowels is taken up in detail in §2.8.

(10) Àkùrẹ́ words with [−ATR] mid vowels followed by high vowels

*ẹCụ *ẹCu *ẹCị *ẹCi

*ẹCụ *ẹCu *ẹCị *ẹCi

Note the absence of words ending with a [−ATR] high vowel. The pattern indicates that the direction of the harmony is anticipatory, that is, right to left, since it has already been established that [−ATR] high vowels are not phonemes in Àkùrẹ́ and only result from harmony. As a result of this, [−ATR] high vowels only surface to the left of [−ATR] vowels.

The vowel /a/ behaves asymmetrically with respect to harmony. While /a/ may precede vowels from either set (11), /a/ is preceded only by [−ATR] vowels (12).

(11) /a/-initial Àkùrẹ́ words (from Fresco 1970, except words in []s are from an Àkùrẹ́ consultant)

aCi:	(no examples found)[5]		
aCu:	arúgbó, 'old man'		
aCe:	àlejò, 'guest'	[ate, 'hat']	[ajé, 'economy']
aCẹ:	alẹ̀, 'ground'	alẹ́, 'night'	[abẹ, 'razor']
aCo:	àgbò, 'ram'	àbó, 'junior sibling'	àgbàdo, 'corn'
aCọ:	aṣọ, 'clothing'	aọ, 'body'	
aCa:	abà, 'village'	àlá, 'dream'	àdàbà, 'dove'

[5] Adetugbo (1967) records *aṭị* 'and' with a [−ATR] high vowel. The existence of one function word with an unexpected pattern is not taken as evidence against the robust patterns presented here.

(12) Àkùrẹ́ words with /a/ in non-initial position (from Fresco 1970, except that words in []s are from an Àkùrẹ́ consultant)

ịCa:	ịlá, 'okra'	ịbà, 'fever'	[ịta, 'pepper']
ụCa:	[ụ̀jà, 'fight']	[ụgbá, 'calabash']	
ẹCa:	ẹja, 'fish'	ẹkà, 'corn'	ẹ̀kpà, 'ground nut'
ọCa:	ọba, 'chief'	ọjà, 'market'	

*iCa, *uCa, *eCa, *oCa

The patterning of words with /a/ confirms that harmony within the word spreads from right to left (Bakovic 2000 makes a contrary claim). Because of the existence of words like *ate*/*àgbò*, *abẹ*/*aṣọ*, and *ẹja*/*ọjà*, but not words with the shape *eCa and *oCa, we know that the [−ATR] value of /a/ spreads leftward and not rightward. If the [−ATR] value spread rightward, we would be correct in expecting words like *abẹ*/*aṣọ*, but we would not expect to find words like *ate*/*àgbò*. Likewise, if the [−ATR] value did not spread leftward from /a/, we would expect such forms as *eCa and *oCa to exist. The behavior of /a/ shown here for Àkùrẹ́ is typical of many other ATR vowel harmony systems, such as Akan (Kwa, Niger-Congo) (Clements 1981), Igbo (Benue-Congo, Niger-Congo) (Green and Igwe 1964), and Lulubo (Central Sudanic, Nilo-Saharan) (Andersen 1987). For example in Lulubo, words with /a/ before [−ATR] vowels and [+ATR] vowels are common, as are words where /a/ follows [−ATR] vowels; words where /a/ appears after a [+ATR] vowel are few, and are probable loan words. This is shown in (13).

(13) /a/ distribution in Lulubo (Andersen 1987)

a-C-[−ATR]		a-C-[+ATR]	
àsị	'fire'	àlí	'short'
àŋgwɛ	'white'	rabolo	'banana'
àlụ	'one'	ámbúrú	'mosquito'

[−ATR]-C-a		[+ATR]-C-a[6]
ị̀ba	'rope'	(not found)
ẹ́dʒá	'pot'	
ụkwa	'wealth'	

The distribution of vowels in disyllabic words in Àkùrẹ́ is summarized in Table 2.5.

[6] Andersen gives some examples of loan words with this pattern, e.g. *ìmòkwà*, 'to iron' (from Arabic).

Table 2.5. Distribution of V1 and V2 vowels in Àkùrẹ́ (C)V$_1$CV$_2$ words. A "✓" indicates existing vowel co-occurrences. A "?" indicates a possible accidental gap or near gap. The cells with diagonal lines indicate vowel combinations with [+ATR] and [−ATR] vowels. The shaded cells indicate words with only [−ATR]; the unshaded cells indicate [+ATR] words.

V1 \ V2		/i/		/e/	/ẹ/	/a/	/ọ/	/o/	/u/	
		[i]	[ị]						[ụ]	[u]
/i/	[i]	✓		✓				✓		?
	[ị]				?	✓	✓			
/e/		✓		✓				✓		✓
/ẹ/					✓	✓	✓			
/a/		?		✓	✓	✓	✓	✓		?
/ọ/					✓	✓	✓			
/o/		✓		✓				✓		✓
/u/	[ụ]				✓	✓	✓			
	[u]	✓		✓				✓		✓

To summarize, within a word, adjacent vowels must belong to the same set, except that while /a/ must be preceded only by [−ATR] vowels, /a/ may precede vowels from either set, as summarized in (14).

(14) aCe, aCẹ, ẹCa, *eCa

aCo, aCọ, ọCa, *oCa

The high vowel phonemes /i/ and /u/ precede vowels of either set, but may only follow [+ATR] vowels (with the possible exception of /a/) as summarized in (15). When followed by a [−ATR] vowel, a high vowel is realized with its [−ATR] allophone.

(15) iCe, ịCẹ, eCi, *ẹCi

iCo, ịCọ, oCi, *ọCi

2.5.2 Àkùrẹ́ vowel harmony across word boundaries

We have seen Àkùrẹ́'s distribution of ATR vowels within the word, now we examine harmony extending to preposed clitics.[7] Singular subject pronouns harmonize with vowels in the following word, as shown for the third person singular, *ó/ọ́*, in (16).

(16)	ó dé	's/he arrived'	ọ́ jẹ	's/he ate'
	ó jó	's/he danced'	ọ́ lọ	's/he went'
	ó kí	's/he greeted'	ọ́ á	's/he came'
	ó kú	's/he died'		

Before [+ATR] vowels (i, u, e, o), the pronoun is realized as [+ATR] *ó*; before [−ATR] vowels (ẹ, ọ, a), the pronoun is realized as [−ATR] *ọ́*. The harmony involves both mid and high vowels. In (17), the mid vowel of the noun is the trigger, with the mid vowel of the pronoun as the target. The apostrophe signifies that a vowel was elided; in this case, the /í/ of *rí*, 'to see', is elided when preceding a vowel. In some cases, the second of the two vowels is elided. In cases where the vowels are the same, it is not possible to determine which vowel is deleted.

(17)	ó r'oko	's/he saw the farm.'	/rí/, 'see'
	ọ́ r'ọkọ̀	's/he saw the car.'	

High vowels act similarly. They undergo harmony by surfacing as distinct allophones, as is the case with all the high vowels in (18). They also trigger harmony, so all the mid and high vowels to the left of a high vowel undergo harmony.

[7] Suffixes are not common in Yorùbá, and are not examined.

(18) High vowels participate in harmony in Àkùrẹ́

(a) ó r'ulé 's/he saw the house.' /rí/, 'see'

(b) ọ́ r'ụgbá 's/he saw the calabash.'

(c) ó di ti jó 'it has burnt again.'

(d) ọ́ dị tị bẹ́ 'it has burst again.'

What we see above in (18) is cross-height harmony, which typifies ATR harmony systems. In this case, a high vowel [ụ] lowers a mid vowel /o/ to a phonetically lower mid vowel [ọ]; that is, a high vowel acts as a trigger to lower a vowel that is already lower than itself. The sentences in (18c, d) above are modeled on Bamgboṣe (1967: 270) whose examples are for Ijesa and Èkìtì,

(19) High vowels participate in harmony in closely related Ìjẹ̀ṣà and Èkìtì, Bamgboṣe (1967)

ó tú ti jó 'it has burnt again.'

ọ́ tụ́ tị bẹ́ 'it has burst again.'

The examples shown so far have mostly used the third person singular pronoun. In addition, (18) shows vowel harmony in the particles *di* 'again' and *ti* '(past marker)'. Harmony is found in other pronouns as well, as shown in (20) where the pronouns are set off in boxes.

(20)	[+ATR] object		[−ATR] object	(from consultants)	
	mo rígi	‘I saw a tree.’	mọ rị́lá	‘I saw okra.’	
	wo rígi	‘you saw a tree.’	wọ rị́lá	‘you saw okra.’	
	ó rígi	‘s/he saw a tree.’	ọ́ rị́lá	‘s/he saw okra.’	
	a rígi	‘we saw a tree.’	a rị́lá	‘we saw okra.’	
	ẽ rígi	‘you (pl) saw a tree.’	ẽ rị́lá	‘you (pl) saw okra.’	
	ã rígi	‘they saw a tree.’	ã rị́lá	‘they saw okra.’	

Pronouns with [+ATR] oral vowels alternate (*mo*, *wo*, *ó*), while those with [−ATR] or nasal vowels do not. Note that the vowels found in the non-alternating pronouns do not alternate in other contexts. In addition to these pronouns and the other prefixes mentioned in examples, there may be other clitics participating in vowel harmony in Àkùrẹ́ that are not listed here.

Vowel harmony does not occur when the subject is a full noun, as in (21).

(21)	òbí r’ọkọ̀	‘The parents saw the car.’	/rí/, ‘see’
	*ọ̀bị́ r’ọkọ̀		
	ọdẹ r’ulé	‘The hunter saw the house.’	
	*ode r’ulé		

Nor does harmony occur between an object and its preceding verb, as in (22).

(22)	ó ti rí ẹ̀bà	‘s/he has seen eba.’	*ọ́ tị rị́ ẹ̀bà
	ó ru obì	‘s/he carried kola.’	
	ó ru ịlá	‘s/he carried okra.’	*ọ́ rụ ịlá
	ó ru’lá	‘s/he carried okra.’	*ọ́ rụ’lá

Bamgboṣe (1967) provides some nice minimal pairs showing that when the vowel of the verb is deleted, as in (23a), harmony proceeds from the object to the subject pronoun.

(23) Interesting vowel elision in Ìjẹ̀ṣà, Èkìtì (Bamgboṣe 1967) and Àkùrẹ́

(a)	ọ́ r'ụgbá	'he bought a calabash'	/ra/, /ugbá/
	ọ́ j'ịlá	'he ate okra'	/jẹ/, /ilá/
(b)	ó ru'gbá	'he carried a calabash'	/ru/, /ugbá/
	ó jí'lá	'he stole okra'	/jí/, /ilá/

This is also the case when the verb's elided vowel is [+ATR], as in (17) repeated as (24). However, when the initial vowel of the object is deleted, as in (23b), the object's remaining vowel does not trigger harmony in the verb. Bamgboṣe's examples are for Ìjẹ̀ṣà and Èkìtì, which are closely related to Àkùrẹ́; I verified that they were identical for Àkùrẹ́.

(24)

ó r'oko	's/he saw the farm.'	/rí/, 'see'
ọ́ r'ọkọ̀	's/he saw the car.'	

ATR harmony in Àkùrẹ́ is also exhibited in word formation, as shown in (25). Here *olí-*, meaning 'possessor of' (Abraham 1958, for SY cognate *oní-*), is prefixed to a noun to form a noun meaning 'seller of' or 'possessor of'. When the noun is consonant initial, the prefix agrees in ATR with the initial vowel of noun. When the following noun starts with a vowel, the /i/ of the prefix elides. The initial vowel of the prefix then assimilates in ATR value with a following high vowel, otherwise it undergoes total assimilation with the following vowel.

(25) Word formation in Àkùrẹ́

ọlị́bàtà	'shoe seller'	/bàtà/
olígi	'wood seller'	/igi/
ọlị́lá	'okra seller'	/ilá/
ọlị́gbǎ	'garden egg seller'	/ìgbá/
olúṣu	'yam seller'	/uṣu/
elépo	'oil seller'	/epo/
olóbì	'kola seller'	/obì/
ẹlẹ́ja	'fishmonger'	/ẹja/
ẹlẹ́mu	'palm wine seller'	/ẹmu/
ọlọ́tín	'wine seller' or 'drunkard'	/ọtín/
aláṣọ	'clothes seller'	/aṣọ/

2.5.3 Domain of Àkùrẹ́ ATR harmony

In determining the domain of ATR harmony in Àkùrẹ́, I assume the prosodic framework as described in Nespor and Vogel (1986), in which the prosodic hierarchy consists of the following levels:

(26) utterance
intonational phrase
phonological phrase
phonological word

It is clear that the domain of ATR harmony does not extent to the utterance or intonation phrase level, since we do not see harmony extend across even simple phrases such as those in (27).

(27)	ó mù'wé sálẹ̀	'He put the book on the ground.'	/ìwé/, 'book'
	òbí r'ọ́kọ̀	'The parent saw the car.'	/rí/, 'see'

No criteria were found for distinguishing between the remaining candidates, the phonological word and phonological phrase in Àkùrẹ́. I follow Zsiga (1992) who found ATR vowel harmony to occur within the phonological word for Igbo, whose behavior with respect to ATR harmony domain is similar to that of Àkùrẹ́ Yorùbá. In the Igbo sentence in (28), the phonological words are bracketed based on ATR specification. The data were obtained by me from an Igbo consultant.

(28) Igbo ATR harmony example:

[ó	nyèrè]$_{+ATR}$	[Àmà]$_{-ATR}$	[éwú]$_{+ATR}$.
3sg	gave	Ama	goat.

'She gave a goat to Ama.'

I thus assume that ATR harmony occurs within the phonological word for Àkùrẹ́. For the purposes of this study, characterizing the precise domain in prosodic hierarchy terms is not essential. See chapter 5 for more discussion on the domain.

2.5.4 Summary of Àkùrẹ́ ATR harmony

The vowel harmony of Àkùrẹ́ Yorùbá can be summarized by the rule in (29) which states that a non-low vowel becomes [−ATR] if before a [−ATR] vowel with an optional intervening consonant, when the vowels are within the phonological word.

(29) $[-\text{low}] \rightarrow [-\text{ATR}] \,/\, \mathbf{[}_{\text{Pwd}} \ldots \_\_\_\_ \; C_0 \; [-\text{ATR}] \ldots \mathbf{]}$

The use of a rule here is not intended to suggest a rule-based approach over any other, it is instead a manner of formulating the facts into a succinct statement. Alternatively this generalization could be stated using a constraint-based theory such as Optimality

Theory. For a constraint based analyses of Yorùbá vowel harmony, see Akinlabi (in preparation) and Bakovic (2000).

2.6 Mọ̀bà vowel harmony

The Mọ̀bà[8] data presented here were collected from a consultant from Ọ̀tùn-Èkìtì, Nigeria, unless otherwise noted. Additional assistance was generously provided by Oladiipo Ajiboye. The most important difference between Mọ̀bà and Àkùrẹ́ Yorùbá vowel harmony is easily stated: Mọ̀bà exhibits harmony in the same contexts as Àkùrẹ́; however, Mọ̀bà harmony does not target high vowels, as shown in (30). While the mid vowel pronoun /e/ surfaces as [−ATR] [ẹ] when directly in front of a [−ATR] mid vowel [ẹ] (30b), the high vowels in (30d) do not undergo harmony from the [ẹ], and block harmony to the pronoun. Compare this with the Àkùrẹ́ patterns in (31), repeated from (18).

(30) High vowels do not alternate in Mọ̀bà

(a)	é jó	'it burnt.'
(b)	ẹ́ bẹ́	'it burst.'
(c)	é ti tí jó	'it has burnt again.'
(d)	é ti tí bẹ́	'it has burst again.'

(31) High vowels participate in harmony in Àkùrẹ́ (repeated from (18)).

ó r'ulé	's/he saw the house.'	/rí/, 'see'
ọ́ r'ụgbá	's/he saw the calabash.'	
ó di ti jó	'it has burnt again.'	
ọ́ dị tị bẹ́	'it has burst again.'	

[8] In addition to Bamiṣilẹ (1987) which I was not able to find, only one published paper on Mọ̀bà was found: Bamiṣilẹ's (1994) short paper on vowel coalescence.

In the rest of this section, we examine in detail the harmony of Mọ̀bà Yorùbá.

2.6.1 Mọ̀bạ̀ vowel harmony within the word

Monosyllabic words in Mọ̀bà may contain any of the seven oral vowels (32).

(32) Monosyllabic Mọ̀bà words (verbs)

rí	'see'	kú	'die'
dé	'arrive'	jó	'dance'
jẹ	'eat'	lọ	'go'
ká	'fold'		

As in Àkùrẹ́, Mọ̀bà words containing only mid vowels have either [+ATR] or [−ATR] vowels, but not both (33).

(33) Mọ̀bà words with only mid vowels

ègbè	'chorus'	ẹ̀gẹ́	'cassava'	*eCẹ	*ẹCe
oko	'farm'	ọkọ	'husband'	*oCọ	*ọCo
ekpo	'oil'	ẹ̀kọ́	'lesson'	*eCọ	*ẹCo
òkè	'mountain'	ọbẹ̀	'soup'	*oCẹ	*ọCe

Like Àkùrẹ́, the vowel co-occurrence patterns exhibited by the low vowel in Mọ̀bà indicate the direction of harmony. In (34), we see that /a/ may precede both [+ATR] and [−ATR] vowels.

(34) /a/-initial Mọ̀bà words

aCi:	àlí, 'thief'	Akí, 'name'	
aCu:	àrù, 'fear' (Ajiboye 1991)		Adú, 'name'
aCe:	agbe, 'begging'	àlè, 'concubine'	àwé, 'friend'
aCẹ:	àgbẹ̀, 'farmer'	akpẹ́, 'applause'	akpẹ, 'type of pot'
aCo:	àgbò, 'ram'	àtò, 'wood for bottom of a pile of firewood'	
aCọ:	aṣọ, 'clothing'	awọ, 'skin'	àkọ̀, 'sheath for sword'
aCa:	ajá, 'dog'	àka, 'type of ant'	aká, 'arm'

However, with /a/ as the final vowel, we find a restriction such that [+ATR] mid vowels may not precede. All other vowels may precede a final /a/, including the low vowel /a/, the [−ATR] mid vowels, and the high vowels, shown in (35).

(35) Mọ̀bà words with /a/ in non-initial position

iCa:	ilá, 'okra'	ikpá, 'epilepsy'	ìka, 'finger'
uCa:	ugbá, 'calabash'	ùgà, 'king's courtyard'	
	ùkà, 'something for protecting small trees'		
*eCa			
ẹCa:	ẹja, 'fish'	ẹkà, 'cassava flour meal, sorghum'	
	ẹ̀kpà, 'ground nut'		
*oCa			
ọCa:	ọba, 'chief'	ọ̀pá, 'cane'	ọgbá, 'type of snake'
aCa:	ajá, 'dog'	àka, 'type of ant'	aká, 'arm'

As in Àkùrẹ́, the absence of *eCa and *oCa suggests that vowel harmony moves from right to left. While the mid vowel behavior is the same for both dialects ([+ATR] mid vowels can only follow /a/, not precede it), this is not the case with the high vowels.

Mọ̀bà allows a [+ATR] high vowel before /a/, while Àkùrẹ́ does not. There are two points to make regarding this fact: first, there are no [−ATR] high vowels in Mọ̀bà; and second, in Mọ̀bà we find [+ATR] (high) vowels preceding [−ATR] vowels, something that does not occur in Àkùrẹ́.

Words containing only high vowels necessarily have only one ATR value, since there is only one phone for each of the high vowel phonemes, namely, the [+ATR] vowels, [i] and [u]; examples are shown in (36). These high vowels are realized as [+ATR] vowels, both impressionistically and when measured instrumentally, which we see in chapter 4.

(36) Mọ̀bà words with only high vowels

igi	'tree'				
ukù	'belly'	ùlù	'drum'	uṣu	'yam'
ìdu	'fibroid cyst'	ìfù	'type of sickness'	ìwù	'fur'
ùkì	'bullet'				

Words containing high vowels followed by mid vowels may contain either [+ATR] or [−ATR] mid vowels, as shown in (37).

(37) Mọ̀bà words with high vowels followed by mid vowels

ìwé	'book'	ìpẹ́	'fish scales'
ìgò	'bottle'	igọ̀	'cavity in roots of big tree'
ulé	'house'	ùgbẹ́	'the bush'
ùgbó	'type of vegetable'	ùkọ́	'hook (n.)'

Again we find in Mọ̀bà, that [+ATR] high vowels may precede [−ATR] vowels, in this case mid vowels. Similarly, in words with a mid vowel followed by a high vowel, both [+ATR] and [−ATR] mid vowels are found, as in (38).

(38) Mọ̀bà̩ words with mid vowels followed by high vowels

ebi	'hunger'	ẹbí	'relations'
eku	'rat'	ẹtu	'type of deer'
obì	'kola nut'	?ọCi	
ojú	'eye, face'	?ọCu	

Here, note that [−ATR] mid vowels may precede [+ATR] (high) vowels, something not found in Àkùrẹ́. I return to this issue in §2.8.

Table 2.6 summarizes the distribution of vowels in disyllabic words in Mọ̀bà.

Table 2.6. Distribution of V1 and V2 vowels in Mọ̀bà (C)V_1CV_2 words. A "✓" indicates existing vowel co-occurrences. A "?" indicates a possible gap, or near gap. The cells with diagonal lines indicate vowel combinations with [+ATR] and [−ATR] vowels. The shaded cells indicate words with only [−ATR]; the unshaded cells indicate [+ATR] words.

V2 / V1	i	e	ẹ	a	ọ	o	u
i	✓	✓	✓	✓	✓	✓	✓
e	✓	✓				✓	✓
ẹ	✓		✓	✓	✓		✓
a	✓	✓	✓	✓	✓	✓	✓
ọ	?		✓	✓	✓		?
o	✓	✓				✓	✓
u	✓	✓	✓	✓	✓	✓	✓

2.6.2 Mọ̀bà̩ vowel harmony across word boundaries

In this section, we look at productive harmony across morphological boundaries. In Mọ̀bà, preposed clitics agree in ATR value with the following vowels. In (39a) and (39b), vowels in the third person singular subject prefix /é/ agree with vowels in the following verb. Note that the third person singular subject prefix in Mọ̀bà (é/ẹ́) differs from that of Àkùrẹ́ (ó/ọ́) and SY (ó).

(39) Right-to-left harmony to preposed clitics

(a)	é dé	's/he arrived'	é jó	's/he danced'
	é rí	's/he saw'	é kú	's/he died'
(b)	ẹ́ jẹ	's/he eats'	ẹ́ lọ	's/he went'
	ẹ́ á	's/he came'		

As shown in (40), harmony involves mid vowels both as triggers and targets: the mid vowels in the nouns trigger harmony in the mid vowels of the subject prefixes.

(40) Harmony involves mid vowels

é r'oko	's/he saw the farm.'
ẹ́ r'ọkọ̀	's/he saw the car.'

But when the subject prefix has a [−ATR] mid vowel, alternation does not occur, as shown in (41).

(41)

ọ̀ọ́ fẹ́ …	'you (sg.) want…'
ọ̀ọ́ jó	'you (sg.) danced.'

The high vowels are opaque: they do not participate in harmony as triggers, nor do they participate as targets. Mid vowels, which alternate when before mid-vowels, are always realized as [+ATR] vowels when before high vowels, as seen in the pronouns in (42). We see that the underlying form of the prefix here contains the [+ATR] mid vowel /é/, in contrast with the [−ATR] vowel above in (41) which does not alternate.

(42)	High vowels do not trigger alternation in Mọ̀bà	
	é r' ulé	's/he saw the house.'
	é r' ugbá	's/he saw the calabash.'

High vowels do not alternate before ±ATR vowels, as shown in (43). Before the [+ATR] vowel [o] in *jó* and before the [−ATR] vowel [ẹ] in *bẹ́*, the vowel of *ti* surfaces as [+ATR] [i].

(43)	High vowels do not alternate in Mọ̀bà	
	é ti tí jó	'it has burnt again'
	é ti tí bẹ́	'it has burst again'

In (44), we see two variants of the future tense. In (a), the harmony is blocked by the high vowel in *ní*, so the negation particle surfaces with its [+ATR] underlying form, /ké/; while in (b), the harmony targets the mid vowels to the left and hence we see alternation.

(44)	(a)	a kè ní lọ	'we will not go.'	
		a kè ní ṣe	'we will not do.'	
	(b)	a kẹ̀ ẹ̀ lọ	'we will not go'	(Oladiipo Ajiboye, pc)
		a kè è ṣe	'we will not do'	

The regular subject pronouns of Mọ̀bà are shown in (45). In Mọ̀bà, as in Àkùrẹ́, only pronouns with [+ATR] oral vowels alternate. While in Àkùrẹ́ we found alternation in three of the six regular subject pronoun clitics, in Mọ̀bà, only the first and third person singular pronouns alternate.

(45)

	Singular	Plural
1st	**mè/mẹ̀**	à
2nd	ọ̀	ḭ̀
3rd	**é/ẹ́**	à̰

Harmony is also found with negation, as in (46), but it does not extend through /ke/ to [−ATR] pronouns, as shown in (47).

(46)

kè rí	's/he didn't see'	kè kú	's/he didn't die'
kè dé	's/he didn't arrive'	kè jó	's/he didn't dance'
kẹ̀ jẹ	's/he didn't eat'	kẹ̀ lọ	's/he didn't go'
kẹ̀ gbà	's/he didn't collect'		

(47)

ọ kẹ̀ lọ 'you didn't go'

ọ kè jó 'you didn't dance'

a kẹ̀ lọ 'we didn't go'

a kè jó 'we didn't dance'

Harmony is also exhibited with the future in (48) and the negative future (49), but again does not extend to [−ATR] or nasal vowel pronouns.

(48) Future in Mọ̀bà; harmonizing prefixes are $\boxed{\text{boxed}}$

[−ATR] verbs		[+ATR] verbs	
$\boxed{\text{ẹ́ẹ̀}}$ lọ	's/he will go'	$\boxed{\text{éè}}$ jó	's/he will dance'
$\boxed{\text{mẹ̀ẹ̀}}$ lọ	'I will go'	$\boxed{\text{mèè}}$ jó	'I will dance'
ọ̀ $\boxed{\text{ẹ̀}}$ lọ	'you (sg.) will go'	ọ̀ $\boxed{\text{è}}$ jó	'you (sg.) will dance'
à $\boxed{\text{ẹ̀}}$ lọ	'we will go'	à $\boxed{\text{è}}$ jó	'we will dance'
àn $\boxed{\text{ẹ̀}}$ lọ	'they will go'	àn $\boxed{\text{è}}$ jó	'they will dance'
ìn $\boxed{\text{ẹ̀}}$ lọ	'you (pl.) will go'	ìn $\boxed{\text{è}}$ jó	'you (pl.) will dance'

(49) Negative Future in Mọ̀bà

[−ATR] verbs	[+ATR] verbs
ọ $\boxed{\text{kẹ̀ẹ̀}}$ lọ ‘you won’t go’	ọ $\boxed{\text{kèè}}$ jó ‘you won’t dance’
ọ kè ní lọ ‘you won’t go’	ọ kè ní jó ‘you won’t dance’

In Mọ̀bà, as in Àkùrẹ́, vowel harmony does not occur when the subject is a full noun, as in (50).

(50)	òbí r’ọkọ̀	‘The parents saw the car.’	*ọ̀bị́ r’ọkọ̀
	ọdẹ r’ulé	‘The hunter saw the house.’	*ode r’ulé

Vowel harmony also does not occur between the object and the verb, as shown in (51).

(51)	é $\boxed{\text{ro}}$ $\boxed{\text{ẹjọ́}}$	‘s/he stated his/her case.’	(Oladiipo Ajiboye, p.c.)
	é $\boxed{\text{re}}$ $\boxed{\text{ọ̀tùn}}$	‘s/he went to Ọ̀tùn.’	

As we have seen, usually when a verb and a vowel-initial object are adjacent, either the vowel of the verb or the initial vowel of the object is elided. When it is the vowel of the verb that is deleted, harmony proceeds from the object to the subject pronoun, as shown in (52). The [+ATR] vowel of the verb is deleted, leaving a [−ATR] vowel of the object to harmonize with the subject.

(52)	ẹ́ r’ẹ́ja.	‘s/he saw the fish.’	/rí/, /ẹja/
	ẹ́ r’ọ́kọ́.	‘s/he saw the hoe.’	/rí/, /ọkọ́/
	é r’óko.	‘s/he saw the farm.’	/rí/, /oko/

Similarly, in (53), the [−ATR] vowel of the verb is deleted, leaving a [+ATR] vowel in the object to harmonize with the subject. Note that the deleted [−ATR] vowel does not trigger harmony in the subject.

(53)	é j'obì	's/he ate kola.'	/jẹ/, /obì/
	é j'ilá	's/he ate okra'	/jẹ/, /ilá/
	é gb'ekpo	's/he collected the oil'	/gbà/, /ekpo/

Finally, in (54), the vowel of the verb remains, while the vowel of the object is elided. In this case, the pronoun harmonizes with the remaining vowel of the verb, and there is no harmony between the object and the verb.

(54)	ẹ́ kọ'jò	's/he refused the snake.'	/kọ/, /ejò/
	ẹ́ kọ'bì	's/he refused the kola.'	/kọ/, /obì/
	é se'ja	's/he prepared the fish.'	/sè/[9], /ẹja/
	ẹ́ kọ'ja	's/he refused the fish.'	/kọ/, /ẹja/

In word formation in Mọ̀bà, we see total vowel assimilation as opposed to simple ATR vowel harmony, as shown in (55), except that mid vowel of the first vowel does not become high in front of a high vowel.

[9] All low toned verbs, such as *sè*, 'prepare', surface with a mid tone in certain environments.

(55) Word formation in Mọ̀bà

ení + item = item seller

eníbàtà	'shoe seller'	bàtà, 'shoe'
enígi	'wood seller';	igi, 'wood'
enílá	'okra seller'	ilá, 'okra'
eníkàn	'garden egg seller'	ikàn, 'garden egg'
elékpo	'oil seller'	ekpo, 'oil'
ẹlẹ́ja	'fishmonger'	ẹja, 'fish'
ẹlẹgẹ̀ẹ́	'cassava seller'	ẹ̀gẹ́, 'cassava'
aláṣọ	'clothes seller'	aṣọ, 'clothes'
ọlọ́tín	'wine seller'	ọtín, 'wine'
olóbì	'kola seller'	obì, 'kola'
olúṣu	'yam seller'	uṣu, 'yam'
olúgbá	'calabash seller'	ugbá, 'calabash'

2.6.3 Vowel harmony rule in Mọ̀bà

In summary, we find that in Mọ̀bà, vowel harmony operates in much the same way as in Àkùrẹ́, except that the high vowels are opaque to harmony—they do not undergo alternation, nor do they trigger harmony to their left. In Mọ̀bà, the patterns found within the word are active across word boundaries within the same domain as found for Àkùrẹ́, which I assume is the phonological word. Thus we can say that the harmony in Mọ̀bà is also active harmony, as opposed to one limited to distribution across the lexicon. The vowel harmony rule in Mọ̀bà can be stated as in (56), and is similar to the rule for Àkùrẹ́ in (29) above, except that in Mọ̀bà, only mid vowels are targeted for harmony, while for Àkùrẹ́, mid and high vowels are targeted.

(56) $[-\text{low}, -\text{high}] \rightarrow [-\text{ATR}] / \mathbf{[}_{\text{Pwd}} \ldots \_\_\_\_ \text{C}_0 [-\text{ATR}] \ldots\mathbf{]}$

Again, the use of such a rule here is intended as a way of concisely capturing the harmony, rather than an indication of a rule-based analysis.

2.7 Standard Yorùbá vowel harmony

Standard Yorùbá exhibits the least degree of vowel harmony of the three dialects in this study. Unlike Àkùrẹ́ and Mọ̀bà, SY shows no harmony to clitics, shown in (57).

(57) Clitics do not alternate in SY

(a)	ó jó	'it burnt.'
(b)	ó bẹ́	'it burst.'
(c)	ó tún ti jó	'it has burnt again.'
(d)	ó tún ti bẹ́	'it has burst again.'

Interestingly, some other North West Èdè dialects do show alternations in the clitics, as in Ọ̀yọ́ and Ẹ̀gbádò, where the singular subject pronouns agree in ATR value with the vowel of the following verb, seen in (58).

(58) ATR harmony in Ọ̀yọ́, Ẹ̀gbádò, and Standard Yorùbá (Bamgbose 1967).

	Ọ̀yọ́, Ẹ̀gbádò	SY	
a)	mọ lọ	mo lọ	'I went'
b)	mo jó	mo jó	'I danced'
c)	ọ fẹ́	o fẹ́	'you (sg.) want'
d)	o dé	o dé	'you (sg.) arrived'
e)	ọ́ wá	ó wá	'he came'
f)	ó kú	ó kú	'he died'

Note that for Ọ̀yọ́ and Ẹ̀gbádò, as in Mọ̀bà, [+ATR] does not spread to [−ATR] vowels, as can be seen in (59) where the [−ATR] pronoun /ẹ/ does not change regardless of the following vowel.

(59) [+ATR] harmony does not spread to [−ATR] pronouns (Bamgbose 1967).

	Ọ̀yọ́, Ẹ̀gbádò	SY	
a)	ẹ lọ	ẹ lọ	'you (pl.) went'
b)	ẹ jó	ẹ jó	'you (pl.) danced'
c)	ẹ fẹ́	ẹ fẹ́	'you (pl.) want'
d)	ẹ dé	ẹ dé	'you (pl.) arrived'

The only indication of harmony in SY is that vowels found in lexical items have a restricted distribution, similar to the patterns we have seen already for Àkùrẹ́ and especially, Mọ̀bà. The facts of Yorùbá vowel harmony and analyses thereof have been presented by Adetugbo (1967), Bamgboṣe (1967), Awobuluyi (1967, 1985), and Oyelaran (1973), Archangeli and Pulleyblank (1989, 1993), Akinlabi (in preparation), Bakovic (2000), and Ọla Orie (2001). Examples in this section are drawn from Abraham's (1958) dictionary, unless otherwise indicated. I first present the facts of the vowel patterns in SY.

2.7.1 SY vowel distribution

Monosyllabic words in SY may contain any of the seven oral vowels (60).

(60) Monosyllabic SY words (verbs)

rí	'see'	kú	'die'
dé	'arrive'	jó	'dance'
jẹ	'eat'	lọ	'go'
ká	'fold'		

In SY, words may not have an initial /u/. Historically, initial /u/s in Proto-Yorùbá became /i/s in Standard Yorùbá and related dialects, for example, *ilé* (SY), 'house', is *ulé* in Mọ̀bà and Àkùrẹ́, *ulí* in Ondo (Fresco 1970), from **ulí,* Proto-Yorùbá (Akinkugbe 1978). While the high vowels are phonetically [+ATR] (impressionistically and from acoustic measurements, see Chapter 4), they are neutral with respect to [ATR]. That is, they occur with both [+ATR] and [−ATR] vowels in either direction. The words in (61) show that /i/ can precede all other vowels.

(61) SY words with /i/ followed by [+ATR] and [−ATR] vowels

i-C-[+ATR]	i-C-[−ATR]
idì, 'eagle'	
ìlú, 'town'	
ilé, 'house'	ilẹ̀, 'ground'
igbó, 'bush'	ìwọ́, 'hook'
	ìka, 'finger'

And in (62), we see that /i/ and /u/ may follow any other vowel.

(62) SY words with [+ATR] and [−ATR] vowels followed by high vowels.

[+ATR]-C-[+high]		[−ATR]-C-[+high]	
igi	'tree'		
ikú	'death'		
ebi	'hunger'	ẹbí	'family, relations'
eku	'mouse'	ẹtu	'type of deer'
obì	'kola nut'	ọtí	'alcoholic drink'
ojú	'eye, face'	?ọCu	
		àdí[10]	'palm nut oil'
		atú	'type of yam'

For words with only mid vowels, SY words must contain either [+ATR] vowels, /e, o/, or [−ATR] vowels, /ẹ, ọ/, as shown in (63), although there are exceptions, which I return to shortly.

(63) SY VCV words with only mid vowels

ètè	'lip'	ẹ̀jẹ̀	'blood'	*eCẹ	*ẹCe
oko	'farm'	ọkọ	'husband'	*oCọ	*ọCo
ekpo	'oil'	ẹ̀dọ̀	'liver'	*eCọ	*ẹCo
òkè	'mountain'	ọbẹ̀	'soup'	*oCẹ	*ọCe

[10] Also *àdín*.

The low vowel /a/ can occur with vowels from either set when /a/ is initial, as shown in (64).

(64) SY VCV words with /a/ followed by mid vowels.

ago, ‘rat’	agọ̀, ‘stupid person’
àjẹ̀, ‘paddle’	abẹ, ‘razor’

However, when /a/ is final, it patterns with [−ATR] vowels, as only [−ATR] mid vowels may appear to its left, as shown in (65). The asymmetric distribution of /a/ suggests a right-to-left directionality in SY, with [−ATR] being the active feature.

(65) SY VCV words with mid vowels followed by /a/.

*eCa	ẹ̀gbà, ‘bracelet’
*oCa	ọ̀gà, ‘chameleon’

The distribution of vowels in VCV words is shown in Table 2.7 which indicates allowable vowel combinations in SY. Note that other vowel combinations, such as aCi, aCu, ọCi, and ọCu are also rare; these apparent gaps do not seem to be related to the topic at hand, and they are usually not discussed in harmony analyses. Some of the patterns marked as empty in this table do appear exceptionally, and I turn to this next.

Table 2.7. Distribution of V1 and V2 vowels in SY (C)V_1CV_2 words. A "✓" indicates existing vowel co-occurrences. A "?" indicates a possible gap, or near gap. The cells with diagonal lines indicate vowel combinations with [+ATR] and [−ATR] vowels. The shaded cells indicate words with only [−ATR]; the unshaded cells indicate [+ATR] words.

V1 \ V2	i	e	ẹ	a	ọ	o	u
i	✓	✓	✓	✓	✓	✓	✓
e	✓	✓				✓	✓
ẹ	✓		✓	✓	✓		✓
a	✓	✓	✓	✓	✓	✓	✓
ọ	✓		✓	✓	✓		
o	✓	✓				✓	✓
u							

2.7.2 Exceptions to harmony

The co-occurrence generalizations above are robust in terms of frequency of occurrence; however, some words violate these patterns.[11] For example, loan words may be disharmonic, as shown in (66).

(66) Disharmonic SY loan words, all from English words (from Bamgbose 1967)

fọ́tò 'photo'

mọ́tò 'car' < motor

télọ̀ 'tailor'

bébà 'paper'

Disharmony is also found in words that have undergone consonant deletion and subsequent vowel assimilation, as seen in the alternations in (67). In the alternations

[11] The equivalent data for Àkùrẹ́ and Mọ̀bà are not found in the literature, and were not investigated.

on the right side, where the consonant and high vowel remain, ATR harmony is blocked by the opaque high vowel. When the vowel and consonant are deleted, the harmony still does not occur.

(67) Disharmony after consonant deletion (Archangeli and Pulleyblank 1989)

eèpẹ̀ ~ erùpẹ̀	‘soil’
oódẹ ~ odídẹ	‘Grey Parrot’
yoòbá ~ yorùbá	‘Yorùbá’
òótọ́ ~ òtítọ́	‘truth’ (Abraham 1958)
òòka ~ òrùka	‘ring’ (Akinlabi, in preparation)

Words with morphological complexity may contain disharmonic sequences (68).

(68) SY morphological complex disharmonic words (from Akinlabi, in preparation, except where noted)

ẹyẹlé	‘pigeon’ < ẹyẹ, ‘bird’, ilé, ‘house’					
ọmọge	‘girl’ < ọmọ, ‘child’, oge, ‘ostentation’					
ọmọdé	‘child, young person’ < ọmọ, ‘child’, + edé ‘shrimp’					
ewébẹ̀	‘vegetables for soup’ < ewé ‘leaf’ + ọbẹ̀ ‘soup’					
ìkọ́lé	‘building’	from	i + prefix	kọ́ + build	ilé house	(Bamgbose 1967)

However, we also find monomorphemic non-loan words which are disharmonic. While having the same appearance as the words in (68), the examples shown in (69) have no apparent internal morphemic structure, suggesting that the disharmonies in these cases have been lexicalized, and not the result of synchronic phonological processes.

(69) Disharmonic monomorphemic words (Abraham 1958, unless noted)

erá	ó já sí erá, 'he disappeared', < ?? + ?rá, 'vanished'
èròjà	'ingredients' < ??
ẹ̀dọ̀foro	'lungs' < ẹ̀dọ̀, 'liver'; foro, ??
kòròba	'bucket' < ?? (possibly a loan word, Akinlabi, pc)
òòlà	'wedge' < ??, là 'to split'
eèrà, èèra	'type of ant' < ??; Fresco (1970)
èépá	'scab on a wound that is healing', pá, 'to cause to shrink'
èèta	'coarse flour remaining after èlùbọ́ has been removed' < ??

The robust vowel generalizations, as well as the exceptions, provide the necessary background for analysis of SY vowel harmony.

2.7.3 Previous analysis of SY vowel patterns

Of the several formal accounts of SY vowel harmony, I examine the rule-based account of Archangeli and Pulleyblank (1989). To a considerable degree, the constraint-based accounts of Bakovic (2000) and Akinlabi (in preparation) contain the similar principles couched in Optimality Theoretic terms. My inclusion of a rule-based account instead of a constraint-based account is motivated by an interest in brevity; I believe the discussion would not differ if a constraint-based account were analyzed. Archangeli and Pulleyblank motivate the radical underspecification of the feature [ATR], with [ATR] being specified at the morpheme level as opposed to the phoneme level. They capture many of the relevant distributional generalizations of Yorùbá by positing a few principled rules and constraints, shown in (70).

(70) (i) a co-occurrence constraint against [+high] and [−ATR],

(ii) the rightward association of ATR,

(iii) the redundancy specification of [−ATR] to [+low] vowels,

(iv) the leftward spreading of [−ATR] within the word, and

(v) the redundant specification of [+ATR] to all unspecified vowels.

These few statements enable them to account for the following generalizations, among others. The absence of [−ATR] high vowels follows from (70i) above. The absence of words with adjacent mid vowels with differing ATR values follows from (70ii), (70iv) and (70v). The absence of words with forms like **eCa*, and the existence of words with forms like *ẹCa*, *aCe*, and *aCẹ* follows from (70iii) and (70iv). Since /a/ is specified as [−ATR] before leftward spread, this results in *ẹCa*, and rules out **eCa*. In the case of *aCe*, the morpheme is unspecified for ATR, with /a/ getting its [−ATR] specification due to (70iii) above. For *aCẹ*, the morpheme is specified for [−ATR], so /a/ gets its [−ATR] specification from spreading. Finally, the existence of words like *èlùbọ́*, 'yam flour', *àkùrọ̀*, 'a type of farmland', and *àbúrò*, 'younger sibling' follows from all of the rules mentioned above. The non-existence of words of the shape in (71) occurs because the [−ATR] feature for the morpheme associates to the rightmost vowel, as per (70ii), and can not spread left due to rule (70i), so the initial mid vowel surfaces as [+ATR] due to (70v).

(71) *<+mid [−ATR]>C<+high>C<−high>, as in *ẹCuCo, *ẹCuCa, *ẹCuCẹ

While capturing this gap is an elegant feature, some words are problematic for this analysis:

(72) Exceptions to *<+mid [−ATR]>C<+high>C<−high>

(a) ọtíkà, ‘guinea corn wine’ > ọtí ‘wine/beer’ + ọkà, ‘guinea corn’
Archangeli and Pulleyblank (1989)

(b) ẹ̀rúnlá/ọ̀rúnlá, ‘okra seed’ > ẹ̀rún ‘particle’, ilá ‘okra’
Archangeli and Pulleyblank (1989)

(c) pọ̀kìyà, ‘fat and round’ (Abraham 1958)

(d) ọlidé, ‘holiday’ (Abraham 1958) > ‘holiday’ English

The first two cases (72a, b) are said to be compounds, so they are polymorphemic and therefore do not violate the principles above. In the case of *ọtíkà* (72a), however, we would expect **ọtọ́kà* if this were a synchronic process, since /i/ always elides in this environment (Pulleyblank 1988a). Instead, the existence of /i/ suggests that *ọtíkà* is monomorphemic, with its diachronic origin from *ọtí* and *ọkà*. In the case of (72b), a synchronic analysis works for *ẹ̀rúnlá*, but not for *ọ̀rúnlá*, and only the latter is found in Abraham’s (1958) dictionary. The remaining examples (72c, d) are monomorphemic, and remain exceptions under Archangeli and Pulleyblank. It is noteworthy that other unexpected gaps or near gaps are not accounted for in their analysis, such as: *VCoC(e|u), *VCeC(i|u), *VCaCe, *VCiCe, and *VCuCi. It is not clear if there is a principled way of determining which gaps if any can remain outside of an account, but the fact that these gaps remain unaccounted for reduces the importance of accounting for the gap in (71).

Awobuluyi (1967) and Stahlke (1976) argue that some SY nouns are made up of a vowel prefix and a CV root, as shown in (73).

(73)	è-rò	'thought'	< rò	'think'
	ẹ̀-rí	'evidence'	< rí	'see'
	ẹ̀-rọ	'machine'	< rọ	'fabricate'
	ì-là	'line'	< là	'split'

Archangeli and Pulleyblank (1989) agree that some or all VCV nouns consist of a vowel prefix and a CV root. In their account of the morphology, they cite these alternations to strengthen their analysis, shown in (74). In particular, because the prefixes agree in [ATR] value with the vowel from the root verb, they view this as evidence of harmony.

(74) Proposed nominalizing prefixes (from Archangeli and Pulleyblank (1989:188).

a.	ọ-dẹ	'hunter'	dẹ	'hunt'
b.	ọ̀-ta[12]	'person who is a good shot'	ta	'shoot'
c.	ò-kú	'corpse of person'	kú	'die'
d.	ẹ̀-rú	'the haft'	rú	'haft'

In these examples, the prefixes agree in ATR with the verbs, with the exception of (74d), for which they posit an underlying floating [−ATR] specification in the stem /rú/. Akinlabi (in preparation) lists many more examples of VCV nouns and related CV stems, which he considers to be bimorphemic, consisting of prefix and stem. He shows that they exhibit the same vowel co-occurrence restraints as monomorphemic VCV nouns, that is, nouns which are not derived from any apparent root.

If these types of patterns above indicate productive phonology, then these are arguably the only productive alternations SY exhibits. However, it is not clear that

[12] Archangeli and Pulleyblank (1989) has *ọta*; Abraham (1958) writes *ọ̀ta*.

these alternations are productive. I present evidence here that argues against the view that the examples above exhibit productive morphology. While it is clear that some VCV nouns are related to CV verbs, as in the examples above, it is more parsimonious to posit a diachronic relationship between the words and their roots, for the following four reasons. First, the phonological behavior of the group of words considered to be bimorphemic does not differ from those considered to be monomorphemic, as Akinlabi (in preparation) points out. Second, the prefixes are not productive. According to Yorùbá consultants, one can not create new words with these prefixes, as shown in (75).

(75) SY verbs with no related nouns

jẹ	'eat'	*ẹ̀jẹ
sọ	'speak'	*ẹ̀sọ
sè	'cook'	èsè, 'journey provisions'; but not 'cooked food in general'

Third, we find a continuum of relationships between VCV nouns and CV verbs ranging from very clearly related (*èrò*, 'thought' < *rò*, 'think'), to slightly related (*èsè*, 'journey provisions'< *sè*, 'cook') , to word pairs with no clear relationship (*èrú*, 'dishonesty, shiftiness' < *rú* 'cause muddle') suggesting a lexicalization of the nouns. Finally, the prefixes only harmonize with the following vowel when the root vowel is mid or low. But when the root has a high vowel as in (76), we find in some cases a [+ATR] [e], and other cases a [−ATR] [ẹ]. This is not what is *a priori* expected. If the prefixes harmonized with the [+ATR] high vowels, we would expect only [+ATR] [e] before high vowels. In light of this fact, Archangeli and Pulleyblank posit an underlying floating [−ATR] specification in the stem, which, unable to link to the high vowel due to the constraints of SY, links instead to the prefix /e/, resulting in the forms shown in (76).

(76)	ẹbí	'relations'	< bí	'give birth to child'
	ẹ̀rúkọ́	'hoe haft'	< rú	'haft (verb)', ọkọ́ 'hoe'
	ẹrù	'load'	< rù	'carry'
	ẹ̀-rú	'the haft'	< rú	'haft'

Nouns of similar form without apparent morphemic derivation are found, such as *ẹrú*, 'slave'. The word *rú* can mean 'caused muddle', 'sprouted', or 'hafted', but not *'enslaved'. Under their analysis, these nouns must presumably either have prefixes and an abstract root with a floating [−ATR] feature, or, they must have a different derivation that arrives at the same surface form. Neither of these solutions is optimal.

Evidence against the floating feature position comes from Mọ̀bà Yorùbá, which shows virtually the same set of facts in relevant areas. That is, *ẹrù* 'load, luggage' and *ẹ̀rí* 'evidence' both exist, as do *rù* 'load' and *rí* 'see'. Presumably, the analysis posited for SY would be required for Mọ̀bà since the motivation for positing the [−ATR] floating feature in Mọ̀bà is as compelling as it is in SY. However, in Mọ̀bà, a further environment can be examined to corroborate this floating feature. If a floating feature is posited in *rù* 'load', and *rí* 'see', then we should expect this feature to be realized in Mọ̀bà's harmonizing environment from (39), above. That is, we would expect the [−ATR] floating feature to be realized on the subject pronoun in the sentences such as in (77).

(77) Predicted harmony in Mọ̀bà, under floating [−ATR] feature analysis

rí   rù
 :   :
[−ATR]   [−ATR]

* ẹ́ rí oko	's/he saw the farm'
* ẹ́ ru ẹja	's/he carried fish'
* ẹ́ bí ọmọ	's/he bore the child'

However, the facts of Mọ̀bà do not support the floating [−ATR] hypothesis, because for these verbs and all high vowel verbs, the pronouns are realized as [+ATR] vowels as seen in (78).

(78)	é rí oko	's/he saw the farm'
	é rú ọkọ́	's/he hafted the hoe'
	é ru ẹja	's/he carried fish'
	é bí ọmọ	's/he bore the child'

This suggests that no floating feature is present underlyingly in these verb stems in Mọ̀bà. The [−ATR] feature must then be present on the initial vowel, that is, the posited prefix. Because of this, it becomes difficult to support the hypothesis that these words are made up of a stem with the same prefix since now the prefixes must differ in [ATR] feature from one word to the other. If this is the case in Mọ̀bà, then one can assume that the same applies to SY, unless additional evidence exists to suggest otherwise.

Additionally, we may expect to find evidence in Àkùrẹ́ that a [−ATR] feature exists in these high vowel words, since [−ATR] high vowels are allowed. However, we do not find these words realized as [−ATR] in Àkùrẹ́; they surface with [+ATR] vowels, as shown in (79).

(79) Àkùrẹ́ nouns

rù	'carry', *rụ̀	erù	'load', *ẹrụ̀
bí	'be born', *bị́	ebí	'family, relations', *ẹbị́
		erú	'slave', *ẹrụ́

A simpler explanation, in light of the evidence presented here, is that at least some of the initial vowels have their origins in prefixes and that synchronically these words such as those in (73) and (74) are monomorphemic. Under this explanation, the vowel distribution patterns found in such words would not require a synchronic analysis. In particular, the appearance of the [−ATR] initial vowel in the examples in (74) would not require an abstract analysis, but would instead be marked in the lexicon.

2.7.4 Analysis

In presenting an analysis of SY vowel harmony, two questions are critical. What are the generalizations to be accounted for, and how are they represented? The major generalizations have been described above. From the lexicon of SY, the co-occurrence patterns discussed above are robust, though not without exceptions, as seen in §0. While Archangeli and Pulleyblank (1989), Bakovic (2000), and Akinlabi (in preparation) motivate a productive phonology, it may be preferable to posit an alternative position without productive phonology since positing phonological constraints or rules for non-productive alternations offers no clear advantage to the learner or speaker. If the analysis of the verb prefixes from §2.7.3 is correct, then the patterns do not indicate productive phonology. Pierrehumbert (2003), and references therein, shows that speakers are aware of many statistical generalizations in their language, and it is likely that Yorùbá speakers internalize the most robust patterns and thus have an awareness of the make up of typical Yorùbá words. It is not clear if the distribution of the SY lexicon necessitates positing such knowledge in phonological terms. A thought experiment may help to clarify this point.

Imagine two SY learners. Learner A has some kind of phonological encoding of the vowel distribution. Learner B has no such encoding, that is, she approaches the vowels of the language in the same way she would approach the vowels of Italian, where there is no pervasive vowel distribution pattern. Each learner is exposed to SY, a language with vowel distribution patterns that exist prior to the learners' exposure. In order to learn the lexicon without the help of a constraining phonology, Learner B would have a task no more difficult than that of an Italian learner. If there are gaps in the data to which she was exposed, then there will be gaps in her lexicon. It is therefore not necessary for a learner to posit a phonology in order to learn the vowel patterns of the language.

If the learner does not require knowledge of the patterns to learn a lexicon, that does not mean that the learner has no awareness of the patterns. To my knowledge, no tests of SY speakers have been performed to determine the extent to which SY speakers have knowledge of the vowel patterns of their language. But it is reasonable to assume that some or all SY speakers have some testable knowledge of SY vowel patterns. However, the vowel patterns in a speaker's lexicon must emerge from the lexicon itself, and may not be an essential part of the speaker's phonological knowledge. That certain patterns exist is not in doubt, the question is whether we must posit an abstract phonology to account for the patterns, or whether it is more parsimonious to assume a simpler view: that learners learn the items of a language, and then later infer patterns from those items, not as a prerequisite for learning the patterns, but rather as a result of hearing and speaking their language.

2.7.5 Summary of SY vowel harmony

The evidence presented above suggests that ATR harmony is not an active part of the phonology of Standard Yorùbá, and is rather an historical remnant of a pattern found in an earlier language resembling perhaps the pattern found in related dialects. This view is not the prevalent view among many Yorùbá phonologists, who instead argue that [ATR] is an active feature in SY, for example, see Pulleyblank (1988a, b), Archangeli and Pulleyblank (1989, 1993, 1994), Akinlabi (in preparation), and Bakovic (2000). The position I have advanced in this section is not critical for the major question investigated in this study, that coarticulation leads to vowel harmony; however, it is related. What I hope to have done is to lay out the generalizations of vowel distribution in SY, and at the same time present the possibility that one does not need to posit a synchronic phonological analysis such as that presented in Archangeli and Pulleyblank (1989) in order to account for the facts of SY vowel distribution.

2.8 Historical evidence for no high vowel harmony in Proto-Yorùbá

In this section, I provide evidence that the high vowel harmony found in Àkùrẹ́ and closely related dialects is an innovation that emerged from a Proto-Yorùbá-Igala (PYIG) system without high vowel harmony. This view has been put forth by Fresco (1970), Oyelaran (1973) and Capo (1985), but see Bamgboṣe (1967), Adetugbo (1967:156) and Akinkugbe (1978) for dissenting views. My use of the term Proto-Yorùbá-Igala refers to the language at the Yoruboid level of Figure 2.1, following Akinkugbe (1978). That is, PYIG is posited as the predecessor of Igala and the Èdèkiri languages.

A critical part of this discussion is determining the placement of Mọ̀bà within the Èdèkiri language group. Above, I tentatively placed Mọ̀bà in the Central Èdè group along with Àkùrẹ́. However, as has been demonstrated above, Mọ̀bà and Àkùrẹ́ differ considerably with respect to high vowel harmony. Adetugbo (1967) lays out some differences between Central and NW Èdè, of which some of the major ones are listed in (80):

(80) Differences between Central Èdè and NW Èdè (Adetugbo 1967)

	CE	NWE
(a)	'u' initial words	no 'u' initial words
(b)	high vowel harmony	no high vowel harmony
	9 oral vowel phones	7 oral vowel phones
(c)	$\boxed{\text{è}}$rù, 'fear'; $\boxed{\text{ẹ̀}}$ụ́rẹ́, 'goat'	$\boxed{\text{ẹ}}$rù, 'fear'; $\boxed{\text{e}}$wúrẹ́, 'goat'

These differences are quite significant, especially (a) and (b), and they are robust. CE dialects allow initial /u/, while in NWY dialects, /u/ has merged into /i/. Mọ̀bà patterns with CE dialects, as shown in (81).

(81) /u/ merges with /i/

SY	Mọ̀bà	Àkùrẹ́	
ilé	ulé	ulé	'house'
iṣu	uṣu	uṣu	'yam'
igbá	ugbá	ụgbá	'calabash'
ikọ́	ukọ́	ụkọ́	'cough'
igi	igi	igi	'tree'

As for the high vowel differences in (80b), Mọ̀bà patterns with NW Èdè dialects in not having high vowel harmony, as we have seen above. Finally, (80c) shows how words with an initial mid vowel and following high vowel differ between the dialect groups. Central Èdè dialects do not allow [−ATR] mid vowels before [+ATR] high vowels; while the other dialect groups do. Table 2.8 shows these words in nine Yorùbá dialects from five dialect groups. What is notable is that in the Central Èdè dialects (Ìjẹ̀ṣà, Èkìtì, Àkùrẹ́, and Ìfàkì) but not Mọ̀bà, mid vowels are always [+ATR] before high vowels. In the other dialects, /ẹ/ is allowed before high vowels.

Table 2.8. ẹọCiu/eoCiu words in 9 Yorùbá dialects[13] (Akinkugbe 1978, Fresco 1970, Abraham 1958). Mọ̀bà data was provided by a language consultant and verified by Oladiipo Ajiboye, personal communication.

SE Èdè	**SW Èdè**[14]	**NE Èdè**	**NW Èdè**	**?**	**Central Èdè**				
Ondo	Ifẹ-Togo	Yagba	**SY**	**Mọ̀bà**	Ìjẹ̀ṣà	Èkìtì	**Àkùrẹ́**	Ìfàkì	gloss
ẹú	arú	arú	ẹrú	ẹrú	erú	erú	erú	erú	slave
		àrù	ẹ̀rù	ẹ̀rù	èrù	èrù	èrù		fear
ẹtu			ẹtu	ẹtu	etu	etu	etu	etu	duiker
ẹụ̀n	àrù		ẹrù	ẹrù	erù	erù	erù	erù	load
		ẹ̀rí	ẹ̀rí	ẹ̀rí	èrí				evidence
ẹ̀wù	àwù		ẹ̀wù	ẹ̀wù	èù	ewu	èwù, èù	èù	clothes, shirt
			ẹbí	ẹbí	ebí		ebí		relations
ẹtù			ẹ̀tù	ètù	ètù		ètù	ètù	gunpowder
			ọtí	(ọtín)			(ọtín)		alcoholic drink
ebi	ebi	ebi	ebi	ebi	ebi	ebí	ebi	ebi	hunger
etí	etí	etí	etí	etí	etí		etí	etí	ear
obì			obì	obì	obì		obì	obì	kola nut

[13] It is not clear to me if Ifẹ̀ (Togo) shares this characteristic; I found only one word of shape ẹọCiu in Armstrong's (1965) word list: *ẹ̀dù* 'chest'.

[14] Ketu, a SW Èdè dialect, does not fit this pattern. From the data in Fresco (1970) and Akinkugbe (1978) it appears more like the Central Èdè dialects: *erú* 'slave', *etu* 'duiker', *erù* 'load', *èwù* 'clothes', *etù* 'gunpowder', *ebi* 'hunger', and *etí* 'ear' are the same as Àkùrẹ́, but *ẹbí* 'relations' is like Mọ̀bà. This does not detract from the argument I advance.

To account for the patterning of Mọ̀bà with respect to (80), I propose the following historical relationship, in (82).

(82) Proposed classification of Mọ̀bà with NW and Central Èdè dialects.

Proto-NWC Èdè (7 surface oral vowel vowels)

<u>NW Èdè</u> (7) Proto-Mọ̀bà-Central Èdè (7)

SY

Ẹ̀gbá Mọ̀bà (7) <u>Central Èdè</u> (9)

Ọ̀yọ́ Àkùrẹ́

Ọ̀ṣùn Ìfẹ̀

Ìjẹ̀ṣà

Èkìtì

With this classification, we account for Mọ̀bà's differences from the Central Èdè dialects; that is, the high vowel harmony innovation would have occurred after the division of Mọ̀bà from the rest of Central Èdè. Similarly, we can see that the merger of initial /u/ and /i/ into /i/ took place in the NW Èdè branch, and thus Mọ̀bà and Central Èdè were not effected. With the high vowel harmony in Central Èdè, came the accompanying merger of ẹ-C-<high> into e-C-<high>, since we only see this merger in Central Èdè, and not in the other dialect groups or Mọ̀bà.

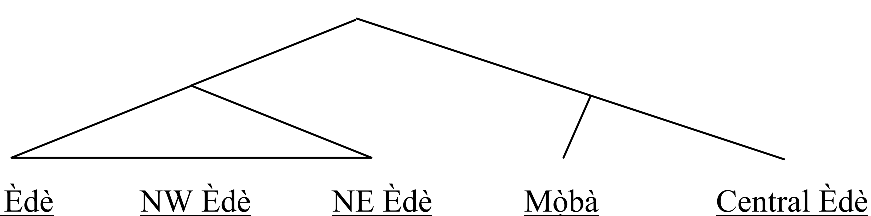
Expanding to a larger view, we see PYIG having no high vowel harmony, and both /u/ and /i/ as initial vowels. Starting with the initial /u/, evidence from Akinkugbe (1978) strongly indicates that Proto-Yorùbá had initial /u/ and /i/, as indicated by looking at the word 'calabash' in several dialects:

(83) ‘Calabash’ in several Yoruboid languages (Akinkugbe 1978, except Mọ̀bà):

Igala:	$\boxed{\text{ú}}$-gbá	(Igala)		
SE-Èdè:	$\boxed{\text{u}}$-gbá	(Itsẹkiri)		
NW-Èdè:	$\boxed{\text{i}}$-gbá	(SY)		
C. Èdè:	$\boxed{\text{ụ}}$-gbá	(Ìjẹ̀ṣà)		
Mọ̀bà:	$\boxed{\text{u}}$-gbá	(Mọ̀bà)		
NE-Èdè:	$\boxed{\text{u}}$-gbá	(Gbẹdẹ)	$\boxed{\text{i}}$-gbá	(Yagba)
SW-Èdè:	$\boxed{\text{u}}$-gbá	(Tsábẹ)	$\boxed{\text{i}}$-gbá	(Ìfẹ̀-Togo)

This example is typical; similar patterns are found for SY *iṣu*, ‘yam’; SY *iṣẹ́*, ‘work’; SY *ilé*, ‘house’; among others. Because there are /u/ initial vowels in Igala, SE-Èdè, and Central Èdè, as well as some in NE-Èdè and SW-Èdè, and because the same words, more or less, show up with an initial /u/ in those dialects that have them, it is uncontroversial that PYIG allowed initial /u/, and that somewhere in the history of the Èdèkiri branch /u/ merged with /i/, possibly more than once. The details of the branching of the Èdèkiri group are beyond this scope of this paper, but it seems plausible that the Central Èdè may have branched off earlier than the others, as shown in (84).

(84) Proposed classification Èdèkiri dialects of Yorùbá



The competing view, that PYIG or PY had high vowel harmony which has only been retained in Central Èdè dialects is harder to maintain. Because high vowel harmony only currently exists in the Central Èdè dialects, then high vowel harmony would have had to disappear three times independently: from Igala, from SE Èdè, from Mọ̀bà, and from the group now containing SW Èdè, NW Èdè, and NE Èdè. This is improbable, especially because there is no trace of high vowel harmony in any Yoruboid dialect outside of Central Èdè.

Corroborating evidence for an emerging high vowel harmony in Central Èdè comes from the words which in SY and Mọ̀bà have the form ẹ-C-<high vowel> and e-C-<high vowel> from above. Akinkugbe (1978) reconstructs the following forms for some of the words in (85).

(85) Some PYIG reconstructions (Akinkugbe 1978):

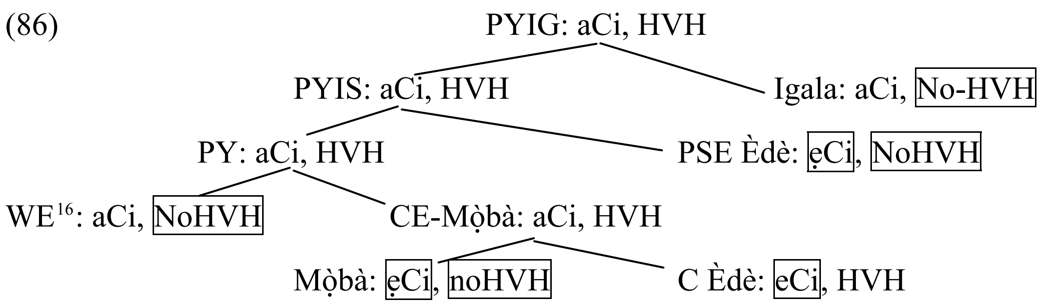
*a-rú	‘slave’	*e-bi	‘hunger’
*a-rù̃	‘load’	*e-tí̃	‘ear’
*à-rù̃	‘fear’		
*à-rí	‘evidence’		
*a-tu	‘duiker, antelope’		

The initial *e- in the right column has remained an /e-/ for all current PYIG dialects in Akinkugbe’s study for ‘hunger’ and ‘ear’. The initial *a- from the left column is clear in the reconstruction, as Igala retains the /a-/, and so do several Yorùbá dialects. The /a/ was raised to an /ẹ/ before a high vowel in Central Èdè, SE Èdè, and NW Èdè[15], a raising whose phonetic motivation is quite plausible. Then, with the advent of high vowel harmony in Central Èdè, the /ẹ/ raised to /e/ before the [+ATR] high vowels, giving rise to the forms found in Central Èdè dialects today.

[15] The fact that SW Èdè and NE Èdè didn’t raise /a-/ to /ẹ-/ in this context is problematic, considering that SE Èdè did raise /a/ to /ẹ/, and SE Èdè is posited to have split off earlier.

We do not expect the existence of ẹCi in a high vowel harmony system, so if high vowel harmony had been a feature of PYIG (or even of PY, excluding Igala), we would have to posit that the /a/ raising to /ẹ/ did not occur until the high vowel harmony had been lost. Again this would indicate a large number of independent changes in the same direction, as shown in (86), where the relevant characteristics appear after the dialect group name. ‘aCi’, ‘ẹCi’, and ‘eCi’ indicate the required form of the group of words under discussion. ‘HVH’ indicates high vowel harmony, and ‘No-HVH’ indicates no such harmony. A box indicates that the feature is an innovation for that branch.

(86)



In (87), we see a pair of words, related to the eCi words from above. They offer yet another piece of evidence in favor of the 7 vowel PYIG.

(87)	**Àkùrẹ́**	**Mọ̀bà**	**SY**	**Ondo**	**Tsabẹ**	
	ẹ̀ụ́rẹ̀	ewúrẹ́	ewúrẹ́	ewúẹ́	e-érẹ́	‘goat’
	ọrụ́kọ	orúkọ	orúkọ	oúkọ	oókọ	‘name’

In Àkùrẹ́, the [−ATR] feature has spread from the final [−ATR] mid vowel through the first, as expected, considering the high vowel harmony. In the case of the other

[16] Western Èdè (WE) is an *ad hoc* label representing the dialects: SW Èdè, NW Èdè, and NE Èdè.

dialects, the [−ATR] has apparently been blocked by the opaque high vowel. If the high vowel harmony had been present throughout PYIS, we would expect to see the forms 'ẹwúrẹ́' or 'ọrúkọ' with a [−ATR] initial vowel after the high vowel harmony had disappeared, especially because the high vowel harmony would have to disappear several times, as discussed above. In fact, we find that this is not the case, except in Isekiri where, unfortunately for my argument, we have ọrúkọ 'name'. Nonetheless, a 7-vowel PYIG hypothesis leads more clearly to the current pronunciations.

Additional support for a seven surface vowel PYIG comes from underlying forms. As discussed above, for all the Yorùbá dialects in the literature discussed here, there exist few words, if any, for which one must posit an underlying [−ATR] high oral vowel (though [−ATR] high nasal vowels exist, see Bamgbose 1965 and Adetugbo 1967). Adetugbo (1967:164-5) mentions two Central Yorùbá words, *wị́* 'say' and *ẹrụ́* 'slave', which may be exceptions to the patterns discussed so far. However, in this study Adetugbo discusses four CY dialects—Ife, Àkùrẹ́, Ijẹ̀ṣa, and Èkìtì—but does not specify which dialect contains these words nor whether there are more words like them. Certainly, a corroboration of these data is in order, especially since the [+ATR] version of *erú* 'slave' has been noted in the Central Yorùbá dialects of Àkùrẹ́ and Ifaki (Fresco 1970), Irun and Ijẹ̀ṣa (Akinkugbe 1978), and Èkìtì (my field data), and [+ATR] *wí* 'say' has been noted in Ifaki (Fresco 1970), Irun and Ijeṣa (Akinkugbe 1978). If the existence of the words with underlying [−ATR] high vowels in CY were verified, this would not be fatal to the line of argument taken here. In fact, finding such forms in only one of many high harmony dialects even suggests innovation. Certainly at some point, [−ATR] high vowels may develop in an Àkùrẹ́-type dialect. However, the absence of such vowels is expected if we accept that the high vowel harmony is a relatively recent innovation, since the harmony would then be the only source of [−ATR] high vowels. Even so, the absence of underlying [−ATR] high vowels in itself does not preclude the existence of Proto-Yorùbá high vowel harmony, for indeed high vowel harmony exists now without underlying [−ATR] high vowels.

2.9 Summary

In this chapter, I have laid out the phonological structure of Àkùrẹ́, Mọ̀bà, and Standard Yorùbá. The main elements of the three vowel systems are summarized in Table 2.9.

Table 2.9. The major elements of vowel harmony of SY, Mọ̀bà, and Àkùrẹ́.

	Harmony to Prefixes	Mid-vowel Split [e, o, ẹ, ọ]	High-vowel Split [i, u, ị, ụ]	Low-vowel Split
SY	X	✓	X	X
Mọ̀bà	✓	✓	X	X
Àkùrẹ́	✓	✓	✓	X

Standard Yorùbá and Mọ̀bà have an ATR distinction in mid-vowels and not high vowels—they have seven underlying and surface oral vowels. While Mọ̀bà has an ATR harmony extending to prefix clitics, SY does not. In Àkùrẹ́, mid vowels and high vowels have ATR distinctions; while Àkùrẹ́ has seven underlying oral vowels, it has 9 surface oral vowels. The [+high, −ATR] vowels, [ị, ụ], are not phonemic. Its harmony extends to prefix clitics. None of the dialects have an ATR split in the low vowel.

The evidence presented in §2.8 indicates that the high vowel harmony of Àkùrẹ́ (and related dialects) has emerged from a proto-language without high vowel harmony, a language with a vowel system similar to Mọ̀bà or SY. The experiment discussed in the rest of this study presents phonetic evidence that the seeds of a vowel harmony like Àkùrẹ́'s are present in Mọ̀bà and SY, and therefore were likely present in an earlier Yorùbá language.

CHAPTER THREE

HYPOTHESES AND METHODOLOGY

3.1 Introduction

This study was designed primarily to investigate the origins of phonological patterns. In three closely related dialects, one dialect contains a phonological pattern that is absent from the others. As discussed in chapter 2, Standard Yorùbá (SY), Mọ̀bà, and Àkùrẹ́ all have ATR harmony in mid vowels. However, only Àkùrẹ́'s harmony extends to high vowels. If the high vowel patterns emerged from phonetic patterns, then we should be able to find corresponding phonetic patterns in the related dialects in which the phonological pattern has not emerged. If these dialects contain phonetic patterns resembling the phonological patterns, and if the phonetic patterns have a physical explanation, then we can conclude that the phonological patterns have emerged over time from earlier phonetic patterns. More specifically, my claim is that high vowel harmony emerged from coarticulation in an older system without harmony.

In Chapter 2 (§2.8), I presented evidence that proto-Yorùbá had no high vowel harmony. If this is the case, the vowel system of proto-Yorùbá would have looked something like the vowel systems of Mọ̀bà and SY. Because I can not directly investigate the acoustic properties of proto-Yorùbá, I instead look at Mọ̀bà and SY as if I were looking at proto-Yorùbá, as illustrated in Figure 3.1.

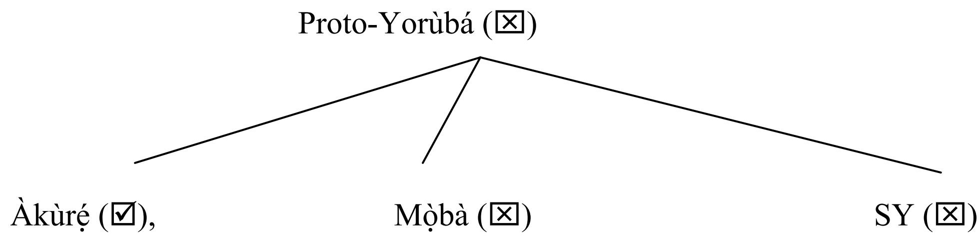


Figure 3.1. Hypothesized relationship between Proto-Yorùbá and Yorùbá dialects of study. ☑ - indicates high vowel harmony, ☒ - indicates no high vowel harmony.

By finding the predicted coarticulatory patterns in Mọ̀bà and SY, I can posit that Proto-Yorùbá exhibits similar patterns, thus supporting the claim that coarticulation is the origin of vowel harmony in Àkùrẹ́.

I expect to find coarticulatory effects in Mọ̀bà and SY that resemble the patterns of high vowel harmony in Àkùrẹ́. I do not expect the coarticulatory effects to be as large and robust as the harmony patterns for two reasons. First, we expect phonological effects to be more pronounced than phonetic ones. I have shown in chapter 2 that [±ATR] high vowels participating in phonology in Àkùrẹ́ but not in SY or Mọ̀bà; therefore it is expected that the Àkùrẹ́ differences will be more robust. If, on the other hand, the Mọ̀bà and SY vowel effects were found to be as great as the vowel harmony patterns of Àkùrẹ́, then we would conclude that Mọ̀bà and SY also have a high vowel harmony, albeit one that differs in scope from Àkùrẹ́'s. Second, speakers of SY and Mọ̀bà do not report that high vowels are pronounced differently in different contexts; neither, for that matter, do linguists. For Àkùrẹ́, on the other hand, some speakers, although not all, recognize that the high vowels are pronounced differently in different contexts.

It is important to keep in mind that although the claim here is that ATR vowel harmony has emerged from coarticulation, it is not the case that coarticulation invariably leads to vowel harmony, even in Yorùbá. In fact, the classification of

Yorùbá dialects in Figure 3.2 shows that it is likely that high vowel harmony emerged only once, in the branch of Central Èdè that includes Àkùrẹ́. The other dialects did not develop high vowel harmony at the time the Central Èdè branch did, nor have they done so since, which suggests that the coarticulatory patterns can not be so prominent that the development of high vowel harmony is inevitable. It is not at all clear just how prominent a phonetic pattern must be if it stands a chance of becoming phonologized, but it is clear that some patterns are more likely than others to become phonologized.

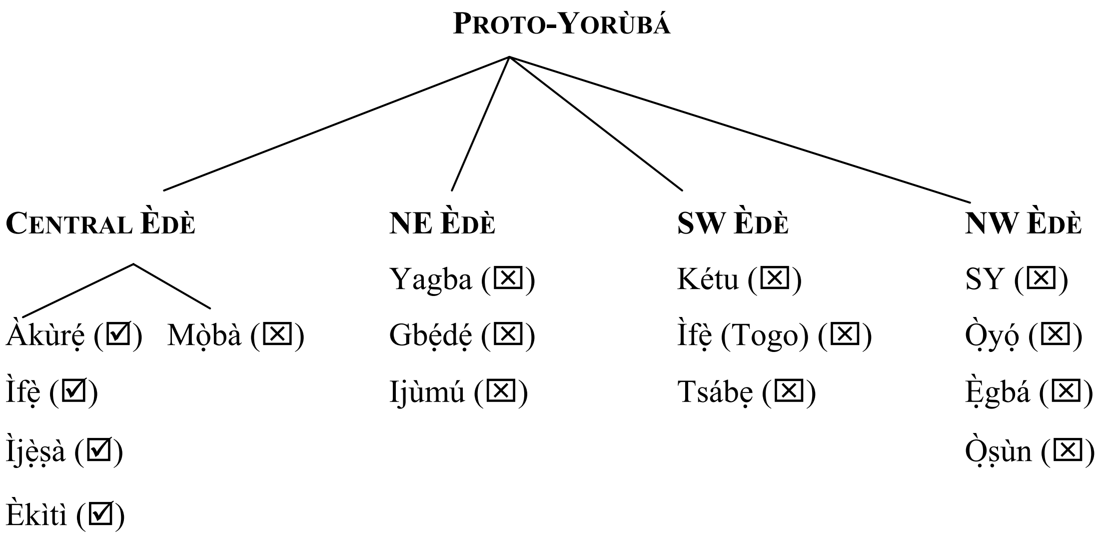


Figure 3.2. Classification of Yorùbá dialects. ☑ - indicates dialect with high vowel harmony; ☒ - indicates no high vowel harmony.

In the rest of this chapter, I lay out the methodology of the experiment. In §3.2, I present a sketch of the experiment. In §3.3, the word lists and carrier phrases are discussed. §3.4 details how the data collection was accomplished. The segmentation and labeling of the data are discussed in §3.5. An important initial step in this study is detailed in §3.6; the goal of this step is to find an acoustic property that can distinguish the [ATR] sets in the three dialects. §3.7 discusses the different kinds of measured effects used to evaluate vowel sets in this study. The conclusion follows in §3.8.

3.2 Overview of the experiment

In (1), I show an example of two Àkùrẹ́ sentences using nonsense words. The second vowels (V2) differ—V2 in (1a) is [+ATR] [e]; in (1b), it is [−ATR] [ẹ]. The initial vowels (V1) exhibit phonetic differences characteristic of vowel harmony, with V1 in (1a) showing [+ATR] characteristics, and V1 in (1b) showing [−ATR] characteristics.

(1) Comparison of environments of Àkùrẹ́

			V1 C V2		
(a)	ó	f' #	$\boxed{\textbf{i}}$----k---**e**	#	s'okè.
(b)	ọ́	f' #	$\boxed{\textbf{ị}}$----k---**ẹ**	#	s'okè.

In the similar environment in (2) shows two sentences for Standard Yorùbá. The V2s differ, as for Àkùrẹ́. However, in SY, the V1 for both (2a) and (2b) is [i].

(2) Comparison of environments of Standard Yorùbá

			V1 C V2		
(a)	ó	f' #	$\boxed{\textbf{i}}$----k---**e**	#	s'ókè.
(b)	ó	f' #	$\boxed{\textbf{i}}$----k---**ẹ**	#	s'ókè.

I predict that the coarticulation in SY and Mọ̀bà will resemble the vowel harmony in Àkùrẹ́, in acoustic effect and direction, although not degree. This is schematized in the next two figures. Figure 3.2 diagrams the high vowel harmony of Àkùrẹ́.

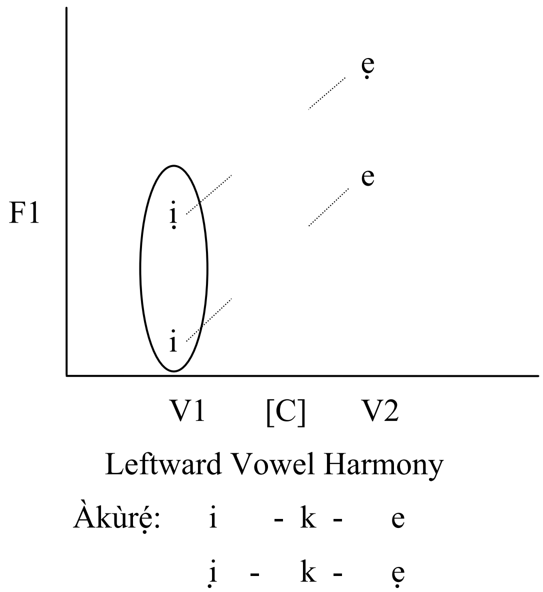


Figure 3.2. Schematic of predicted F1 value for vowels in Àkùrẹ́ tokens /*ike*/ and /*ịkẹ*/.

The F1 values clearly differ in [ike] compared with [ịkẹ], for both V1 and V2. Figure 3.3 diagrams the dialects with no high vowel harmony.

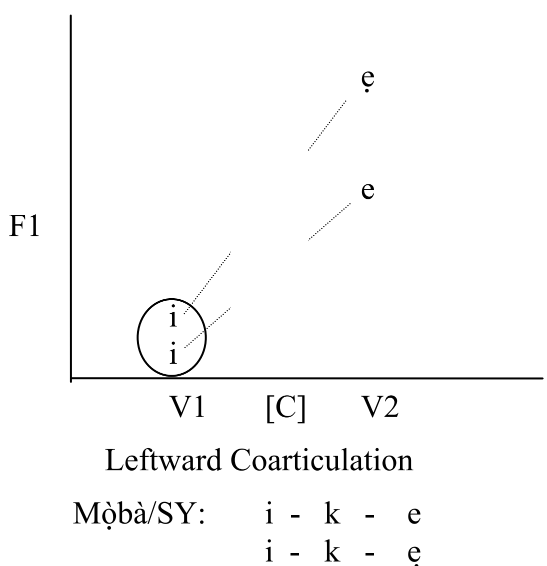


Figure 3.3. Schematic of predicted F1 value for vowels in Mọ̀bà and SY tokens /*ike*/ and /*ikẹ*/.

Whereas the F1 values for V2 are the same as found in Figure 3.2, the values for the V1s differ. Because they are the same phone, the F1 values are nearly the same.

However, because the speaker is anticipating the following vowel during the first vowel and the following consonant, the anticipatory coarticulation results in a slightly lower F1 for [i] in [ike] compared with [i] in [ikẹ]. For the same reasons, /u/ is predicted to behave the same way.

In this study, I investigate how local coarticulation may resemble and lead to local high vowel harmony. However, I do not attempt to address how a nascent harmony spreads beyond the immediate coarticulation domain; for example, if /ikẹ/ becomes /ịkẹ/, why does /ikikẹ/ become /ịkịkẹ/?

3.3 Word lists

Vowel to vowel coarticulation effects are often quite small. In order to measure the effects one must control for factors that could affect the degree of coarticulation, factors such as intervocalic consonant, tone pattern, and stress,[1] among other things. As it may be difficult or impossible to find a set of real words that differ minimally in these properties, non-words are often chosen for coarticulation experiments, as in Öhman (1966), Recasens (1987), Manuel (1990), Choi and Keating (1991), and Majors (1998), among others. In this study, a list made of VCV words and non-words was developed for each of the three dialects. The items have the general shape shown in (3) and include all allowable vowel combinations.

(3)	V1	–	C	–	V2
	/i, e, ẹ, a, ọ, o, u/	–	C	–	/ i, e, ẹ, a, ọ, o, u /

VCV words containing mid vowels with contrastive ATR values, such as eCẹ, ẹCe, oCẹ, and ẹCo, are not found in any of the three dialects,[2] so these combinations are

[1] Stress is not considered in this study, since it is not a distinctive feature in Yorùbá.
[2] With some exceptions, noted in Chapter 2, §2.7.2.

not included in any of the lists. Other combinations are found in words of some dialects but not others. For example, the list for Mọ̀bà and SY speakers includes forms with the shape {e/o}-C-{i/u} and {ẹ/ọ}-C-{i/u} since words with these patterns are attested in those dialects; on the other hand, Àkùrẹ́ speakers were given only the former set, since Àkùrẹ́ has no words of the latter set, {ẹ/ọ}-C-{i/u}. Words with initial /u/ are not found in SY, so they are not included in the SY word list.[3] In Table 3.1-Table 3.3, the vowel combinations used for the three word lists are shown.)See Chapter 2 for a discussion of co-occurrence of vowels in VCV words.) Note that the vowels indicated are phonemic. The surface vowels differ, although only for Àkùrẹ́ high vowels, where /i/ and /u/ have [±ATR] allophones [i], [i̧]; and [u], [ụ], respectively.

Table 3.1. Combinations of phonemic vowels for VCV word list: Àkùrẹ́.

		V2						
		i	e	ẹ	a	ọ	o	u
	i	✓	✓	✓	✓	✓	✓	✓
	e	✓	✓				✓	✓
	ẹ			✓	✓	✓		
V1	a	✓	✓	✓	✓	✓	✓	✓
	ọ			✓	✓	✓		
	o	✓	✓				✓	✓
	u	✓	✓	✓	✓	✓	✓	✓

[3] In retrospect, it would have been interesting to record items with patterns that are rare or are absent from the lexicon.

Table 3.2. Combinations of phonemic vowels for VCV word list: Mọ̀bà.

		V2						
		i	e	ẹ	a	ọ	o	u
	i	✓	✓	✓	✓	✓	✓	✓
	e	✓	✓				✓	✓
	ẹ	✓		✓	✓	✓		✓
V1	a	✓	✓	✓	✓	✓	✓	✓
	ọ	✓		✓	✓	✓		✓
	o	✓	✓				✓	✓
	u	✓	✓	✓	✓	✓	✓	✓

Table 3.3. Combinations of phonemic vowels for VCV word list: Standard Yorùbá.

		V2						
		i	e	ẹ	a	ọ	o	u
	i	✓	✓	✓	✓	✓	✓	✓
	e	✓	✓				✓	✓
	ẹ	✓		✓	✓	✓		✓
V1	a	✓	✓	✓	✓	✓	✓	✓
	ọ	✓		✓	✓	✓		✓
	o	✓	✓				✓	✓
	u							

Impressionistically, tonal patterns in Yorùbá are sometimes accompanied by non-tonal characteristics such as breathiness or vowel lengthening. It is likely that tonal patterns interact with vowel quality, and for this reason, tone was controlled. A mid-mid tonal pattern was selected because it is the most common tonal pattern in Yorùbá, and because no remarkable voice quality is associated with this tonal pattern, unlike some other patterns; for example, low-low may exhibit a breathy voice quality.

While Yorùbá orthography has tone marks for high and low tone, tone is frequently unmarked in written Yorùbá—for example, many newspapers do not mark tone—and readers use contextual clues to determine tone. Because the experiment used nonsense words, it was anticipated that readers would have difficulty reading the non-word tonal patterns correctly if they varied from item to item since contextual clues would necessarily be absent. In a pilot experiment where marked tone varied from token to token, the subject did indeed have difficulty with the tone. However, in an experiment with a fixed tonal pattern, the same subject had no difficulty reading non-words once the fixed pattern was established. Indeed, after some initial practice, the speakers in this experiment had little trouble reading the non-words since the tone of the items did not vary.

The intervocalic consonant was controlled since the identity of the intervening consonant greatly affects the realization of the vowels, both due to the influence of the consonant on the vowel (Keating et al. 1994), and due to the differing amounts of vowel-to-vowel coarticulation allowed by the consonant (Öhman 1966, Recasens 1987, Choi and Keating 1991). Pilot studies corroborated the difference in consonant interference (see chapter 4). An additional pilot study suggested that consonant voicing by itself, in this case [t] versus [d], was also found to affect the F1 of the adjacent vowel, and therefore consonants that differed only in voicing were not grouped together. The segment boundaries of obstruents are considerably easier to identify during speech analysis than those of sonorants. I therefore chose to use obstruents in this study. Of the obstruents, stops were chosen over fricatives, since fricatives tend to inhibit vowel-to-vowel coarticulation more than stops do due to their constrained articulation (Recasens 1984, 1985). I selected more than one consonant to ensure that the patterns found were not specific to one consonant, but to keep the word lists a manageable length for the speakers, the number of consonants was limited to two. The consonants /k/ and /b/ were chosen because they were obstruents that differ in more than one feature—voicing and place. A pilot study with one Mọ̀bà speaker, Mb10, showed that the general pattern of results found for /b/ and /k/ were also found with other intervocalic stops. For details, see chapter 4, §4.3.2.1.

For Àkùrẹ́, there are 35 allowable vowel combinations, for a total of 70 words (due to two consonants); for Mọ̀bà, there are 39 allowable combinations, for 78 words; and for SY, 32 combinations, for 64 words. For each dialect, the words were randomized, and then minimally reordered so that like words were not adjacent. Filler words were added at the beginning and end of each list, these included real words, which served to model the tonal pattern for the subject. The filler words were not ultimately measured. The word list for each dialect is found in the appendix.

3.3.1 Carrier phrase

The words and non-words were inserted into the carrier phrases listed in (4)—in this example, the non-words *eke* and *ẹkẹ* have been inserted.

(4)	(a)	SY:	ó	f’	_eke_	s’ókè.
		Mọ̀bà:	é	f’	_eke_	s’ókè.
		Àkùrẹ́:	ó	f’	_eke_	s’okè.
			3sg	put	[word]	on top
			‘S/he put “eke” on top.’			
	(b)	SY:	ó	f’	_ẹkẹ_	s’ókè.
		Mọ̀bà:	ẹ́	f’	_ẹkẹ_	s’ókè.
		Àkùrẹ́:	ọ́	f’	_ẹkẹ_	s’okè.
			3sg	put	[word]	on top
			‘S/he put “ẹkẹ” on top.’			

The carrier phrases met four criteria. First, the meaning was the same across the three dialects and amenable to many different target words and non-words. Second, they allowed the target word to be flanked by voiceless consonants, in particular fricatives, allowing for easier segmentation. In an earlier pilot study, when the VCV token was immediately preceded by a vowel of the carrier phrase, it was difficult to determine the start of the target vowel. For example, had I used the SY carrier phrase that Laniran (1993) used for a tone experiment, shown in (5), it would have been problematic and time consuming to determine where the [ọ] of 'sọ' ended and the first vowel of the token began.

(5) sọ ______ kan sí i. (from Laniran 1993)
"Say ____ once more."

Because all words end in a vowel in Yorùbá, the only way to precede a VCV word with a consonant is for the previous vowel to elide. Since /i/ is the only vowel which, in certain domains, is systematically elided before any other vowel, a word with the shape Ci was selected to precede the target word. The /i/ of *fi* is always deleted, leaving *f'VCV* in all three dialects. The third criterion was that the target word should not be phrase final, thus avoiding the problems of final lengthening (Gaitenby 1965; Oller, 1973; Klatt, 1975; Wightman et al. 1992). Finally, the selected carrier phrases differ minimally from dialect to dialect, so that comparisons can be made across dialects. Note that the pronouns differ in all three dialects, and the pronouns in Mọ̀bà and Àkùrẹ́ harmonize with the ATR value of the following vowel while the SY pronouns do not. In addition, the mid tone of the first vowel of *s'okè* in Àkùrẹ́ differs from the other dialects, in which the words have a high tone, *s'ókè*. One criterion that was not met was to have identical consonants flanking the VCV token as for the 'f's in the hypothetical sentence *o f'*$\boxed{\textit{ike}}$ *f'oke*. This criterion was sought in order to balance the peripheral coarticulatory influences on V1 and V2, so that differences in the vowel qualities could be attributed to position and not coarticulation. I found no such carrier phrase that met this criterion as well as the more important criteria already mentioned. However, this criterion was not essential for the purposes of this study.

The carrier phrases were chosen with the assistance of consultants from each dialect. In the case of Mọ̀bà, however, two other speakers later said that the phrase *fi…s'ókè* was not a pure Mọ̀bà phrase, but rather a SY phrase, with the Mọ̀bà equivalent being *mu…s'ókè* 'to put up high'. Because this was mentioned when recording had already begun, and because the speakers were nonetheless able to use the phrase while maintaining Mọ̀bà phonology (most notably, the pronoun alternations), I continued to use the phrase as written above. In fact, most of the Mọ̀bà speakers did not note this problem, and no one appeared to have any difficulty with the phrase.

3.4 Data collection

3.4.1 Subjects

I traveled to Nigeria in January, 1999 to record speakers for this study. The SY speakers were predominantly students from the University of Ibadan. For the Mọ̀bà speakers, I visited the town of Ọ̀tùn-Èkìtì. For the Àkùrẹ́ speakers, I went to the city of Àkùrẹ́. For each of the three Yorùbá dialects, ten adult native speakers were recorded, of which nine were male and one female. Since formant values for vowels of men and women differ considerably, it is easier to make comparisons when all subjects are of the same sex. Due to cultural considerations, it was somewhat easier for me (a male) to work with male speakers.

I asked speakers to answer several questions about their language use and history (see appendix) in order to determine their language background. In the cases where the speaker spoke little English, I was assisted by a translator who spoke the relevant dialect. For each Mọ̀bà subject, Mọ̀bà Yorùbá was the language spoken in their childhood home. All of the Mọ̀bà speakers, except Mb10, spoke Mọ̀bà in the home at the time of the recording. The same pattern was true for the Àkùrẹ́ and SY speakers. Most of the Àkùrẹ́ and Mọ̀bà subjects also speak Standard Yorùbá. In addition, most of the subjects speak English as a second language.

Subjects were offered payment for their participation, though many refused. In the case of the Mọ̀bà speakers, the payments went directly to a community development fund, as suggested by my primary Mọ̀bà consultants, and agreed to by the speakers.

3.4.2 Recordings

One speaker, Mọ̀bà speaker Mb10, was recorded at the Cornell University Phonetics Lab sound-treated room using an Electro-Voice RE20 microphone and a Carver TD-1700 analog tape deck. The rest of the recordings were made in Nigeria with a Sony WM-D6C Professional Walkman analog tape recorder and a Shure SM10A head-mounted microphone[4] onto a Maxell XLII (CrO_2) cassette. These speakers were recorded in their homes, their workplaces, or in a hotel room. As there was often considerable ambient noise present, the close proximity of the head-mounted microphone to the speaker's mouth ensured that the noise did not appreciably affect the quality of the recordings. The head-mounted microphone is adjustable and conforms easily to different head sizes. The microphone was positioned about 5 centimeters from the corner of the speaker's mouth and not directly in front of the mouth. Despite some reservations I had, the speakers had no trouble getting used to wearing the head-mounted microphone, and seemed very comfortable with it.

[4] At least one linguist has told me that some consonants (particularly labial stops) are difficult to accurately record with head-mounted mikes; this was not a concern of mine, since I was primarily interested in vowels.

3.4.3 Procedure

Each subject was presented with the list, a combination of words and non-words. He was told that some of the words were real and others were not, and that they all had the same tonal pattern, which many Yorùbá speakers know as 're-re' (from the musical scale, *do re mi fa so la ti do*). The first few entries in the word list were common words with a mid-mid tone. In order to demonstrate the intended tonal pattern, I read these common words and some non-words, and then asked the subject to repeat them. The subject was asked to read each word in the carrier phrase. To give the sentences a clearer meaning, I demonstrated the sentences using a stack of cards with real and nonsense words written on them. I suggested that as they were saying the equivalent of "she put *ike* up high", they take a card with the non-word *ike* on it, and imagine putting it someplace high up. I had been apprehensive that some subjects might hesitate to utter these somewhat unusual sentences, but in fact none of the subjects had any trouble pronouncing them. The subject was instructed to speak his or her home dialect, and to speak in a natural manner. The sentences were practiced until the subject was habituated with the task, which usually took about a minute. The subject was informed that he could stop at any time, and that he should repeat a sentence if he was not satisfied with his pronunciation.

Each subject repeated each sentence four times. If I judged that the speaker pronounced a sentence unclearly or incorrectly, I asked him to repeat the sentence again. This was done only for obvious errors such as an incorrect tone or vowel, or a pause within the sentence. Since the experiment used mostly non-words, speakers occasionally pronounced the tone incorrectly when the non-words had real word counterparts with different tones. For example, because *oke* (with tone mid-mid) is not a Yorùbá word and *òkè* 'mountain' is, some speakers pronounced *oke* as *òkè*. In that case, the speaker was reminded to use a mid-mid tonal pattern. Occasionally, I modeled the non-word with the correct tonal pattern. If the speaker's rate of speech changed noticeably, I asked him to slow down or speed up, accordingly.

Some speakers, especially of Àkùrẹ́, did not produce the expected harmony patterns. One speaker's data were not included because the harmony did not show any consistent pattern, especially when the target words started with a high vowel. While this was a problem for non-words, it was not a problem for real words. These problems may have been due to the mixing of dialects, since the speakers all knew SY, in addition to Àkùrẹ́. The orthography may have contributed to the problem, since harmony is not marked on high vowels. In any case, because a sufficient number of speakers consistently harmonized in the expected pattern, I was not concerned with the speech of those who did not. Even allowing for the possibility that some Àkùrẹ́ speakers do not harmonize high vowels, this would not impact the results of this study, but instead might reflect the influence of Standard Yorùbá phonology on some Àkùrẹ́ speakers.

3.5 Segmentation and labeling

3.5.1 Digitization

The recordings were digitized at a sampling rate of 11,025 kHz and low pass filtered to a Sun SparcStation at the Cornell Phonetics Lab. Each token with its carrier phrase was stored in a separate sound file for analysis. I disregarded recorded tokens with obvious problems, such as the speaker saying the wrong form, stuttering, or pausing. Of the remaining tokens for each form, I measured the first three or as many as were left if fewer than three remained.

3.5.2 Segmentation

The XLABEL utility associated with XWAVES allows the user to insert labels at specified times while viewing the waveform and spectrogram of a sound file, thus inserting the label name and corresponding time value into a text file for measurements and for future reference. The onset and offset for both vowels in the target word were labeled +V1, −V1, +V2, and −V2, respectively. In addition, the onset and offset of the preceding vowel (+V0, −V0) were labeled, as well as the start of the intervocalic consonant burst (+burstC2) for the [k] of VkV words. An illustrative labeled utterance is provided in Figure 3.4.

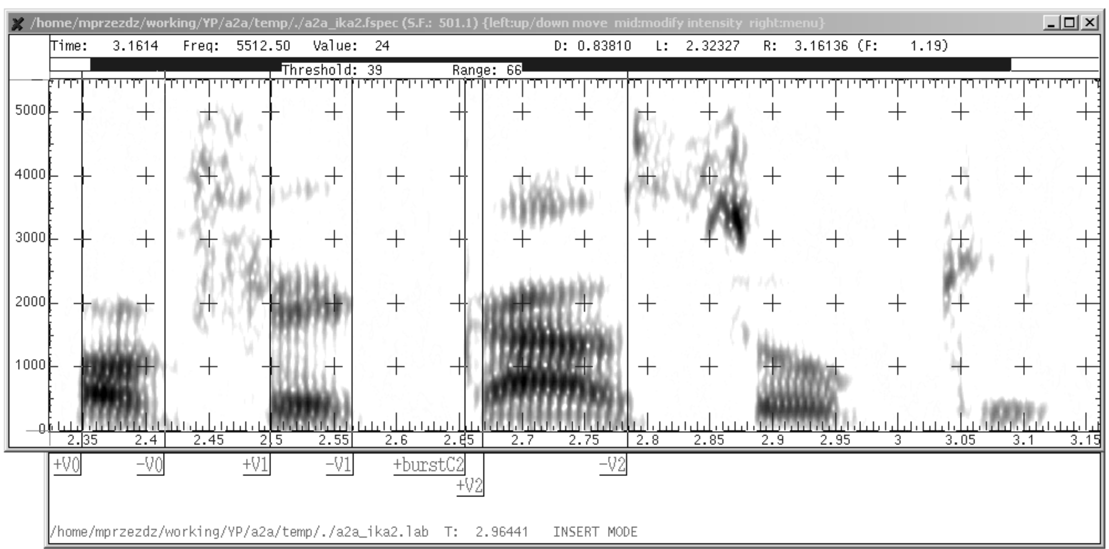


Figure 3.4. Spectrogram of Àkùrẹ́ speaker Ak2 *ọ́ f'ika s'okè*.

The vowel endpoints were labeled in order to measure the vowel durations and also to serve as the basis for determining points for measuring vowel characteristics. The onsets and offsets were usually marked at the start and end of F2 energy going into the consonant as observed in the spectrogram. In the case of the high vowels, which often have very short duration with low amplitude and with little formant structure evident on the spectrogram, I inspected the waveform and also used cues (such as frication) from adjacent consonants. For the vowel onset, neither the burst nor the aspiration of the consonant were labeled as the start of the vowel. There are two reasons for this. First, I am interested in the formant structure of the vowels, which does not start at the burst, and is very difficult to measure during the aspiration. Second, in many tokens, the burst is either not present or is difficult to discern. The use of the F2 onset allowed for consistent and objective measurements across tokens.

Once the endpoints of the vowel were determined, the vowel was automatically divided into ten equal sections with a label at 10% into the vowel, 20% into the vowel, and so forth until 90% into the vowels, the labels being inserted into the existing label file for each token. The points were labeled V1+10%, V1+20%, and so forth up to V1+90%. The same was done for V2. An example of a resulting label file is shown in (6); each point corresponds to a point in the spectrogram in Figure 3.5.

(6) partial contents of label file for speaker Ak2, token ika2

```
signal a2a_ika2
2.3483   +V0
2.4142   −V0
2.4989   +V1
2.5055   V1+10%
2.5121   V1+20%
2.5188   V1+30%
2.5254   V1+40%
2.5321   V1+50%
2.5387   V1+60%
2.5453   V1+70%
2.5520   V1+80%
2.5586   V1+90%
2.5653   −V1
2.6542   +burstC2
2.6691   +V2
2.7832   −V2
```

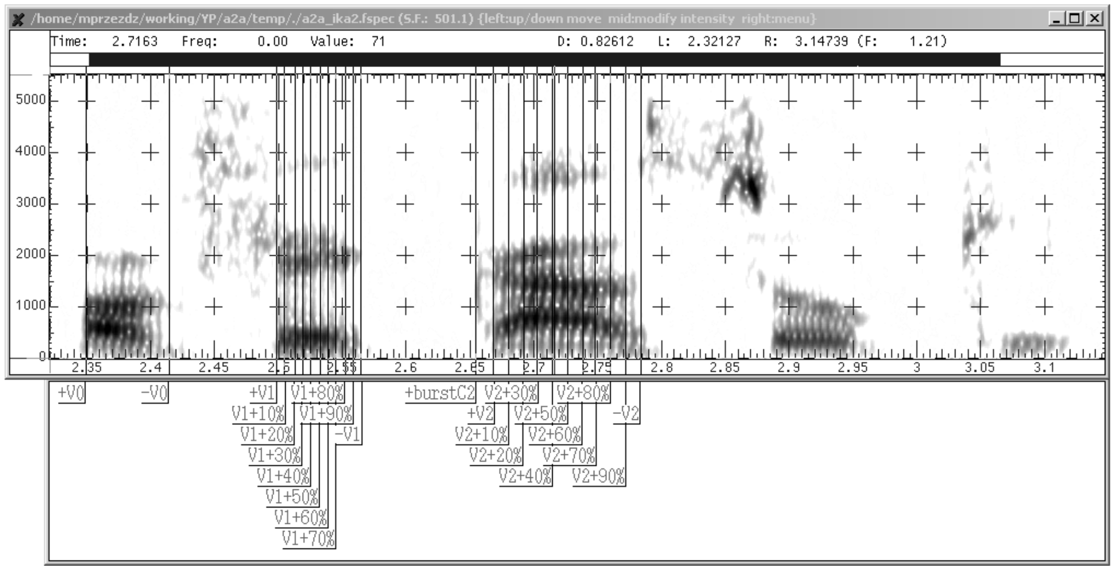


Figure 3.5. Spectrogram after running labeling script of sentence *ọ́ f'ika s'okè* for Àkùrẹ́ speaker Ak2.

3.6 Acoustic measures

I examined a number of acoustic measures to determine which ones best correlate with the two phonological sets we assume to be [+ATR] and [−ATR], or more specifically, how do [i] and [u] differ from [ị] and [ụ] in Àkùrẹ́, and how do [e] and [o] differ from [ẹ] and [ọ] in the three dialects. One of the purposes of this endeavor is to determine if the acoustic characteristics distinguishing the two sets of vowels are consistent with those found in other [ATR] languages. In this section, the results of this part of the study are presented, with an examination of data for one speaker for each dialect. Once an acoustic correlate is found, the second step is to determine if phonetic differences between identical phones in different contexts differ in the same acoustic correlate as is found with phones which differ only in [ATR]. This step is examined in chapter 4.

Hess (1992) found, in one speaker of Akan, that the first formant (F1) and first formant bandwidth (BW1) were the best acoustic correlates of the phonological feature [ATR]—that is, the [±ATR] vowels pairs {i/ị, u/ụ, e/ẹ, o/ọ} could be best distinguished acoustically by these two factors, with [+ATR] vowels having lower F1s and lower BW1s than their [−ATR] counterparts. The other factors she considered were vowel duration and vowel spectral balance. Fulop, Kari, and Ladefoged (1998) found that F1 was the best acoustic correlate for four out of five vowel pairs in Degema (Edoid, Niger-Congo). The low vowel pair could not be distinguished by F1. Following Hess, I look at a number of measures to determine which measurements were correlated with ATR differences in the Yorùbá varieties studied here. The following measurements were taken for one speaker of each dialect:

(7) Acoustic Measurements

- Vowel Duration
- 1st Formant (F1)
- 2nd Formant (F2)
- 1st Formant Bandwidth (BW1)
- Fundamental Frequency (F0)
- Spectral Measures:
  - Amplitude of the fundamental frequency (F0) minus amplitude of the 2nd harmonic (H2)
  - Amplitude of the harmonic in F1 minus amplitude of F0

These measurements were chosen because they could plausibly be expected to differ from one vowel set to the other. Some of the measures were *a priori* more likely than others to be involved in the relevant contrast. In particular, F1 and BW1 were the primary candidates, due to their correlation to ATR from other languages, such as Akan and Igbo (Hess 1992; Lindau 1978; Ladefoged 1964). However, the spectral measures were also under consideration since the vowel differences between [+ATR] and [−ATR] could conceivably be related to voice quality differences such as breathiness and creakiness, which have been found to be correlated with spectral tilt, an indication of the state of the glottis. Two quantitative acoustic measures typically used to determine phonation type are listed above. They will be described in detail below.

The data were imported into a spreadsheet program (Microsoft Excel) and into a statistical analysis program (SPSS) for analysis. In the following sections, results are presented for each measure. Similar to Hess's (1992) findings for Akan, F1 was found to be the best acoustic correlate of the feature [ATR] in all three varieties of Yorùbá. However, unlike Hess's study of Akan, BW1 was not a good correlate for the feature [ATR] in the three varieties of Yorùbá studied here. The other measures taken here were also not found to be correlated to the ATR sets. In what follows, I describe how each of these measurements was taken, and determine how well they correlate with the phonological [±ATR] distinctions.

3.6.1 Vowel duration

The vowel duration was measured as the difference between the vowel offset (labels: −V1 or −V2) and onset (+V1 or +V2), shown in Figure 3.6. This measure was easily calculated from the label files with the use of csh and awk scripts.

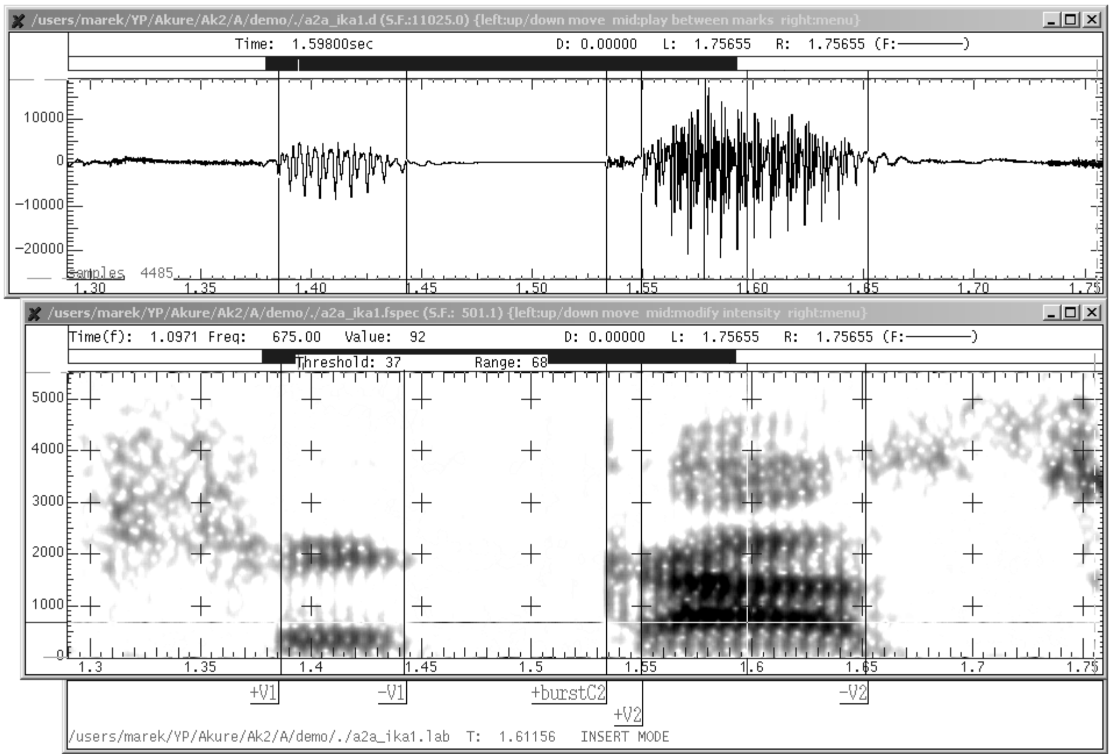


Figure 3.6. Example of vowel duration measure for Àkùrẹ́ speaker Ak2, token ika1, for V1 and V2.

The average duration of each V1 vowel for Àkùrẹ́ speaker Ak2 is shown in Figure 3.7. For all similar figures, the diagonally striped bars indicate [+ATR] vowels; the dotted bars indicate [−ATR] vowels. As expected, lower vowels are longer than higher ones. For front vowels, [+ATR] vowels are shorter then the [−ATR] ones. For the back vowels, there is virtually no difference between [+ATR] and [−ATR] vowels. I performed one-way ANOVAs with vowel duration as the dependent variable and ATR of V1 the factor, which showed the differences to be significant for the front vowel pairs: for {i/ị}, [$F(1, 41) = 10$, $p < .01$]; for {e/ẹ}, [$F(1, 38) = 10$, $p < .01$]. However for the back vowels, the differences were not significant: for {o/ọ}, [$F(1, 41) = .004$, $p > .95$]; and for {u/ụ}, [$F(1, 41) = 1.5$, $p > .23$].

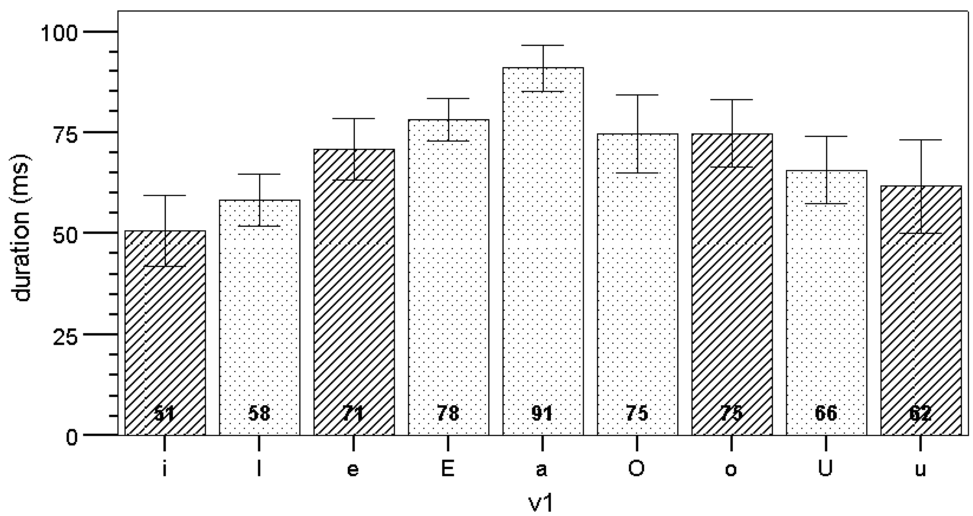


Figure 3.7. Mean duration for V1 for Àkùrẹ́ speaker Ak2.[5] The error bars mark one standard deviation. The [+ATR] vowels are [i, e, o, u]; the [−ATR] vowels are [I, E, O, U, a], corresponding to [ị, ẹ, ọ, ụ, a].

[5] Due to apparent software limitations, phones normally written with an underdot are displayed as capital letters in all figures. For example, [−ATR] [ụ] is shown as U.

The average duration of each V1 vowel for Mọ̀bà speaker Mb10 is shown in Figure 3.8. The [+ATR] mid-vowels have a slightly shorter duration than their [−ATR] counterparts. One-way ANOVAs with vowel duration as the dependent variable and ATR of V1 the factor showed the differences not to be significant for either mid vowel pair: for {e/ẹ}, [F(1, 53) = 3.3, p = .07]; and for {o/ọ}, [F(1, 53) = .50, p = .48].

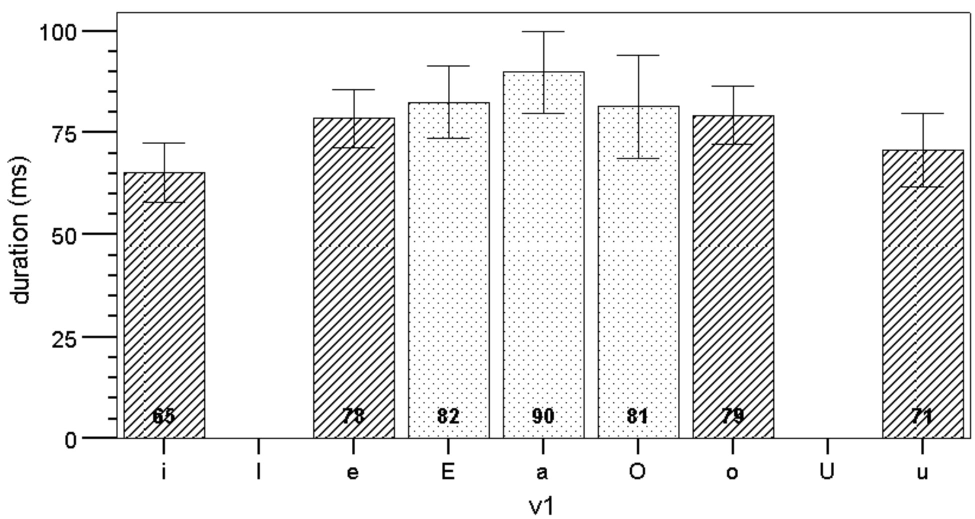


Figure 3.8. Mean duration for V1 for Mọ̀bà speaker Mb10. The error bars mark one standard deviation. The [+ATR] vowels are [i, e, o, u]; the [−ATR] vowels are [E, O, a], corresponding to [ẹ, ọ, a].

The average duration of each V1 vowel for SY speaker SY1 is shown in Figure 3.9. The pattern is similar to the pattern for Mb10 above: the [+ATR] mid-vowels have a shorter duration than their [−ATR] counterparts, but the differences are not significant, as shown by one-way ANOVAs with vowel duration as the dependent variable and ATR of V1 the factor: for {e/ẹ}, [F(1, 53) = 4.8, p = .031]; and for {o/ọ}, [F(1, 52) = 5.0, p = .029].

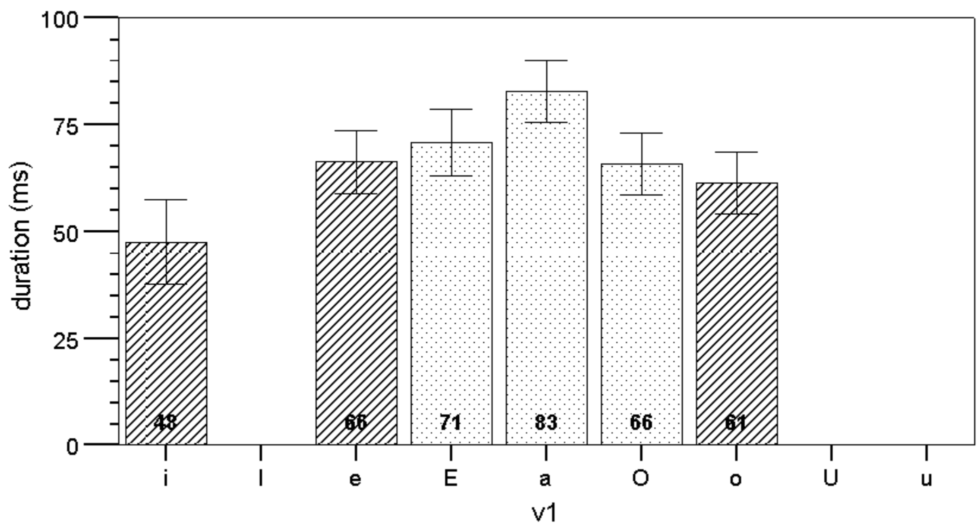


Figure 3.9. Mean duration for V1 for SY speaker SY1. The error bars mark one standard deviation. The [+ATR] vowels are [i, e, o]; the [−ATR] vowels are [E, O, a], corresponding to [ẹ, ọ, a].

While there is a tendency for [+ATR] vowels to be shorter than [−ATR] for these three speakers, the correlation is rather weak. It is clear that vowel duration does not sufficiently distinguish [+ATR] and [−ATR] vowels for these speakers.

3.6.2 Fundamental frequency

In this section, I examine the fundamental frequency (F0) of the vowels. The F0 is obtained by using the ESPS program get_f0 which uses a cross correlation algorithm. A script was developed to automatically calculate the f0 at each of the specified labels. Figure 3.10 shows an example of an f0 trace created in xwaves using get_f0. The f0 is reported here only for the midpoint of the vowels.

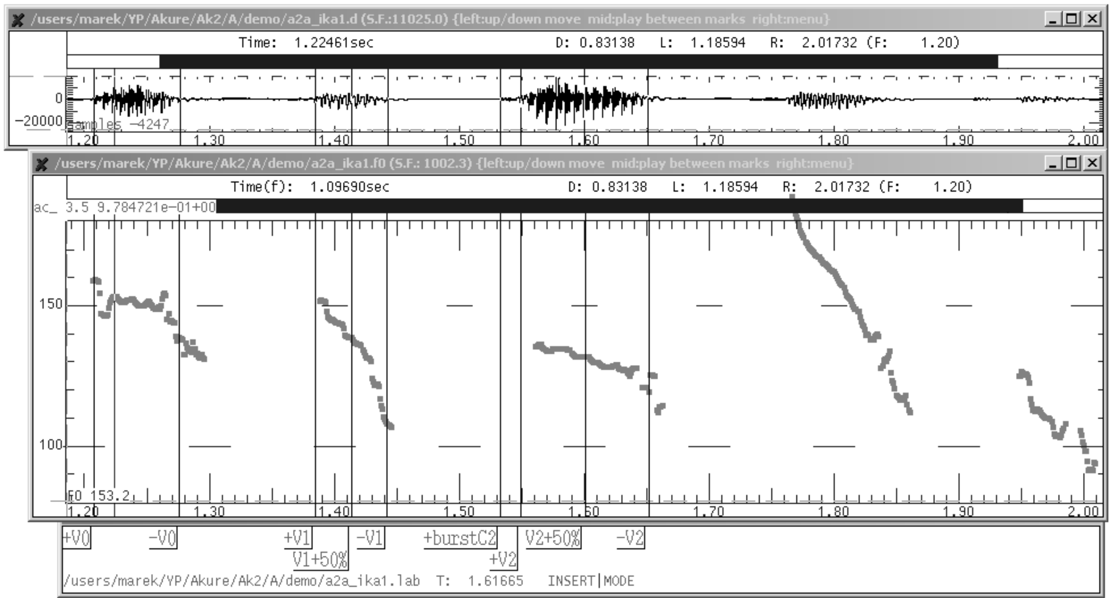


Figure 3.10. Waveform, F0 trace, and labels for Àkùrẹ́ speaker Ak2 saying *ọ́ f'ika s'okè*. For this study, the f0 is measured at labels marked "V1+50%" and "V2+50%".

For Àkùrẹ́ speaker Ak2, we observe, in Figure 3.11, a pattern where the lower vowels have a lower F0, as for [a], [ẹ], and [ọ]. For the mid vowels, the [−ATR] vowels [ẹ], [ọ], have a lower F0 than the [+ATR] vowels [e], [o]. However, for the [+ATR] and [−ATR] pairs, the differences are only significant for [o] versus [ọ], as shown by one-way ANOVAs with f0 as the dependent variable and ATR of V1 the factor: {e/ẹ}, [$F(1, 37) = 3.4$, $p = .073$]; and for {o/ọ}, [$F(1, 41) = 18.6$, $p < .01$]. For the high vowels, the [+ATR] vowels do not have higher F0s than the [−ATR] vowels. For the front pair [i] and [ị], the F0s are virtually equal, while for the back pair, it is the [−ATR] vowel [ụ] which has the higher F0. The results for the one-way ANOVA show that neither difference is significant: for {i/ị}, [$F(1, 41) = .03$, $p = .87$]; and for {u/ụ}, [$F(1, 41) = 1.99$, $p = .16$].

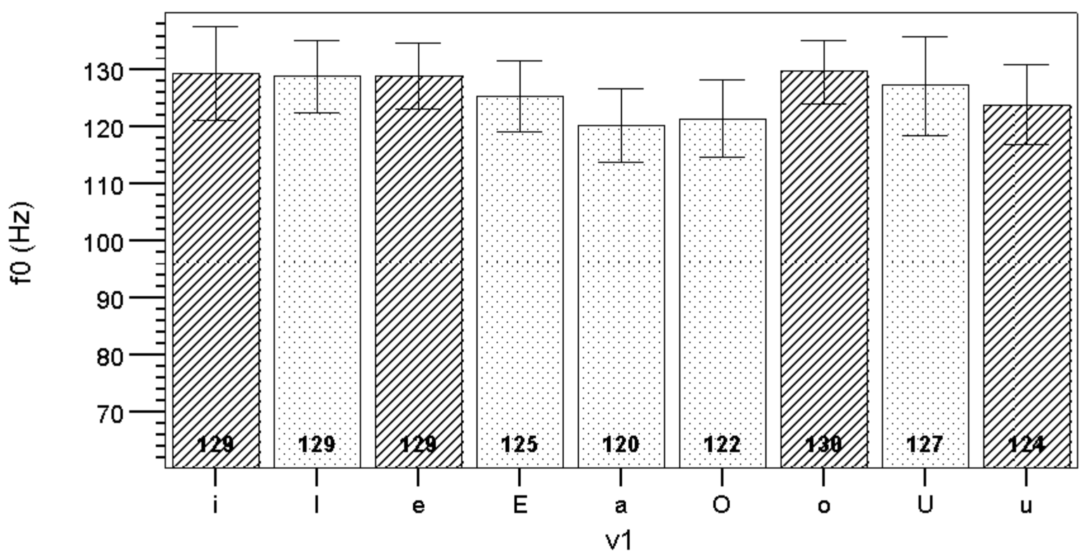


Figure 3.11. F0 for V1 for Àkùrẹ́ speaker Ak2. The error bars mark one standard deviation. The [+ATR] vowels are [i, e, o, u]; the [−ATR] vowels are [I, E, O, U, a], corresponding to [ị, ẹ, ọ, ụ, a].

For Mọ̀bà speaker Mb10, shown in Figure 3.12, the lower vowels [ẹ], [ọ], and [a] do not have lower F0s than do the high vowels. There is also no significant difference between the [+ATR] vowels and the [−ATR] vowels, as shown by one-way ANOVAs with F0 as the dependent variable and ATR of V1 the factor: {e/ẹ}, [$F(1, 53) = 1.46$, $p = .232$]; and for {o/ọ}, [$F(1, 53) = 4.63$, $p = .036$].

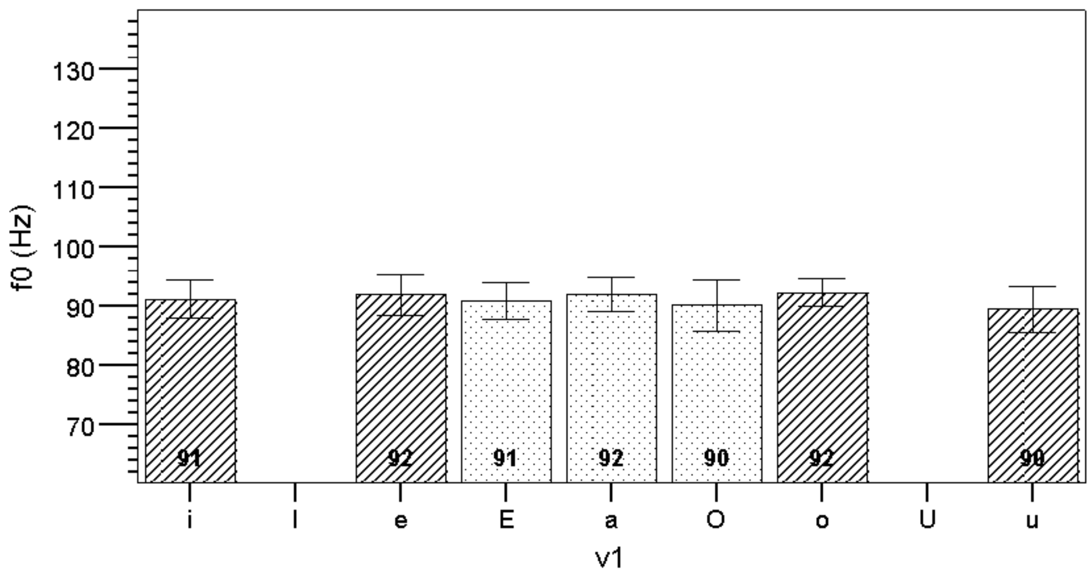


Figure 3.12. F0 for V1 for Mọ̀bà speaker Mb10. The error bars mark one standard deviation. The [+ATR] vowels are [i, e, o, u]; the [−ATR] vowels are [E, O, a], corresponding to [ẹ, ọ, a].

For SY speaker SY1, shown in Figure 3.13, the lower vowels [ẹ], [ọ], and [a] have a slightly lower F0 than the others, though not significantly. The [−ATR] mid vowels have a slightly lower average F0 than the [+ATR] mid vowels, but again the difference is not significant as shown by one-way ANOVAs with F0 as the dependent variable and ATR of V1 the factor: {e/ẹ}, [$F(1, 53) = 2.80$, $p = .101$]; and for {o/ọ}, [$F(1, 52) = 5.19$, $p = .027$].

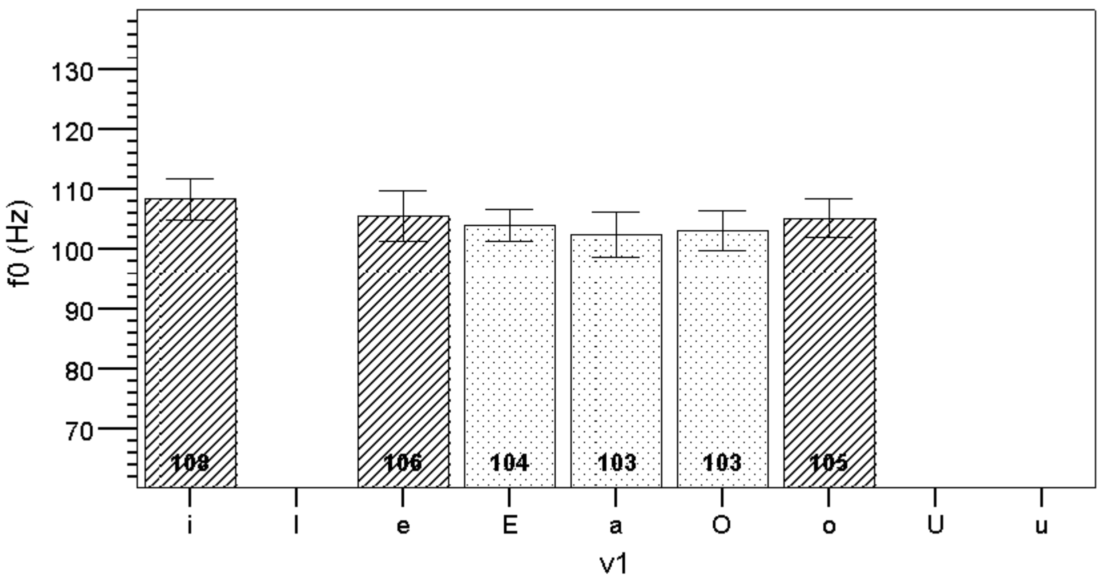


Figure 3.13. F0 for V1 for SY speaker SY1. The error bars mark one standard deviation. The [+ATR] vowels are [i, e, o]; the [−ATR] vowels are [E, O, a], corresponding to [ẹ, ọ, a].

It is clear that F0 is not a reliable measure for distinguishing between the two sets of vowels in any of the three subjects examined. The results for the Àkùrẹ́ speaker and the SY speaker show that in general the higher vowels have a higher fundamental frequency, consistent with Hombert's (1977) examination of f0 in Yorùbá vowels. This pattern, known as the intrinsic f0 of vowels, has been observed in many or perhaps all languages in which it has been studied (Peterson and Barney 1952, Lehiste and Peterson 1961, and see Whalen and Levitt 1995 for an overview). Although the Mọ̀bà speaker does not show the same effect, this may be due to the low f0 for that speaker, since it has been reported that at a lower pitch range the intrinsic pitch may disappear (Ladd and Silverman 1984, Whalen and Levitt 1995).

3.6.3 Formants and formant bandwidth

In this section, we examine the F1 measurements; a more detailed look of F1 in all three dialects is found in chapter 4. The formants and bandwidths were measured using the xspectrum utility associated with ESPS xwaves, using a smoothed LPC analysis by the autocorrelation method with a Hamming window of 20 ms and an order of 14. The spectrum program produces a graphical view of the spectrum at a given point in time. In addition, for each label it computes a list of poles consisting of frequency and the corresponding bandwidth (BW) in hertz. Poles with a bandwidth greater than 600 Hz for F1 and F2 were discarded, as these were clearly not formants. In addition, data points that did not match the value seen on the spectrogram were removed, as well as ones that were not plausible values for the vowel. An example of the output of the spectrum program is shown in Table 3.4, with the corresponding spectrogram and spectrum shown in Figure 3.14; in this case, three pairs of values were removed, as the large bandwidth clearly indicated that these were not formants.

Table 3.4. Output of spectrum program for midpoint of [i̯] in *ika* for speaker Ak2. The poles that are not formants are crossed out.

FREQUENCIES	BANDWIDTHS
406.364	48.826
~~747.722~~	~~1095.158~~
1893.287	158.154
2309.103	195.662
~~3486.195~~	~~1973.380~~
3793.000	135.413
~~4648.315~~	~~561.446~~

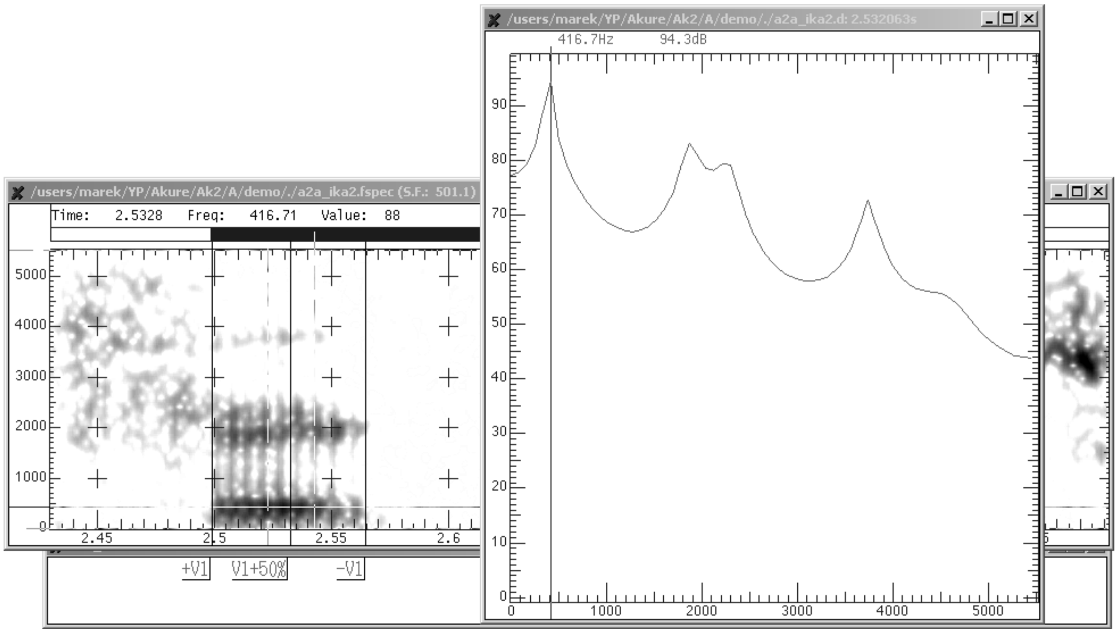


Figure 3.14. Spectrogram and LPC spectrum for speaker Ak2, token ika2, middle of /i/ at label V1+50% (20 ms Hamming window, Autocorrelation, order 14).

I used the LPC spectrum value instead of estimating the value from a spectrogram for two reasons. First, the F1 value for high vowels was difficult to ascertain from the spectrogram with any consistency and objectivity. Second, the spectrogram method was considerably slower than the spectrum method. While the spectrogram method involves hand inputting the values, the spectrum method was automated using various scripting languages (c-shell, xwaves, awk, perl). LPC analysis has been used by Disner 1983, Manuel 1990, Hess 1992, Bradlow 1993, and Zsiga 1993, 1997. The choice of a 20-25 ms. data window is a common one (Bradlow 1993, Manuel 1990, Zsiga 1993). Choi and Keating (1991) used a smaller sized window, 5 ms., because they were looking at languages with very small and short lasting intervocalic coarticulatory effects. In this way, they were able to measure the vowel formants close to the intervocalic consonants. With a smaller window, one has greater precision in time, but lower precision in terms of frequency and bandwidth. Thus, more local discrepancies are seen in the data, especially for a speaker with a low F0. As I was not primarily concerned with coarticulation at the edges of the vowels, a larger window was acceptable.

3.6.4 F2 results

In this section, the F2 results are examined, first for Àkùrẹ́ speaker Ak2, shown in Figure 3.15. For the front vowel pairs, the [+ATR] vowel exhibits a higher F2 than its [−ATR] counterpart. On the other hand, for {o/ọ}, the [−ATR] vowel [ọ] has a higher F2 than [+ATR] [o]. The vowels {u/ụ} have nearly identical F2 averages, with the [−ATR] [ụ] slightly higher than [u] . One-way ANOVAs with F2 as the dependent variable and ATR of V1 the factor, showed the differences to be highly significant for three vowel pairs, but not for {u/ụ}: for {i/ị}, [$F(1, 41) = 14.5$, $p < .001$]; for {e/ẹ}, [$F(1, 38) = 119$, $p < .001$]; and for {o/ọ}, [$F(1, 41) = 110$, $p < .001$]; but not for {u/ụ}, [$F(1, 41) = 2.7$, $p = .11$].

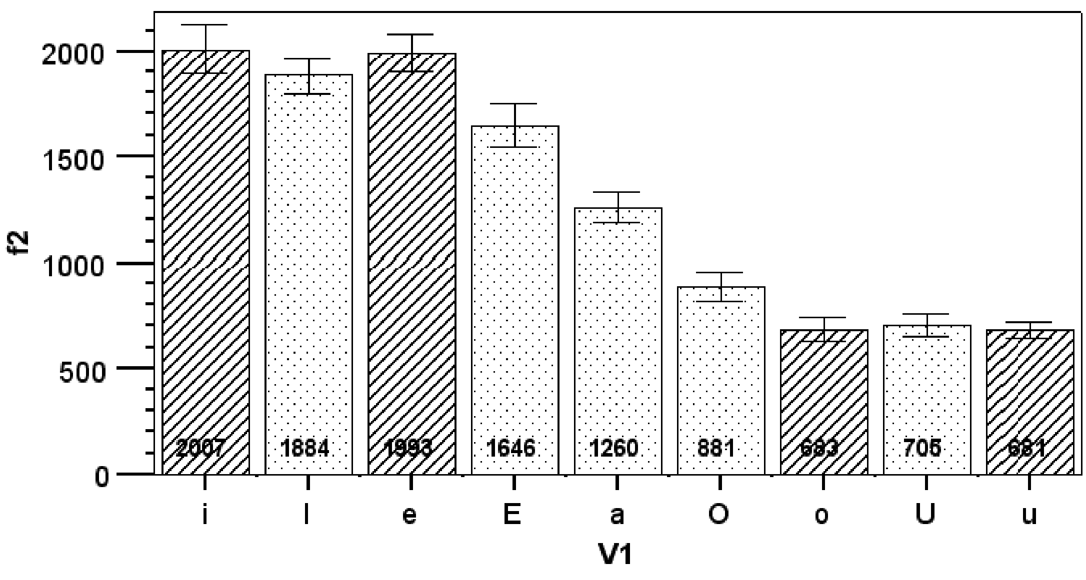


Figure 3.15. Mean F2 for V1 midpoint for Àkùrẹ́ speaker Ak2. The error bars mark one standard deviation. The [+ATR] vowels are [i, e, o, u]; the [−ATR] vowels are [I, E, O, U, a], corresponding to [ị, ẹ, ọ, ụ, a].

The F2 results for Mọ̀bà speaker Mb10 are shown in Figure 3.16. The mid vowels show the same pattern of ATR differences as the mid vowels in Àkùrẹ́. The [+ATR] [e] has a significantly higher F2 than [−ATR] [ẹ]; and the [+ATR] [o] has a significantly lower F2 than [−ATR] [ọ]. One-way ANOVAs with F2 as the dependent variable and ATR of V1 the factor showed the differences to be highly significant for both vowel pairs: for {e/ẹ}, [$F(1, 53) = 75$, $p < .001$]; and for {o/ọ}, [$F(1, 53) = 192$, $p < .001$].

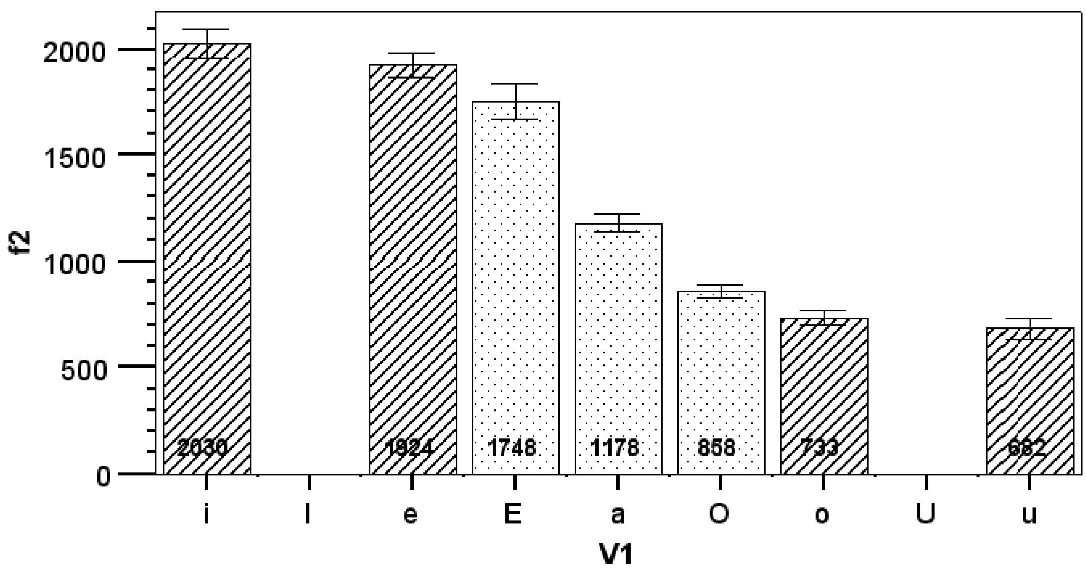


Figure 3.16. Mean F2 for V1 midpoint for Mọ̀bà speaker Mb10. The error bars mark one standard deviation. The [+ATR] vowels are [i, e, o, u]; the [−ATR] vowels are [E, O, a], corresponding to [ẹ, ọ, a].

The same pattern is seen for the F2 results of SY speaker SY1 in Figure 3.17. The [+ATR] /e/ has a significantly higher F2 than [−ATR] /ẹ/; and the [+ATR] /o/ has a significantly lower F2 than [−ATR] /ọ/. One-way ANOVAs with F2 as the dependent variable and ATR of V1 the factor showed the differences to be highly significant for both vowel pairs: for {e/ẹ}, [$F(1, 52) = 25$, $p < .001$]; and for {o/ọ}, [$F(1, 52) = 68$, $p < .001$].

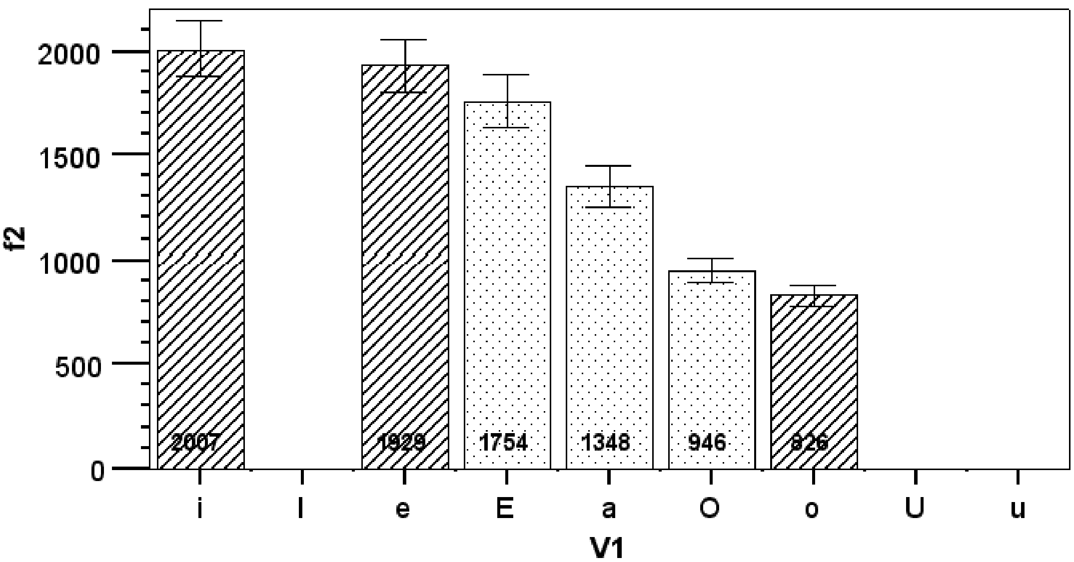


Figure 3.17. Mean F2 for V1 midpoint for SY speaker SY1. The error bars mark one standard deviation. The [+ATR] vowels are [i, e, o]; the [−ATR] vowels are [E, O, a], corresponding to [ẹ, ọ, a].

For the three speakers, [+ATR] front vowels have a higher F2 than their [−ATR] counterparts. The mid back vowels {o/ọ} show the opposite effect. The one pair of high back vowels, in Àkùrẹ́, is not statistically different. The results show that F2 is not a reliable correlate of ATR.

3.6.5 F1 results

F1 was found to be a reliable correlate of ATR in Àkùrẹ́. For speaker Ak2, differences in mean F1 between [+ATR] and [−ATR] vowel pairs were large, as seen in Figure 3.18. For the four vowel pairs, the [+ATR] vowel exhibits a significantly lower F1 than its [−ATR] counterpart. For example, [+ATR] [i] has an average F1 of 271 Hz, while for [−ATR] [ị] the average F1 is 408 Hz. The same pattern is found for [e]/[ẹ], [o]/[ọ], and [u]/[ụ]. The [−ATR] low vowel [a] has no [+ATR] counterpart. In order to determine that these formant values differ for each vowel in the pair, I performed one-way ANOVAs with F1 as the dependent variable and ATR of V1 the factor, which showed the differences to be highly significant for all four vowel pairs: for {i/ị}, [$F(1, 41) = 546$, $p < .001$]; for {e/ẹ}, [$F(1, 38) = 555$, $p < .001$]; for {o/ọ}, [$F(1, 41) = 787$, $p < .001$]; and for {u/ụ}, [$F(1, 41) = 106$, $p < .001$].

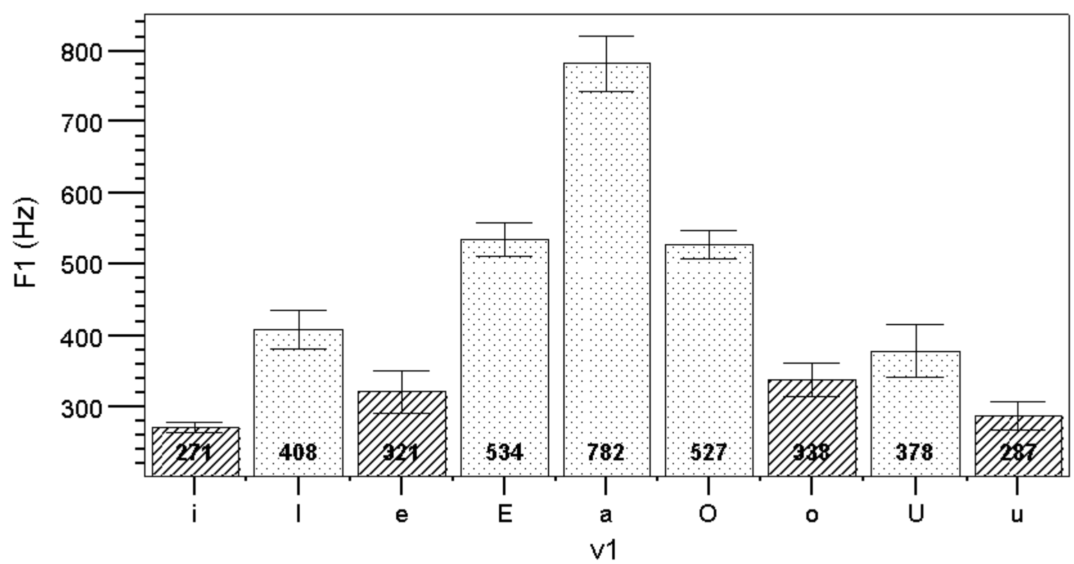


Figure 3.18. Mean F1 for V1 midpoint for Àkùrẹ́ speaker Ak2. The error bars mark one standard deviation. The [+ATR] vowels are [i, e, o, u]; the [−ATR] vowels are [I, E, O, U, a], corresponding to [ị, ẹ, ọ, ụ, a].

Data for the other three Àkùrẹ́ speakers show similar results (see chapter 4, §4.2). These results are consistent with Hess (1992) in that the ATR distinction is strongly correlated with F1. For Mọ̀bà and SY, where the high vowels do not exhibit ±ATR allophonic variants, we can only look at the mid vowels for ±ATR differences. As expected, for Mọ̀bà speaker Mb10, the mid vowels show the same type of ATR differences as seen in Àkùrẹ́. That is, the [+ATR] vowels have a lower F1 than their [−ATR] counterparts: 369 Hz for [e], 508 Hz for [ẹ]; 363 Hz for [o], 491 Hz for [ọ]. One-way ANOVAs with F1 as the dependent variable and ATR of V1 the factor showed the differences to be highly significant for both vowel pairs: for {e/ẹ}, [$F(1, 53) = 270$, $p < .001$]; and for {o/ọ}, [$F(1, 53) = 404$, $p < .001$].

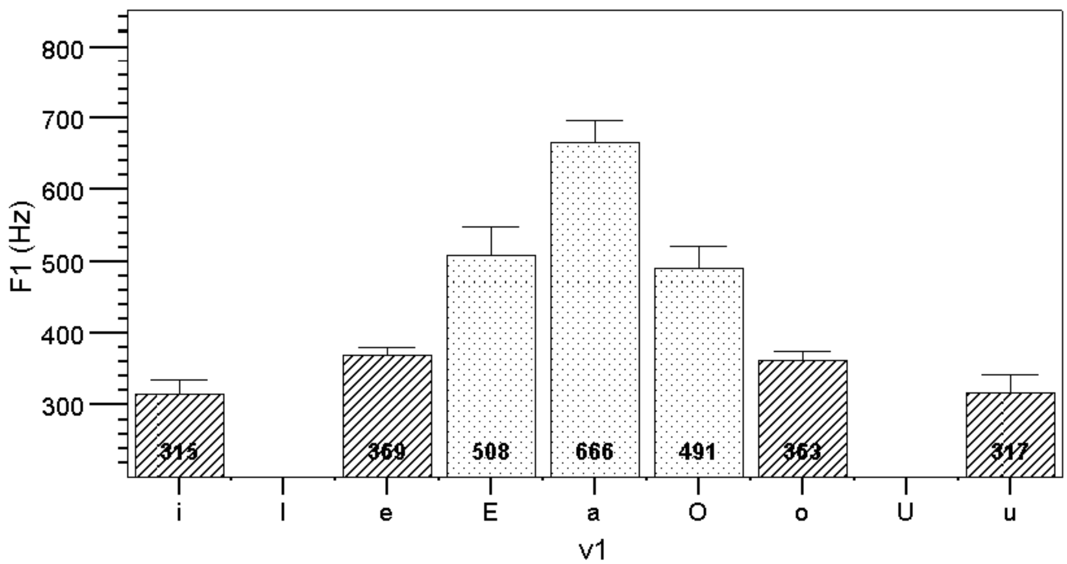


Figure 3.19. Mean F1 for V1 midpoint for Mọ̀bà speaker Mb10. The error bars mark one standard deviation. The [+ATR] vowels are [i, e, o, u]; the [−ATR] vowels are [E, O, a], corresponding to [ẹ, ọ, a].

The same picture is seen for SY speaker SY1 in Figure 3.20, that is, the [+ATR] vowels have a lower F1 than their [−ATR] counterparts: 371 Hz for [e], 483 Hz for [ẹ]; 386 Hz for [o], 516 Hz for [ọ]. One-way ANOVAs with F1 as the dependent variable and ATR of V1 the factor showed the differences to be highly significant for both vowel pairs: for {e/ẹ}, [$F(1, 53) = 54$, $p < .001$]; and for {o/ọ}, [$F(1, 52) = 45$, $p < .001$].

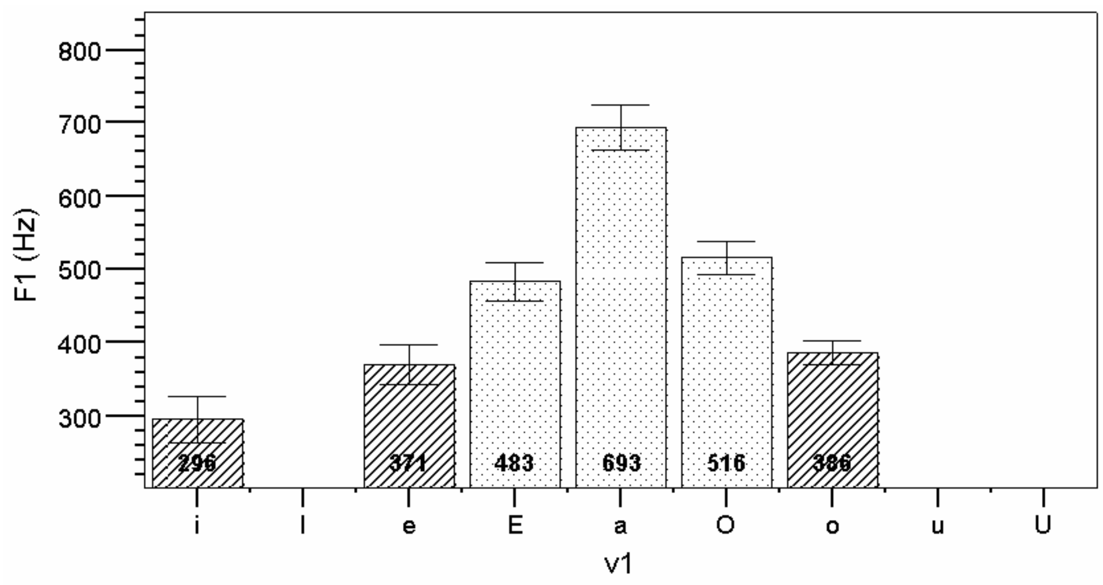


Figure 3.20. Mean F1 for V1 midpoint for SY speaker SY1. The error bars mark one standard deviation. The [+ATR] vowels are [i, e, o]; the [−ATR] vowels are [E, O, a], corresponding to [ẹ, ọ, a].

In summary, for one speaker of each of the three dialects, F1 is correlated with ATR differences in the mid vowels; and in Àkùrẹ́, the one dialect with a high vowel allophonic ATR split, F1 is also correlated with the ATR differences. This contrasts with the results of the F1 bandwidth analysis, which follows.

3.6.6 BW1 initial results

Figure 3.21 shows the mean F1 bandwidth (BW1) for each vowel at the midpoint of V1 for Àkùrẹ́ speaker Ak2. In three out of four vowel pairs, [+ATR] vowels have a lower average BW1 than the [−ATR] counterparts. Using a one-way ANOVA with BW1 as the dependent variable and ATR of V1 the factor, the BW1 difference for the vowel pair [i, ị] was found to be significant [$F(1, 41) = 42$, $p < .001$]. However, the difference between [+ATR] and [−ATR] vowels was not significant for {e/ẹ}, [$F(1, 38) = 4.3$, $p = .044$]; nor for {o/ọ}, [$F(1, 41) = 1.9$, $p = .174$]. Moreover, for the pair {u/ụ}, the BW1 difference was reversed, with [+ATR] [u] having a greater bandwidth than [−ATR] [ụ], though not significantly [F(1, 41) = 6.8, p < .013]. Thus, for speaker Ak2, BW1 is not a reliable correlate of the ATR distinction.

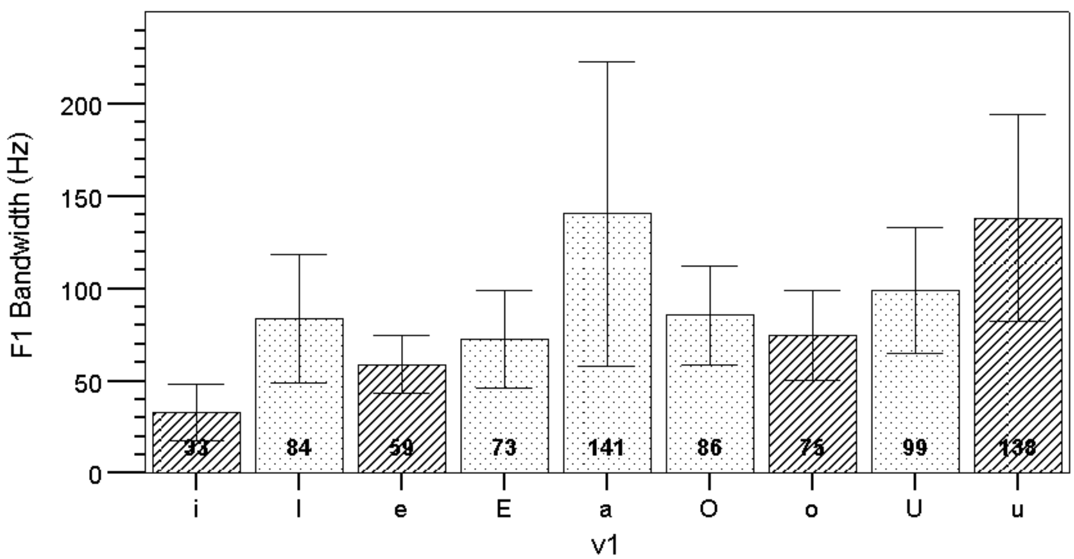


Figure 3.21. Mean F1 bandwidth for V1 midpoint for Àkùrẹ́ speaker Ak2. The error bars mark one standard deviation. The [+ATR] vowels are [i, e, o, u]; the [−ATR] vowels are [I, E, O, U, a], corresponding to [ị, ẹ, ọ, ụ, a].

For the three other Àkùrẹ́ speakers, we see that this is the case for them as well. Table 3.5 shows the mean BW1 values for each surface vowel for four Àkùrẹ́ speakers, as well as the p-values from a single-factor ANOVA which was performed for each [±ATR] vowel pair for each speaker with BW1 as the dependent variable and vowel as the factor. The shaded blocks indicate that for the specified vowel pair for that particular speaker, the effect of ATR on BW1 is significant and in the same direction that Hess found for Akan vowels. It is clear from this table, that BW1 is not correlated to ATR in these speakers of Àkùrẹ́.

Table 3.5. Mean F1 bandwidth values (in Hz) for [±ATR] vowel pairs for four speakers of Àkùrẹ́ Yorùbá. '†' indicates that the difference between [±ATR] vowels is in the opposite direction from that found in Hess (1992).

	Ak2		Ak3		Ak6		Ak7	
i	33	p < .01	54	p < .01	70	p > .65†	40	p < .01
ị	84		79		66		72	
e	59	p < .05	36	p > .26	63	p > .19	59	p < .05†
ẹ	73		42		70		44	
o	75	p > .17	67	p < .01†	82	p < .01†	98	p < .01†
ọ	86		44		67		62	
u	138	p < .05†	67	p < .01	120	p < .01	167	p < .57
ụ	99		104		174		196	

For the Mọ̀bà and the SY speaker, we can only look at the mid vowels. For Mọ̀bà speaker Mb10 in both pairs of mid vowels, BW1 is lower for [+ATR] vowels. One-way ANOVAs with F1 as the dependent variable and ATR of V1 the factor showed the differences to be highly significant both for {e/ẹ}, [$F(1, 53) = 46$, $p < .001$]; and for {o/ọ}, [$F(1, 53) = 135$, $p < .001$].

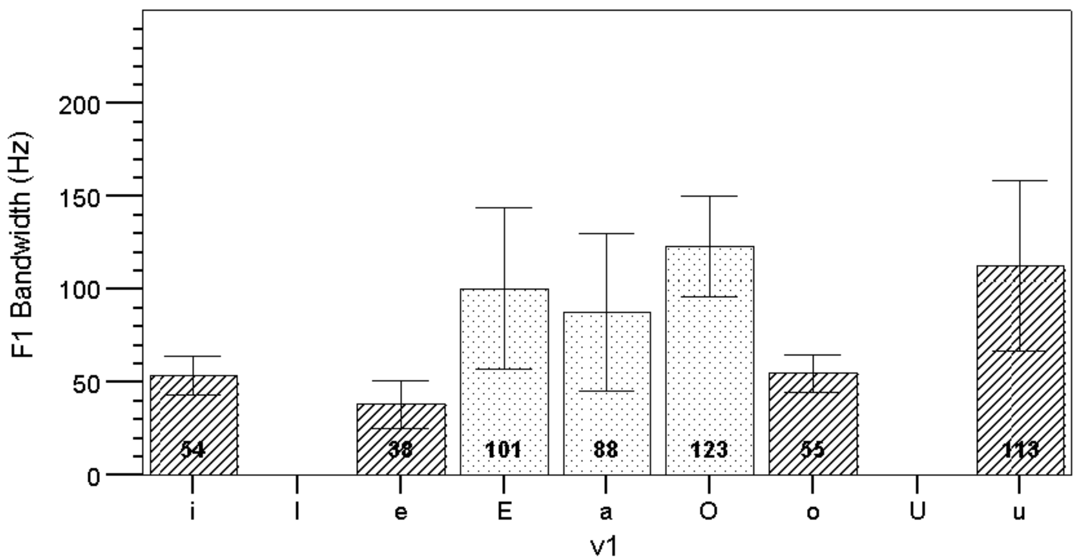


Figure 3.22. Mean F1 bandwidth for V1 midpoint for Mọ̀bà speaker Mb10. The error bars mark one standard deviation. The [+ATR] vowels are [i, e, o, u]; the [−ATR] vowels are [E, O, a], corresponding to [ẹ, ọ, a].

For SY speaker SY1, the mid vowel pairs show little difference between the [+ATR] and [−ATR] vowels. One-way ANOVAs with F1 as the dependent variable and ATR of V1 the factor showed the differences to be not significant for {e/ẹ}, [$F(1, 53) = .45$, $p = .51$]; and for {o/ọ}, [$F(1, 52) = 1.3$, $p = .26$].

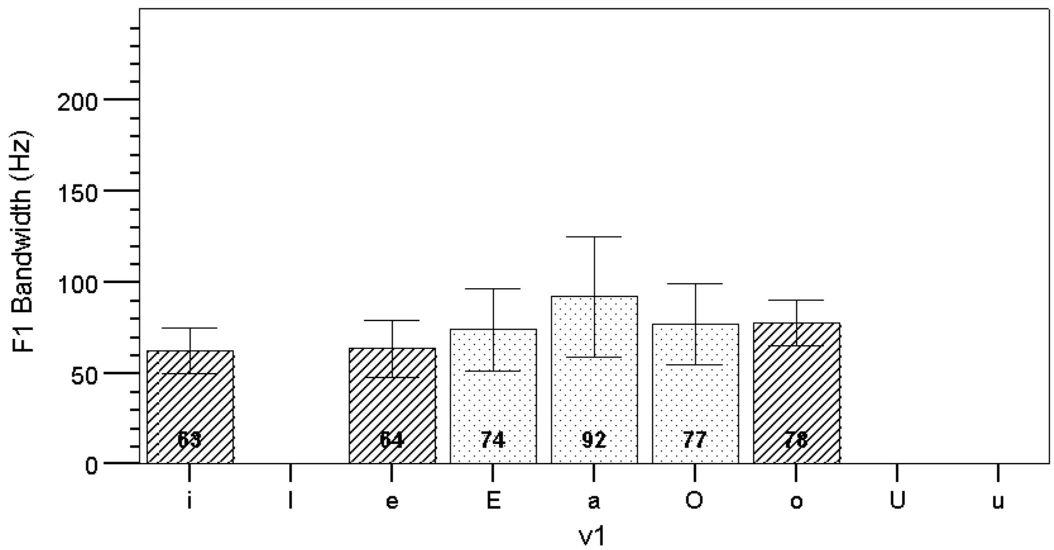


Figure 3.23. Mean F1 bandwidth for V1 midpoint for SY speaker SY1. . The error bars mark one standard deviation. The [+ATR] vowels are [i, e, o]; the [−ATR] vowels are [E, O, a], corresponding to [ẹ, ọ, a].

In summary, for one speaker of each of the three dialects, BW1 is not found to be a reliable correlate with ATR differences, in contrast to Hess's (1992) findings for Akan.

3.6.7 Spectral values

Halle and Stevens (1969) suggest that [±ATR] vowels can be distinguished by their phonation type, with [+ATR] vowels having a breathy quality resulting from the advancement of the tongue root. Hess (1992) found this was not the case in one speaker of Akan. In this section I explore the possibility that the [±ATR] sets in these Yorùbá speakers differ in phonation type. Of course, having found a correlation between F1 and [ATR] sets in the previous section does not preclude finding additional acoustic correlates, such as phonation type. Halle and Stevens suggest that the voice quality difference is secondary to the F1 lowering caused by the advanced tongue root. This may either be because the properties are physiologically related as Halle and Stevens propose, or because one set of differences may have developed to enhance the other perceptually.

Phonation type is a characteristic, mostly of vowels, originating in the larynx. To simplify, a vowel may be breathy, modal, or creaky depending on the size of the opening of the glottis and the tightness of the vocal folds (Klatt and Klatt, 1990). In a breathy vowel, the glottal opening is rather large, with the vocal folds vibrating loosely without completely closing, resulting in vocal cords which are not as tense producing less energy in the higher frequencies. A spectrum from a breathy vowel shows a steeper downward slope as the frequencies increase. In a creaky vowel, the glottal opening is small and the vocal folds are relatively tight. Creaky vowels have relatively high energy in the higher harmonics. A creaky vowel will still slope down, though not as much as a breathy one. For a modal vowel, the glottal opening is large, but unlike a breathy vowel, each pulse brings the vocal folds together. The slope for a modal vowel is between a breathy and a creaky vowel.

A spectrum derived from normal recordings does not directly indicate glottal vibrations, because it necessarily has the resonance of the supralaryngeal vocal tract superimposed on it. A number of methods have been used to quantitatively determine the relative phonation type of vowels from the spectrum. I use two methods here: the difference in amplitudes of the first and second harmonics (Bickley 1982); and the difference in amplitude between the first harmonic and the strongest harmonic in the first formant (Ladefoged 1983). For the first measure, the amplitude of the second harmonic (H2) is subtracted from the amplitude of the first harmonic, which is the fundamental frequency (F0). An example of this measurement is shown in Figure 3.24. In this case, and all the cases we examine, the difference will be negative, since the second harmonic peak (H2) is greater than the peak of the fundamental (F0). The larger the relative difference (that is, the smaller the magnitude, for negative values), the quicker the decline, indicating a breathier vowel.

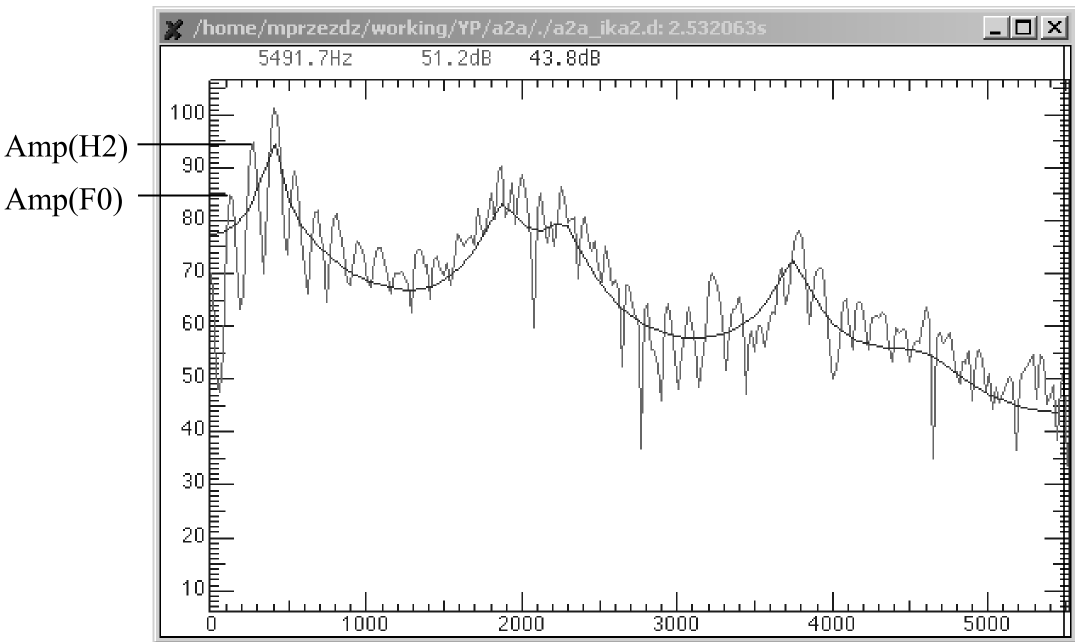


Figure 3.24. Amp(F0) − Amp(H2)

The second measure is similar, although instead of taking the amplitude of the second harmonic, we take the amplitude of the harmonic with the largest peak within the first formant of the vowel, shown in Figure 3.25. Subtracting the amplitude of F1 from the amplitude of F0 yields a negative value for all the subjects here. A relatively higher value indicates a breathier vowel.

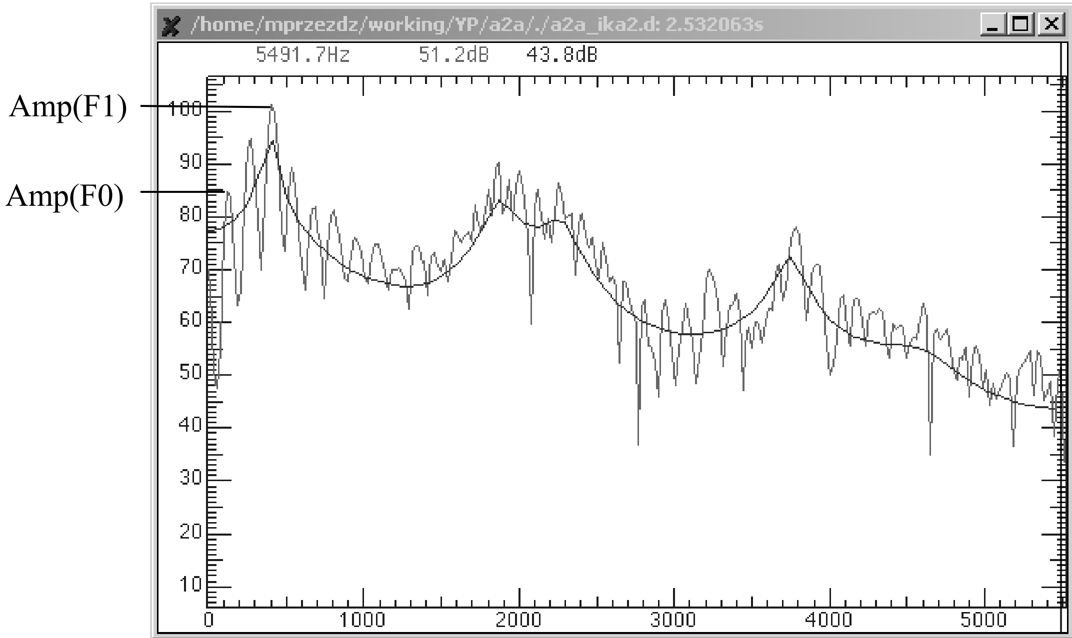


Figure 3.25. Amp(F0) − Amp(F1)

For both of these measures, it is not the case that a particular range can be said to be breathy while another range is creaky. In languages which make a phonological distinction between breathy and creaky vowels, there may be overlap between the breathy vowels of one speaker and the creaky vowels of another as in !Xóõ (Ladefoged and Antoñanzas–Barroso, 1984). For further discussion of phonation types, see Huffman 1987, Ladefoged, Maddieson, and Jackson 1988, Silverman 1997, Watkins 1997, Fischer-Jørgensen 1970, and Gordon and Ladefoged 2001, and references therein.

It is important to note that these measures may only be used for vowels of similar formant values (Ladefoged and Antoñanzas–Barroso, 1984), since a first formant value otherwise may raise the amplitude of F1 in the case of Amp(F0) − Amp(F1), or the amplitude of H2, in the case of Amp(F0) – Amp(H2). When comparing vowels that differ in F1 frequency values, H2 and F1 amplitudes will be boosted more for one vowel than the other, thus giving the false impression that those vowels are creakier than the others. It is not the actual values of the spectrum that indicate phonation type, but rather, these spectral values are estimates in determining of the state of the glottis, which is ultimately what this diagnostic seeks to measure. Therefore, when comparing values, it is important that the influence of the supralaryngeal vocal tract be relatively constant across vowels. Unfortunately, as we have seen in §3.6.5, the F1 values are distinct in the corresponding vowels in the two sets of vowels for each of the three speakers, rendering these measures unreliable. The results of these measures are nonetheless shown below for the three speakers. In order to more reliable measure phonation type, I provide details of a second Àkùrẹ́ speaker, Ak6, for which a comparison can be made between [ị] and [e], and [ụ] and [o], since for these vowel pairs, the F1 values are rather similar.

3.6.7.1. Results of amplitude F0 – amplitude H2

Let's first examine the difference of the first two peaks in the narrow band spectrum at the midpoint of V1. We compare the [+ATR] and [−ATR] sets of vowels pairwise—that is, [i] to [ị], [e] to [ẹ], and so forth—to determine if phonation type is a distinctive characteristic of these sets. Figure 3.26 shows the difference of the first two peaks for Àkùrẹ́ speaker Ak2. For each pair of vowels, the [+ATR] vowel has a lower value than the [−ATR] vowel. Note that the values are negative, indicating that the amplitude of the second harmonic is higher than the amplitude of the first. Using a one-way ANOVA with F0-H2 (dB) the dependent variable and ATR of V1 the factor, the differences were found to be significant for {i/ị}, [$F(1, 40) = 9.3$, $p < .01$]; for {e/ẹ}, [$F(1, 37) = 28.2$, $p < .01$]; for {o/ọ}, [$F(1, 40) = 31.1$, $p < .01$]; but not for {u/ụ}, [$F(1, 40) = 1.0$, $p = .326$].

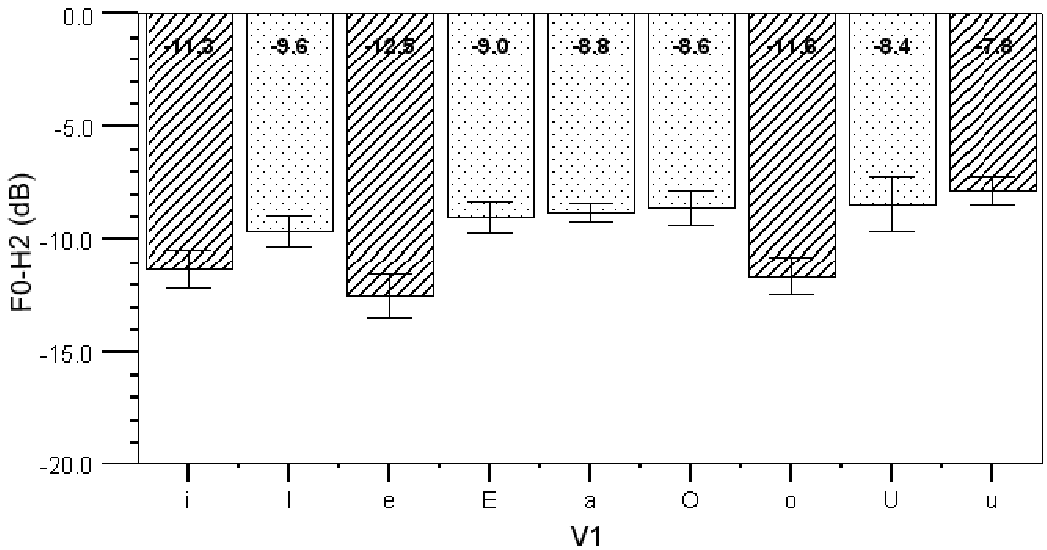


Figure 3.26. Mean F0 − H2 (dB) for V1 midpoint for Àkùrẹ́ speaker Ak2. The error bars mark one standard deviation. The [+ATR] vowels are [i, e, o, u]; the [−ATR] vowels are [I, E, O, U, a], corresponding to [ị, ẹ, ọ, ụ, a].

The direction of the differences might indicate that with vowel pairs, [+ATR] vowels are breathier than [−ATR] vowels, with the exception of [ụ] and [u]. However, as discussed above, the different F1 values influence the spectral peaks independently of phonation type. By examining the high front vowel pair, [i/ị], we can see the influence of F1 on H2, in particular. As listed in Table 3.6, the [+ATR] [i] vowel has an average F1 of 271 Hz, while for [−ATR] [ị], it is 408 Hz. With an approximate F0 of 128 Hz, the second harmonic (H2) occurs at about 256 Hz.

Table 3.6. Measurements for Àkùrẹ́ speaker Ak2

	F1 (Hz)	F0 (Hz)	Amp(F0)	Amp(H2)
i	271	128	82.7	94.1
ị	408	129	83.0	92.7
e	321	129	84.0	96.5
ẹ	534	125	81.6	90.7

Clearly, the F1 will boost the amplitude of H2 for [i], since the H2 frequency is close to F1. For [−ATR] [i̙], F1 is not as close to H2, so the boost will be smaller. We predict that the H2 of [i] will thus be greater than the H2 of [i̙] due to F1 alone, and this is consistent with the actual measurements. Therefore, the results of this measurement do not reliably show that [+ATR] vowels are more breathy than their [−ATR] counterparts. The same is true for the Mọ̀bà and SY speakers, discussion of which follows.

For the Mọ̀bà speaker Mb10, shown in Figure 3.27, we compare only the mid vowel pairs. As for the Àkùrẹ́ speaker, the [+ATR] value is lower than the [−ATR] value. One-way ANOVAs show that the difference is significant for both vowel pairs: for {e/ẹ}, [$F(1, 52) = 13.8$, $p < .01$], and for {o/ọ}, [$F(1, 52) = 39.4$, $p < .01$].

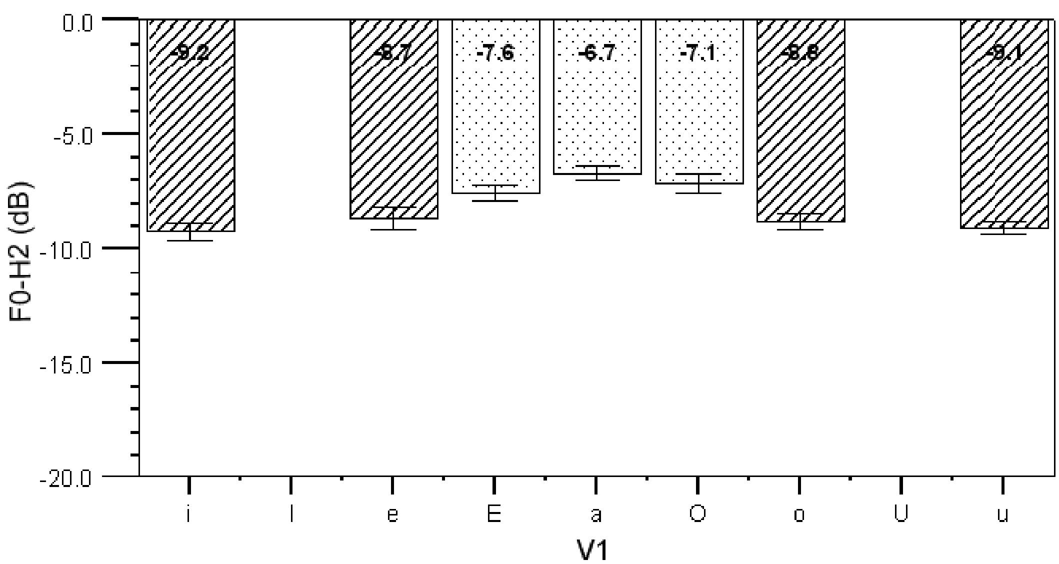


Figure 3.27. Mean F0 − H2 (dB) for V1 midpoint for Mọ̀bà speaker Mb10. The error bars mark one standard deviation. The [+ATR] vowels are [i, e, o, u]; the [−ATR] vowels are [E, O, a], corresponding to [ẹ, ọ, a].

Finally, for SY speaker SY1, shown in Figure 3.28, we see that the [+ATR] value is again lower than the [−ATR] value for the mid vowels. One-way ANOVAs show that the difference is significant for both vowel pairs: for {e/ẹ}, [$F(1, 52) = 44.3$, $p < .01$], and for {o/ọ}, [$F(1, 51) = 31.2$, $p < .01$].

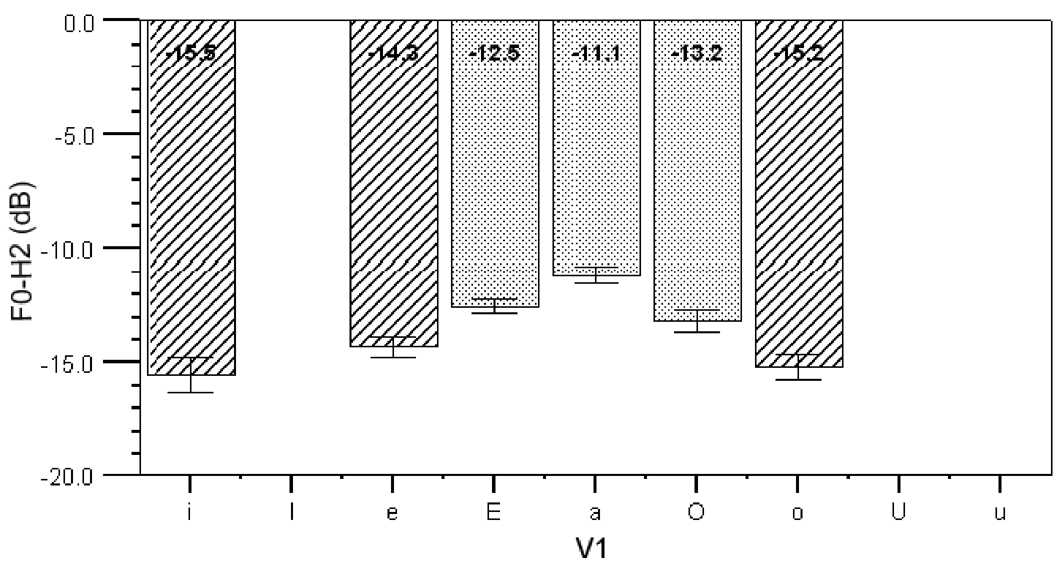


Figure 3.28. Mean F0 − H2 (dB) for V1 midpoint for SY speaker SY1.

To summarize for this measurement, the three speakers showed a correlation between ATR value and the measurement, F0-H2 (dB). The measurement would be consistent with [+ATR] vowels being breathier than their [−ATR] counterparts. However, the correlation is more likely attributed to the lower F1 values in the [+ATR], which results in [+ATR] vowels having a higher amplitude for H2, and thus a greater amplitude increase from F0 to H2.

3.6.7.2. Results of amplitude F0 − amplitude F1

The second spectral measure examined is the difference between the amplitude of the first harmonic and the amplitude of the highest harmonic within the first formant. Breathy vowels have a relatively higher value than creaky vowels. However, for these data, this measure suffers from a similar problem as occurred with the previous measure. Because the [+ATR] vowels have lower F1 values than their

[−ATR] counterparts, we expect that the amplitude of the F1 values will differ due to frequency alone. The graph in Figure 3.29 shows that this is the case for Àkùrẹ́ speaker Ak2. As the frequency of F1 increases, the amplitude of the highest peak within F1 increases.

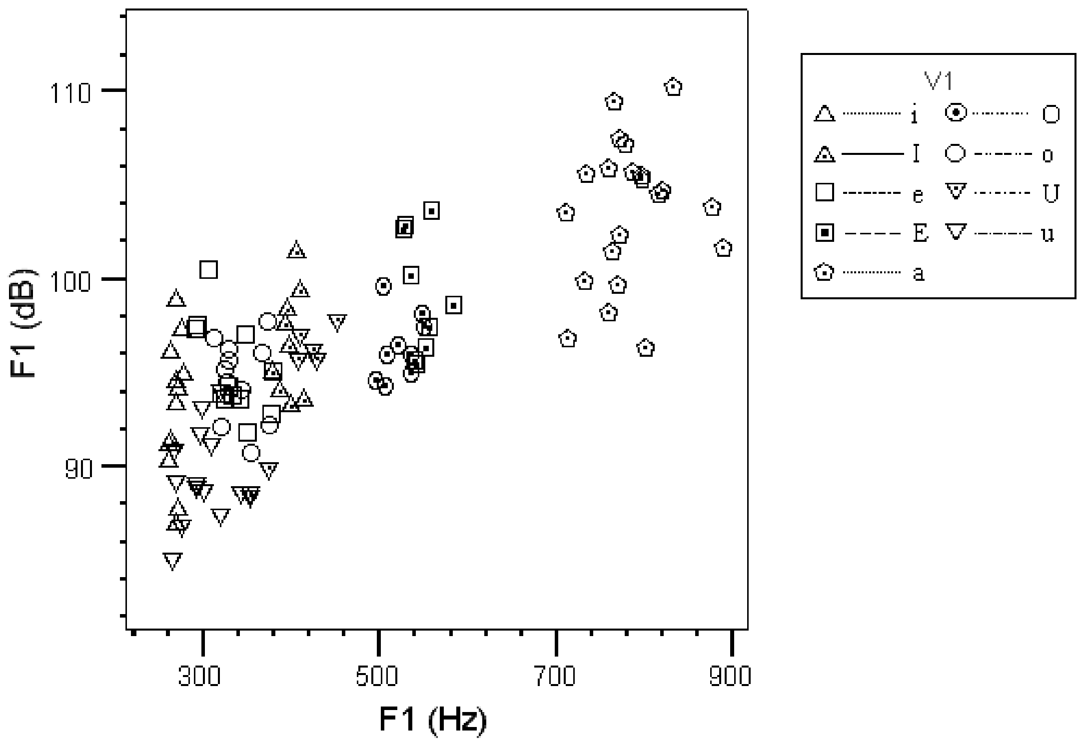


Figure 3.29. F1 (dB) versus F1 (Hz) for V1 in VkV tokens for Àkùrẹ́ speaker Ak2.

Because the vowel pairs being examined differ in F1 frequency, and the F1 frequency correlates to some degree with F1 amplitude, the measure in question can not reliably indicate phonation type for these vowel pairs. That is, a higher F1 frequency will result in a lower value for F0-F1 (dB), which is what is found for Ak2 and SY1. The results of this measure on all three speakers are shown below.

In Àkùrẹ́ speaker Ak2, this measure shows a clear distinction for the pairwise comparison, with each pair showing a lower value for the [−ATR] value than for the [+ATR] value, as seen in Figure 3.30. One-way ANOVAs show that the difference is significant for all pairs: for {i/ị}, [F(1, 40) = 10.5, p < .01]; for {e/ẹ}, [F(1, 37) = 66.0, p < .01], for {o/ọ}, [F(1, 40) = 138.3, p < .01]; and for {u/ụ}, [F(1, 40) = 58.3, p < .01]. The differences were in the direction expected by Halle and Stevens, with [+ATR] vowels having higher value indicating breathiness; however, the difference may instead be attributed to the F1 frequency's influence on F0-F1 (dB) and so no conclusion may be drawn regarding phonation type.

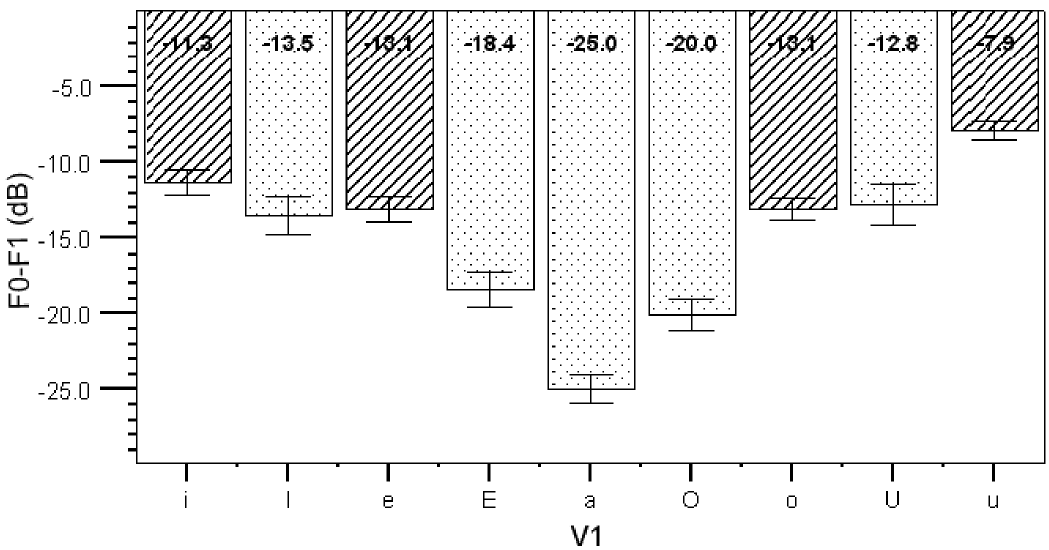


Figure 3.30. Mean F0 − F1 (dB) for V1 midpoint for Àkùrẹ́ speaker Ak2. The error bars mark one standard deviation. The [+ATR] vowels are [i, e, o, u]; the [−ATR] vowels are [I, E, O, U, a], corresponding to [ị, ẹ, ọ, ụ, a].

The SY speaker, SY1, shows a similar pattern for its mid vowel pairs, shown in Figure 3.31, with [−ATR] vowels having a lower value than [+ATR] vowels. Once again, one-way ANOVAs show a significant difference, again in the expected direction: for {e/ẹ}, [$F(1, 52) = 7.8$, $p < .01$]; and for {o/ọ}, [$F(1, 51) = 32.4$, $p < .01$].

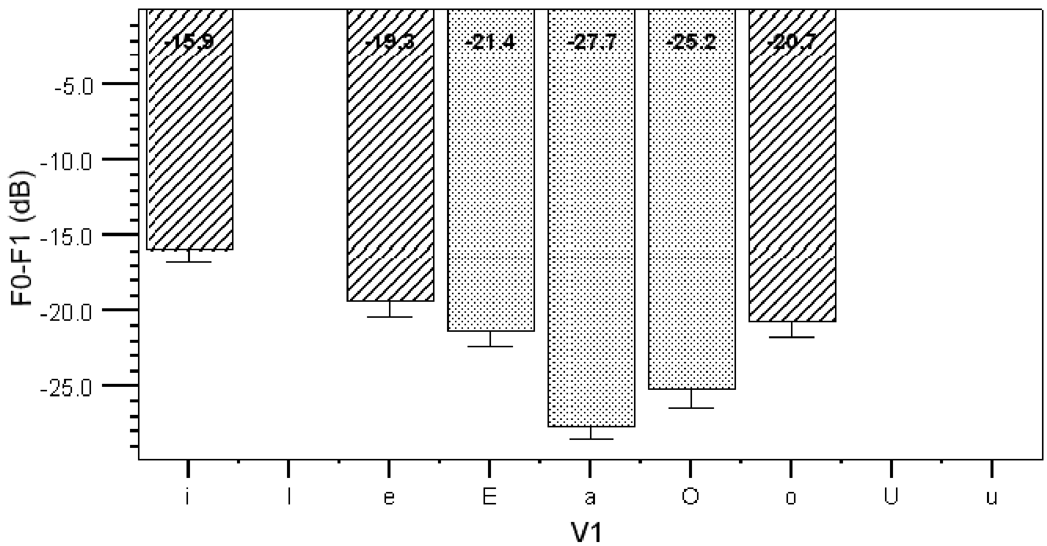


Figure 3.31. Mean F0 − F1 (dB) for V1 midpoint for SY speaker SY1. The error bars mark one standard deviation. The [+ATR] vowels are [i, e, o]; the [−ATR] vowels are [E, O, a], corresponding to [ẹ, ọ, a].

Mọ̀bà speaker Mb10 patterns differently than the Àkùrẹ́ and SY speakers, as shown in Figure 3.32. While for both pairs of mid vowels, there are different values for [+ATR] versus [−ATR], the differences are in the opposite direction from the other dialects, with [+ATR] vowels having a lower value than [−ATR] vowels. One-way ANOVAs show the difference as significant for {e/ẹ}, [$F(1, 52) = 28.2$, $p < .01$]; and nearly significant for {o/ọ}, [$F(1, 52) = 7.1$, $p = .01$]. The difference in the direction of differences between the two sets of vowels is not investigated.

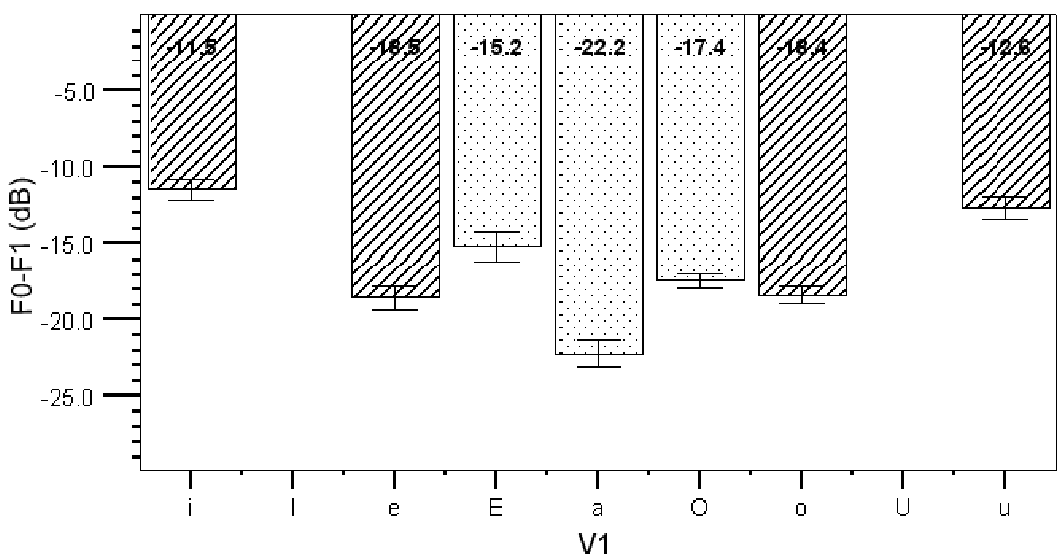


Figure 3.32. Mean F0 − F1 (dB) for V1 midpoint for Mọ̀bà speaker Mb10 The error bars mark one standard deviation. The [+ATR] vowels are [i, e, o, u]; the [−ATR] vowels are [E, O, a], corresponding to [ẹ, ọ, a].

To summarize the results of this measurement, the vowel pairs in Ak2 and SY1 showed a difference that is consistent with Halle and Stevens' claim regarding the relative breathiness of [+ATR] vowels. However, due to the correlation of the F1 amplitude with F1 frequency, and the differences in F1 between vowels across the sets, no such conclusion can be drawn. Because the results for the previous two measures were inconclusive, due to the requirement that F1 values be similar between the vowels being compared, I examined the vowels of a second Àkùrẹ́ speaker.

3.6.7.3. Examining a second Àkùrẹ́ speaker

For Àkùrẹ́ speaker Ak6, F1 values for [+ATR] mid vowels were relatively close to the F1 values of [−ATR] high vowels, as shown in Figure 3.33. The average F1 values for [ị] and [e] are 366 Hz and 393 Hz, respectively; and for [ụ] and [o] are 422 Hz and 390 Hz, respectively. The values for these pairs are the closest of any Àkùrẹ́ speaker measured. The approximately 30 Hz difference may still render these measures unreliable.

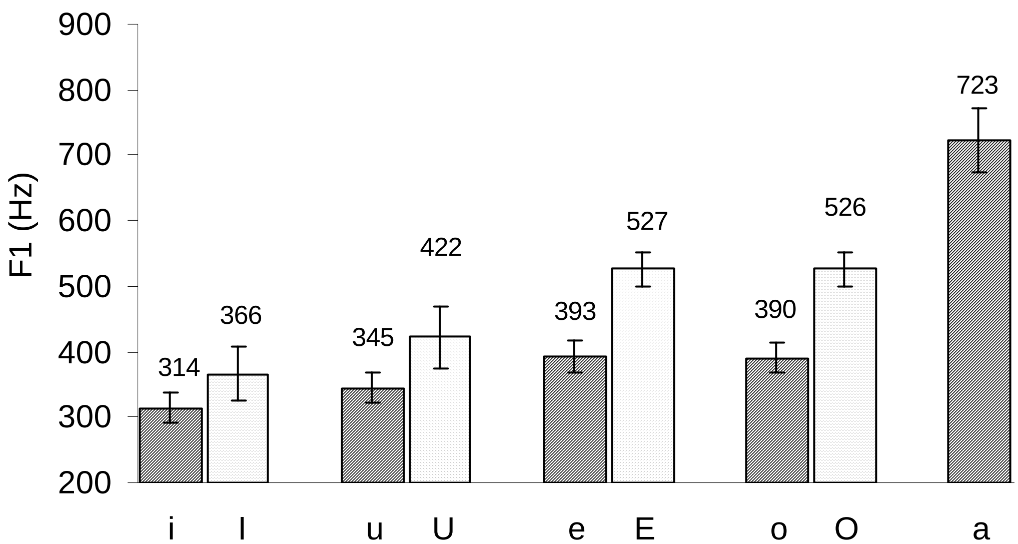


Figure 3.33. Mean F1 values at V1 midpoint for VCV words for Àkùrẹ́ speaker Ak6. The error bars mark one standard deviation. The [+ATR] vowels are [i, e, o, u]; the [−ATR] vowels are [I, E, O, U, a], corresponding to [ị, ẹ, ọ, ụ, a].

The results of the first measure, F0 - H2 (dB), for speaker Ak6 are shown in Figure 3.34. The [±ATR] vowel pairs I compare are [e] versus [ị] and [o] versus [ụ]. For the front pair [ị/e], the difference is small, and not significant: [$F(1, 40) = 4.5$, $p = .04$]; for [ụ/o], the difference is significant [$F(1, 40) = 15.1$, $p < .01$], however the direction is opposite that suggested by Halle and Stevens, with [ụ] having a higher value than [o] thus indicating a breathier vowel.

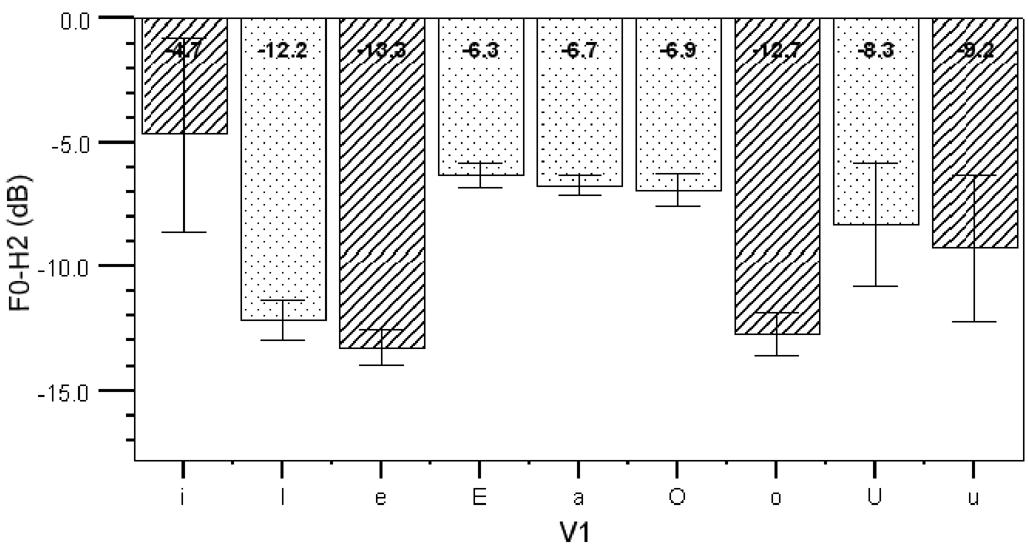


Figure 3.34. Mean F0 − H2 (dB) for V1 midpoint for Àkùrẹ́ speaker Ak6. The error bars mark one standard deviation. The [+ATR] vowels are [i, e, o, u]; the [−ATR] vowels are [I, E, O, U, a], corresponding to [ị, ẹ, ọ, ụ, a].

The results of the second measure, F0 – F1 (dB), for speaker Ak6 are shown in Figure 3.35. For [ị/e], the difference is small, and not significant: [$F(1, 40) = 3.1$, $p = .09$]; for [ụ/o], the difference is significant [$F(1, 40) = 8.6$, $p < .01$]. Again, the direction is opposite that suggested by Halle and Stevens, with [ụ] having a higher value than [o] thus indicating a breathier vowel.

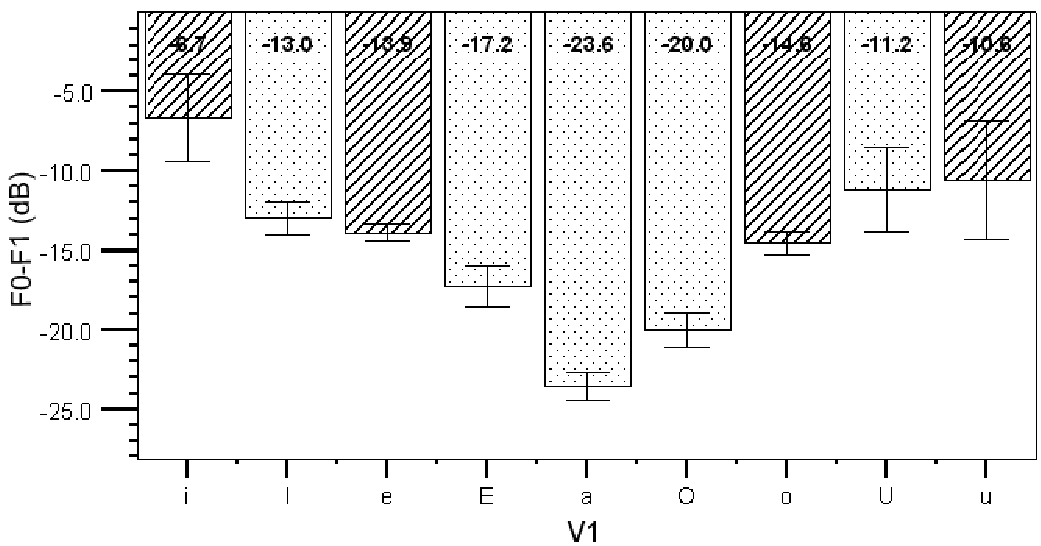


Figure 3.35. Mean F0 − F1 (dB) for V1 midpoint for Àkùrẹ́ speaker Ak6. The error bars mark one standard deviation. The [+ATR] vowels are [i, e, o, u]; the [−ATR] vowels are [I, E, O, U, a], corresponding to [ị, ẹ, ọ, ụ, a].

At best, these measures indicate that there is a phonation type distinction for speaker Ak6 in the back vowels, specifically, [o] versus [ụ], but in the opposite direction from that suggested by Halle and Stevens. Moreover, as the measure does not distinguish the front vowels, the spectral measures are not effective for characterizing [±ATR] vowel pairs for this speaker.

3.6.8 Summary of acoustic measures

The spectral measures tested here were not found to correlate to the two vowel sets under investigation. The measures here are rather difficult to use, requiring other acoustic characteristics to be similar—chiefly F1, but also F0—between the sets in order to ensure a reliable measure. An examination of an additional speaker provided results that did not show a clear pattern of phonation type correlated with vowel sets. While there may be some correlation with phonation type and [±ATR] vowel sets, either for some dialects, some speakers, or some vowels, the pattern was clearly not robustly demonstrated across these three speakers. F1 on the other hand shows a clear and exceptionless difference between the two sets of vowels for each of the three speakers.

3.7 Different kinds of effects

Now that the acoustic measure to be used in the rest of the study has been determined, I use this measure to closely examine vowel to vowel interactions in the three dialects of Yorùbá. In this study, the comparison of the acoustic measure—F1 from now on—is always between two sets of vowels. In some cases, the sets comprise tokens from different phonemes, as a comparison between /e/ and /ẹ/ in all dialects; in some other cases, the sets comprise tokens from different allophones of the same phoneme, as for /i/ and /ị/ in Àkùrẹ́, and in the other cases, the sets comprise tokens from a phoneme with no allophonic split, as in the comparison of [a] before [+ATR] vowels and [a] before [−ATR] vowels in all three dialects.

For the vowel set comparisons of this study, there are three characteristics of acoustic effects that are presented here. First, the magnitude of the acoustic effect is measured. For simplicity, this is measured at one point—the midpoint of the first vowel—as an approximation of the magnitude for the first vowel as a whole. The magnitude is measured in the difference of the average of acoustic effect in one set versus the other. Accompanying this difference of the means is a measure of the statistical significance. Second, the overlap of the effect is evaluated. The overlap is not quantified, *per se*, but is graphically presented. Third, the temporal characteristic of the effect is presented. This characteristic involves the first two characteristics—magnitude and overlap—extended throughout the duration of the vowel. An example of these three characteristics follows.

In this example, I compare the F1 values in the phonemes [e] and [ẹ] for Àkùrẹ́ speaker Ak2. Figure 3.36 shows the mean F1 values for all nine Àkùrẹ́ phones at the V1 midpoint. The magnitude of the difference between the means of [e] and [ẹ] is large, at 213 Hz. A single-factor ANOVA indicates that the difference is highly significant: [F(1,38) = 555, p < .001].

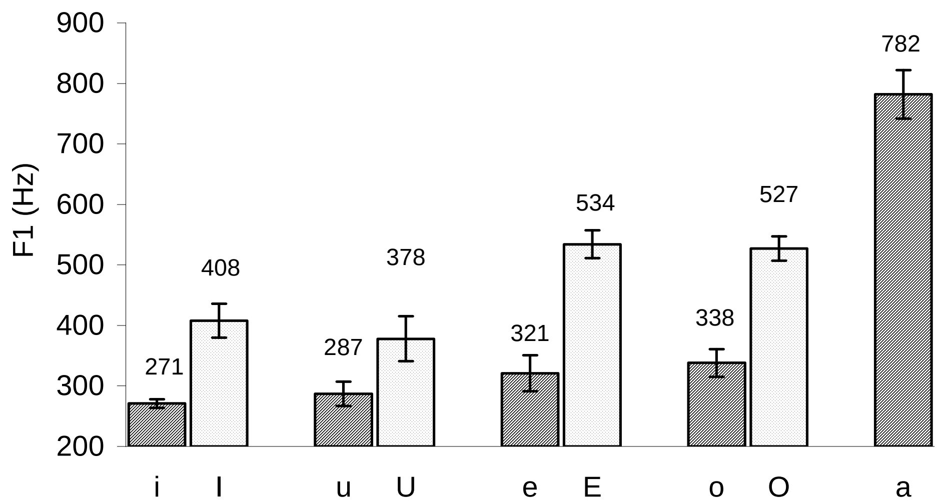


Figure 3.36. Mean F1 values at V1 midpoint for VCV words for Àkùrẹ́ speaker Ak2. The error bars mark one standard deviation. The error bars mark one standard deviation. The [+ATR] vowels are [i, e, o, u]; the [−ATR] vowels are [I, E, O, U, a], corresponding to [ị, ẹ, ọ, ụ, a].

In order to establish the overlap between the two sets of tokens, the individual values must be examined. Figure 3.37 shows the midpoint F1 values for the four Àkùrẹ́ speakers for the vowels [e] and [ẹ]. It is clear from this figure that for all Àkùrẹ́ speakers there is no F1 overlap between the [e] and [ẹ] tokens; that is, for each speaker, all the [e] values are lower than all the [ẹ] values.

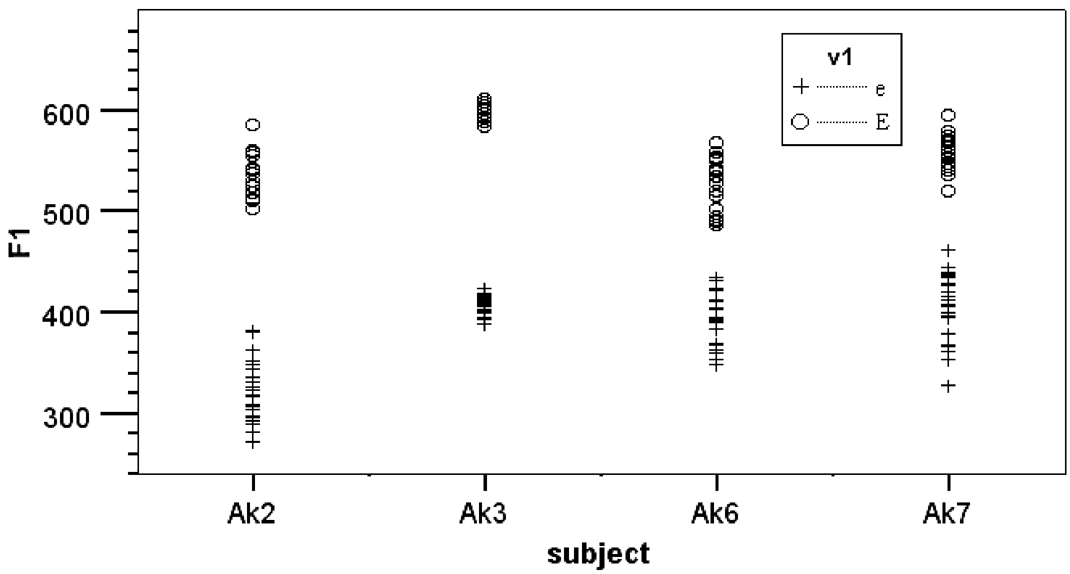


Figure 3.37. Individual F1 values at V1 midpoint for eCV and ẹCV tokens for four Àkùrẹ́ speakers.

Finally, the temporal characteristic is examined. Figure 3.38 shows the mean F1 values and standard deviation throughout the duration of the vowel. The difference between these vowel sets extends throughout the vowels.

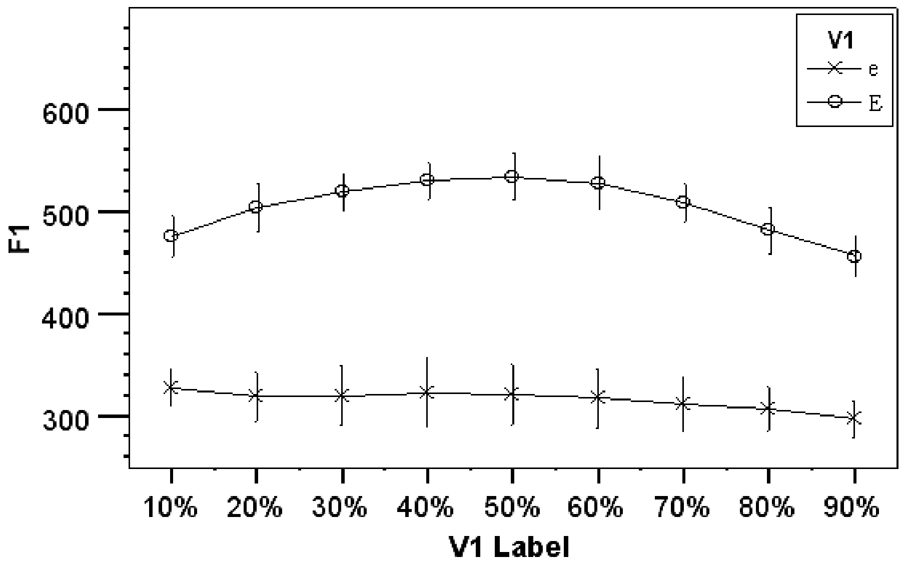


Figure 3.38. Mean F1 values at nine points throughout the V1 for eCV and ẹCV tokens for Àkùrẹ́ speaker Ak2. The error bars mark one standard deviation.

The standard deviations shown in the figure suggest no overlap throughout the vowel; this is corroborated by Figure 3.39 which shows the individual token values throughout the vowel. From this series of graphs on the Ak2 vowels /e/ and /ẹ/, it is clear that there is a large magnitude difference. These types of graphs are presented as needed in chapter 4.

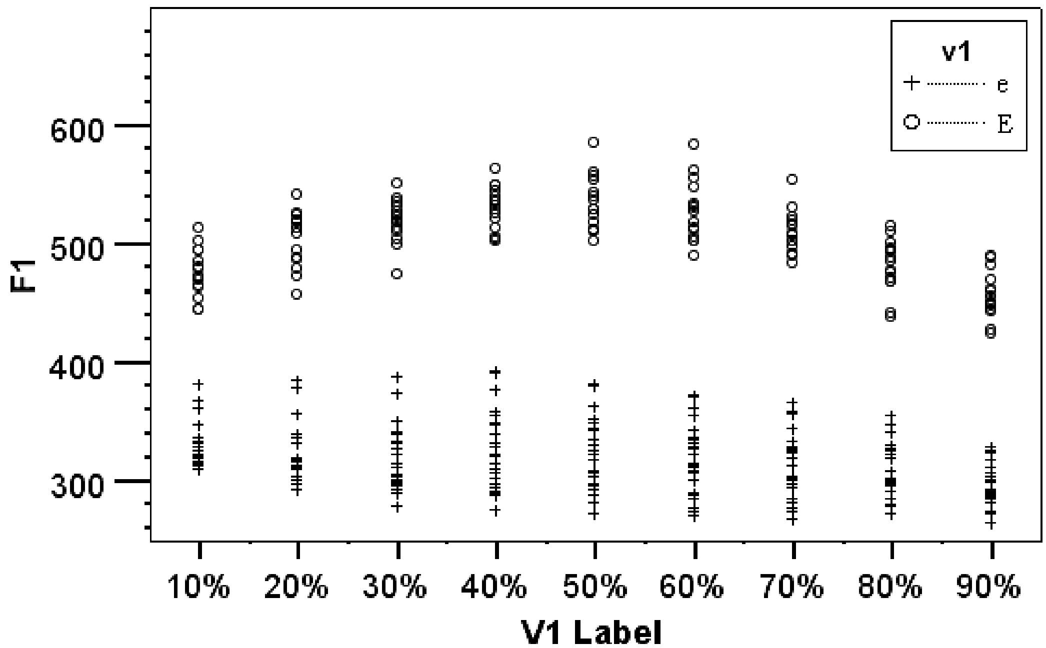


Figure 3.39. Individual F1 values at nine points throughout the V1 for eCV and ẹCV tokens for Àkùrẹ́ speaker Ak2.

It is well known that a frequency difference of, say, 50 Hz may be rather small for a low vowel, but relatively larger for a high vowel due to the non-linear nature of the perception of frequency (Stevens, Volkmann, and Newman 1937). However, as this is an acoustic study and not a perceptual one, no attempt has been made to transpose the measures from physical units such as hertz into more perceptually relevant (and non-linear) units such as mels.

It is not the case that a correlation between an acoustic parameter and a potential phone pair indicates phonological contrast. A robust contrast would necessarily have values in one or several acoustic parameters (or a combination thereof) that would differ significantly. However, such a significant difference by itself does not necessarily imply contrast. A test for contrast requires more. Given a pair of phonemes, we certainly may encounter tokens that are incorrectly classified, for we know that some phonemes are confusable with others. But to the extent that these phonemes are perceivably different, we expect that the acoustic parameter distinguishing the phonemes should not overlap, and the difference should be highly significant. Some phonemes may have more overlap than others, as in the case of [a] and [ɔ] ('chock' versus 'chalk') which are merging in some variants of American English, and have merged completely in others (Labov 1991). However, in this case, a linguist may know ahead of time, from consultant feedback, that these phonemes are sometimes confused. Thus, assuming a given speaker speaking at the same rate, pitch, and voice quality, and idealizing one contrastive acoustic parameter, we expect that a histogram of tokens from both sets will exhibit a bimodal distribution with no overlap, as illustrated in Figure 3.40.

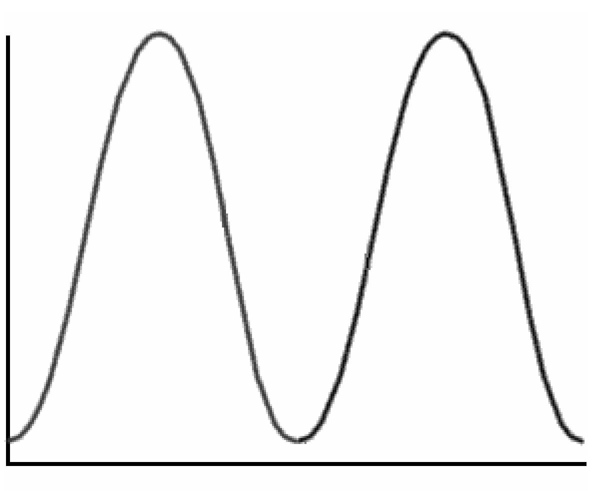


Figure 3.40. Idealized histogram of measurements of tokens from two phonemes, with the x-axis being the distinguishing acoustic measure.

The difference between the two curves would necessarily be large enough to be perceptually salient. Coarticulatory differences will differ from phonemic differences. There will probably be overlap between the tokens of each set, as schematized in Figure 3.41. The difference between the sets may or may not be significant, but will be less significant than the phonemic differences.

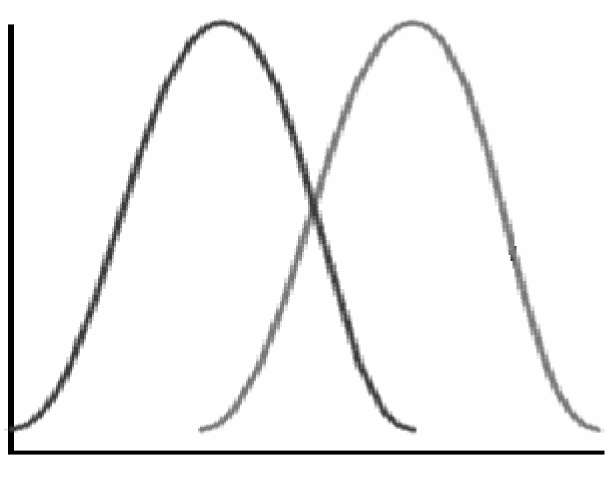


Figure 3.41. Idealized histogram of measurements of tokens from two phone sets showing some overlap, with the x-axis being the distinguishing acoustic measure.

With phonemic differences and coarticulatory differences being the two endpoints in behavior, allophonic differences will occur anywhere between. In the case of allophones, we could have the same sort of distribution as occurs with phonemes. However, it can also be the case that there is some overlap between the two distributions, as there often may be overlap in the case of allophones. Allophonic pairs do not have to exhibit contrast, as they generally do not bear the functional burden of contrast.[6] Sproat and Fujimura (1993) found considerable overlap in the light and dark allophones of English /l/; however they conclude that these are in fact not allophonic

6 This is not always the case, as in ‘writer’ versus ‘rider’ in some variants of American English. In this case, the two allophones of /ay/ exhibit the only contrast between the words due to the neutralization of /d/ and /t/.

variants. Certainly some allophonic differences have contrasts with differences resembling phonemic differences. Of course, in order to be called allophones, phones must differ enough to be recognized as distinct by the speaker, or more relevantly in this case, the linguist. If such a difference is recognizable, then there will be some degree of difference between the two sets. While allophonic behavior that resembles phonemic behavior can easily be classified as allophonic, allophonic behavior that approaches coarticulatory behavior can not be as easily classified. In fact, it does not seem possible to posit criteria to distinguish between allophonic and coarticulatory behavior, as can be seen in the results of Sproat and Fujimura (1993). In the case of Àkùrẹ́ high vowels, some speakers I consulted recognize that [i] and [u] differ from [ị] and [ụ], while others do not. A similar result was obtained from discussions of Ijẹṣa speakers, Ijẹṣa being a closely related dialect with the same phonological behavior with respect to high vowel harmony.

In this chapter, I identified F1 as the acoustic parameter that best correlates with the two sets under investigation. We see in the next chapter that in some cases the differences underlie phonological differences (that is, phonemic or allophonic), while in other cases, the sets indicate differences similar to coarticulatory ones.

3.8 Summary

This chapter has presented the methodology used in this experiment. In addition, a pilot experiment looking at one speaker of each of the three dialects has determined that F1 is the most reliable correlate to the feature [ATR] in the high vowels in Àkùrẹ́ and mid vowels in all three dialects. In the next chapter, I use F1 to characterize vowels involved in harmony, most notably in the high vowels of Àkùrẹ́. I then examine the high vowels in the SY and Mọ̀bà, where high vowel harmony is not found, to investigate the vowel to vowel coarticulation in order to show that coarticulation in these dialects resembles high vowel harmony in Àkùrẹ́. Vowel to vowel interactions between vowel sets are compared with respect to magnitude of acoustic effect, overlap between sets, and, to a lesser degree, temporal characteristics.

CHAPTER FOUR

RESULTS

4.1. Introduction

In this chapter, the results of the acoustic study are presented. In order to show that vowel harmony originates from vowel-to-vowel coarticulation, I show that the two phenomena resemble each other. Learners of Àkùré must be able to deduce the patterns of its high vowel harmony from the acoustic characteristics of the linguistic input they are exposed to. If vowel-to-vowel coarticulation is to have led to high vowel harmony, then these patterns must also have been present in the acoustic characteristics of earlier forms of Yorùbá which lacked the high vowel harmony, though to a much lesser degree than is found in Àkùrẹ́. Since earlier forms are not available, I examine the vowels of SY and Mọ̀bà—two dialects that are likely similar to the proto-language—for evidence of the patterns of the vowel harmony from the acoustic characteristics of the coarticulation.

In the previous chapter, we saw that F1 was the best acoustic correlate to ATR in the dialects of this study. In the first section of this chapter, I examine the phonetic patterns of harmony in Àkùrẹ́. The vowel harmony patterns of Àkùrẹ́ are then summarized according to the characteristics shown in (1), which speakers of Àkùrẹ́ must know at some level:

(1) Characteristics to identify in Àkùrẹ́ harmony

- targets — what are the targets affected by vowel harmony;
- contexts — in what contexts does the vowel harmony occur;
- effect — what is the acoustic effect realized in the harmony.

By examining the phonetic evidence, we corroborate the contexts and targets of the harmony in Àkùrẹ́ that had been determined using phonological evidence in Chapter 2; and we present the robust acoustic effect of the harmony. Once we have seen the phonetic realization of vowel harmony, we then examine the same environment in Mọ̀bà and SY. These dialects share the same vowel phoneme inventory as Àkùrẹ́, and, with the exception of high vowel harmony, are phonologically quite similar. If Àkùrẹ́ developed its high vowel harmony from vowel-to-vowel coarticulation in an earlier language, then the earlier language would have looked like Mọ̀bà and SY (but recall that SY lacks /u/ in vowel initial position) at least in terms of the vowel system. Since we are unable to study the acoustics of the vowels of Proto-Yorùbá, we instead must examine the vowels of Mọ̀bà and SY to look for the seeds of harmony. The results for these two dialects, presented in §4.3 and §4.4, show that vowel to vowel coarticulation effects share some of the characteristics found in Àkùrẹ́ vowel harmony. The effect is quantitatively much smaller for Mọ̀bà and SY, but statistically significant. The effect can be characterized as affecting the same target vowels in the same contexts.

In addition to examining the parallels between harmony and coarticulation, an additional goal of this study is to provide a detailed cross dialectal examination of the vowels of these dialects of Yorùbá. Acoustic data from Àkùrẹ́ and Mọ̀bà Yorùbá have not been presented before this study, and although Standard Yorùbá has a rich phonological literature, it has been the subject of few acoustic studies (Disner 1983, Lindau and Wood 1977a, b).

4.1.1. Overview of this chapter

In §4.2, I examine the vowels of Àkùrẹ́ speakers, starting with contrastive mid vowels in §4.2.1, and moving to allophones within the same phoneme in §4.2.2. Of particular interest here is the high vowel harmony. In §4.3, I examine the vowels for four speakers of Mọ̀bà. In §4.4 I examine SY. Finally, in §4.5, I compare the **phonetic** effects of coarticulation in high vowels in Mọ̀bà and SY with the **phonological** effects of harmony in Àkùrẹ́.

4.2. Àkùrẹ́ results

I start by presenting the results from Àkùrẹ́, the dialect exhibiting the greatest degree of phonological vowel harmony. Starting with a broad view, Figures 4.1-4.5 show vowel charts for each Àkùrẹ́ speaker for the midpoint of the initial vowel in VCV. All formant measurements presented here are taken at the midpoint of the vowel, unless otherwise noted, following the methodology described in chapter 3. For each vowel phone, data points are shown for up to three tokens for each word, that is, up to 42 per vowel. For [a], there are three tokens each for *aka*, *akẹ*, *ake*, *akọ*, *ako*, *aki*, *aku*, *aba*, *abẹ*, *abe*, *abọ*, *abo*, *abi*, and *abu*, or 42 data points. For [e], there are three tokens each for *eke*, *eko*, *eki*, *eku*, *ebe*, *ebo, ebi*, and *ebu*, or 24 data points; for [o], [i], and [u], there are also 24 data points. And for [ẹ], there are three tokens each for *ẹka*, *ẹkẹ*, *ẹkọ*, *ẹba*, *ẹbẹ*, and *ẹbọ*, or 18 data points; for [ọ], [ị], and [ụ], there are also 18 data points. Occasionally, fewer data points were measured when recording problems occurred, or when the vowels were not able to be measured. In these figures, the ranges of the F1 and F2 axes were adjusted for each speaker to more clearly show the speaker's vowel space; Figure 4.5 shows all four Àkùrẹ́ speakers with the same F1 and F2 ranges. Keep in mind that these vowel charts represent the vowel tokens within the deliberately limited phonetic environment of this experiment (for example, after an [f] and before [k] or [b]), and thus likely differ from a more general vowel chart one could construct using a wider variety of contexts.

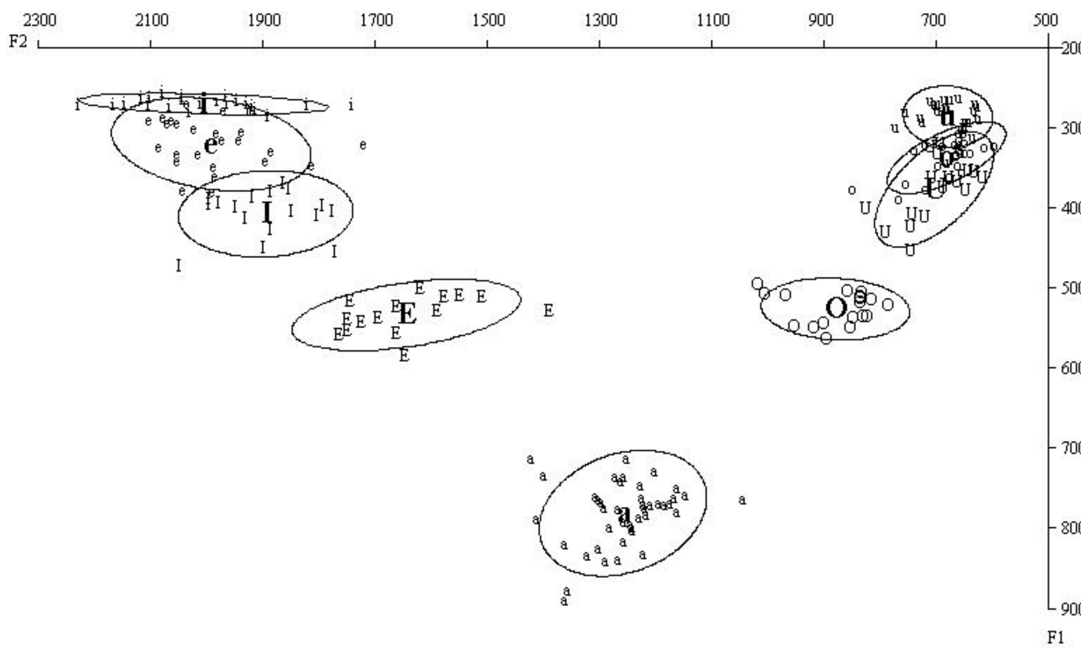


Figure 4.1. Vowel Chart (F1 versus F2) of Àkùrẹ́ Speaker Ak2 for the midpoint of V_1 in V_1CV_2.The small letters indicate the values for individual tokens; the large letters show the mean for the vowel; and the ellipses show one standard deviation from the mean for each vowel. The [+ATR] vowels are [i, e, o, u]; the [−ATR] vowels are [I, E, O, U, a].

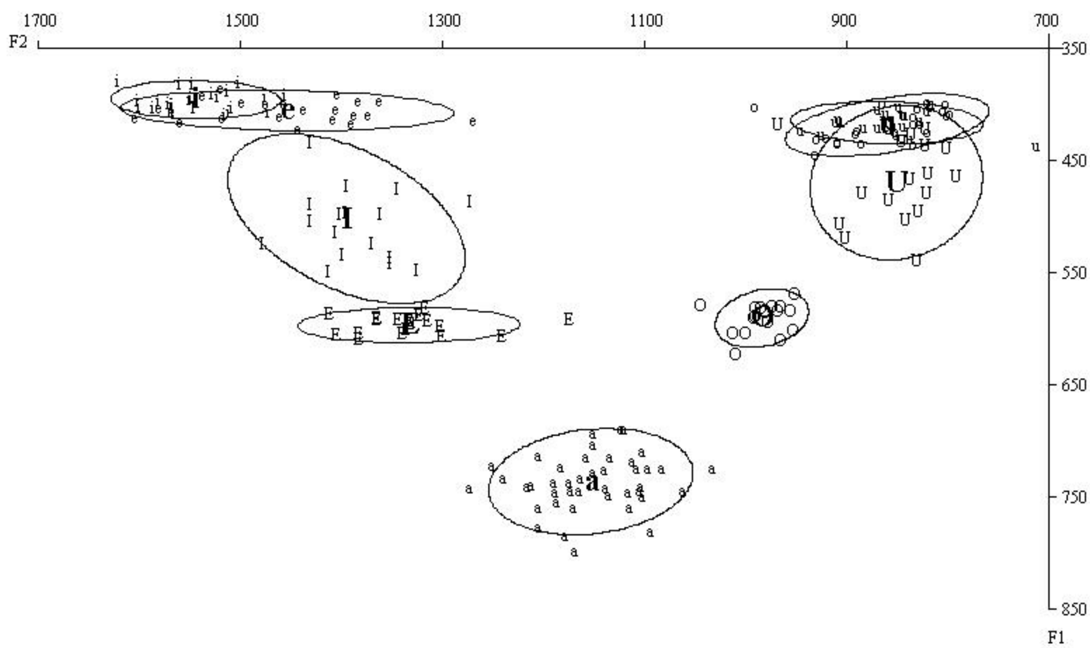


Figure 4.2. Vowel Chart (F1 versus F2) of Àkùrẹ́ Speaker Ak3 for the midpoint of V_1 in V_1CV_2.The small letters indicate the values for individual tokens; the large letters show the mean for the vowel; and the ellipses show one standard deviation from the mean for each vowel. The [+ATR] vowels are [i, e, o, u]; the [−ATR] vowels are [I, E, O, U, a].

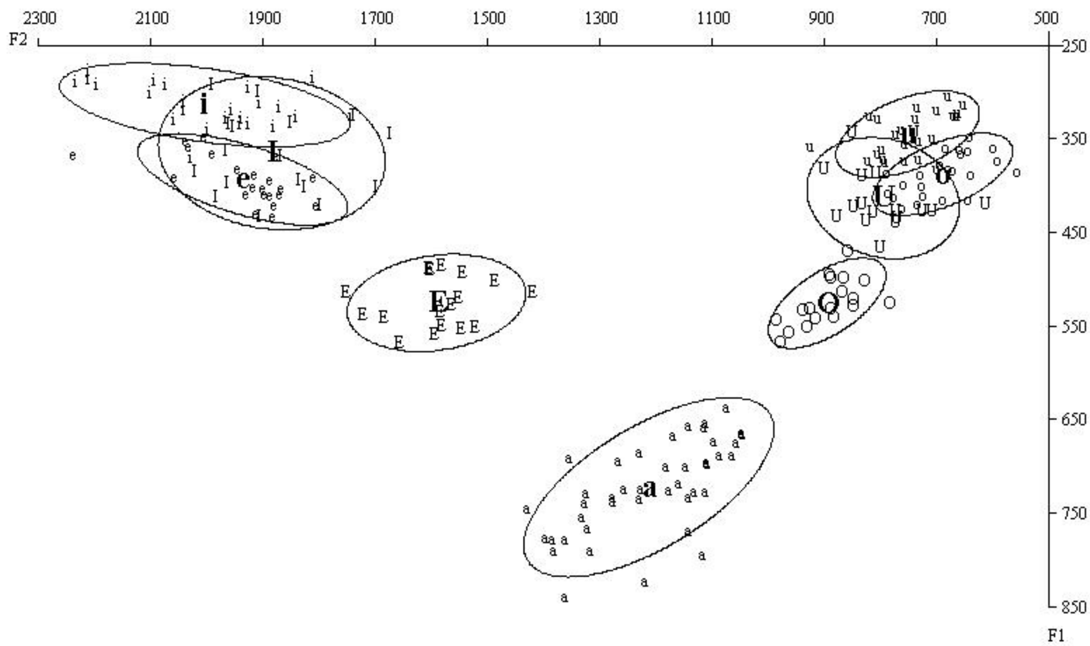


Figure 4.3. Vowel Chart (F1 versus F2) of Àkùrẹ́ Speaker Ak6 for the midpoint of V_1 in V_1CV_2.The small letters indicate the values for individual tokens; the large letters show the mean for the vowel; and the ellipses show one standard deviation from the mean for each vowel. The [+ATR] vowels are [i, e, o, u]; the [−ATR] vowels are [I, E, O, U, a].

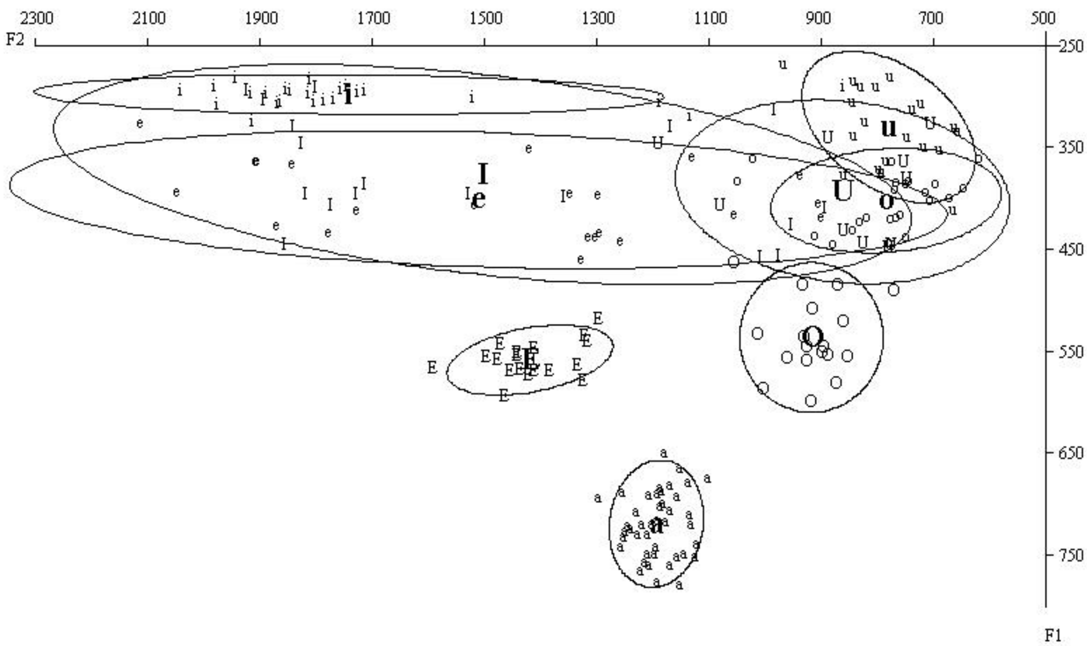


Figure 4.4. Vowel Chart (F1 versus F2) of Àkùrẹ́ Speaker Ak7 for the midpoint of V_1 in V_1CV_2.The small letters indicate the values for individual tokens; the large letters show the mean for the vowel; and the ellipses show one standard deviation from the mean for each vowel. The [+ATR] vowels are [i, e, o, u]; the [−ATR] vowels are [I, E, O, U, a].

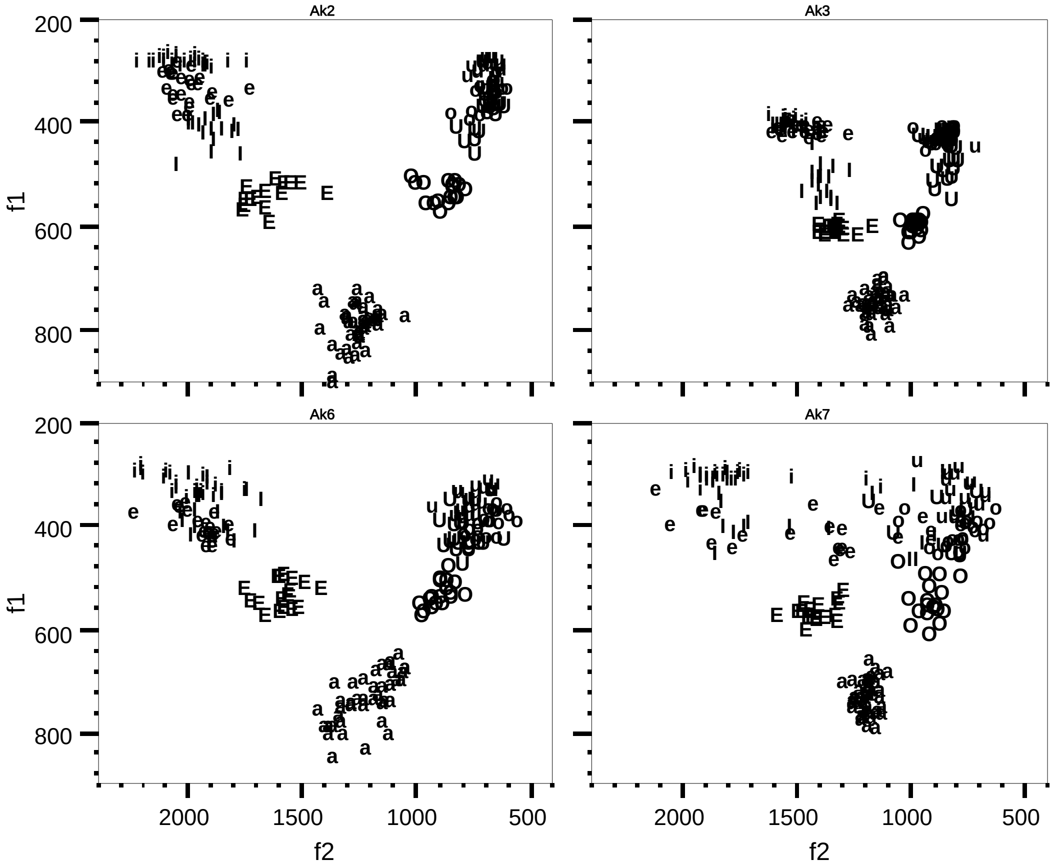


Figure 4.5. Vowel Charts (F1 versus F2) of Àkùrẹ́ Speakers Ak2, Ak3, Ak6, Ak7 for the midpoint V_1 of V_1CV_2. The [+ATR] vowels are [i, u, e, o]; the [−ATR] vowels are [ị, ụ, ẹ, ọ, a], shown here as [I, U, E, O, a]. The ranges of the axes are identical for all speakers.

In all four Àkùrẹ́ speakers, the low vowel, [a], and the [−ATR] mid vowels, [ẹ] and [ọ], show little or no overlap with other vowels. The remaining six vowels show considerable overlap in the F1 dimension with adjacent vowels, with back vowels [u], [o], and [ụ] showing the most overlap, and the front vowels [i], [e], and [ị] also grouping together. As is typical with data such as these, some speakers exhibit a neat idealized pattern, while others exhibit patterns that are less clear, with less compact vowels. Speakers Ak2, Ak6, and to some degree Ak3, show relatively clear patterns, while Ak7 shows a greater variation and overlap in the vowels. Note, for example, the [e] of Ak7 which is found in the [i] area, as well as in the [o] area. On reconsideration of Ak7's recordings, it is clear that Ak7 spoke very quickly—the vowel durations confirm he is speaking much more quickly than the other subjects—and with a great deal of poor enunciation, which is consistent with the pattern exhibited in Figure 4.4. These vowel patterns, therefore, may not be characteristic of Ak7's speech, but rather they may be due to the faster speech rate. It is interesting to note that for this speaker, the formant averages in Figure 4.4 look reasonable; it is only by observing the individual points or the standard deviation that it becomes clear that this speaker's data are atypical. In light of these aberrant data, the results from Ak7 are left out of further discussion.

We examine how the distinctions between vowels are realized both impressionistically from the vowel charts above, and using quantitative measurements. Three types of differences are presented between [+ATR] and [−ATR] vowels: phonemic, allophonic, and coarticulatory. First I present the F1 differences between phonemes: [+ATR] and [−ATR] mid vowels. Next, the [+ATR] and [−ATR] high vowels are presented; the differences here are allophonic. Finally, I present the coarticulatory effects on [a] from following vowels. Figures 4.6-4.8 show the mean values and standard deviations for the vowels shown in the charts above, and should be used for reference in the following sections.

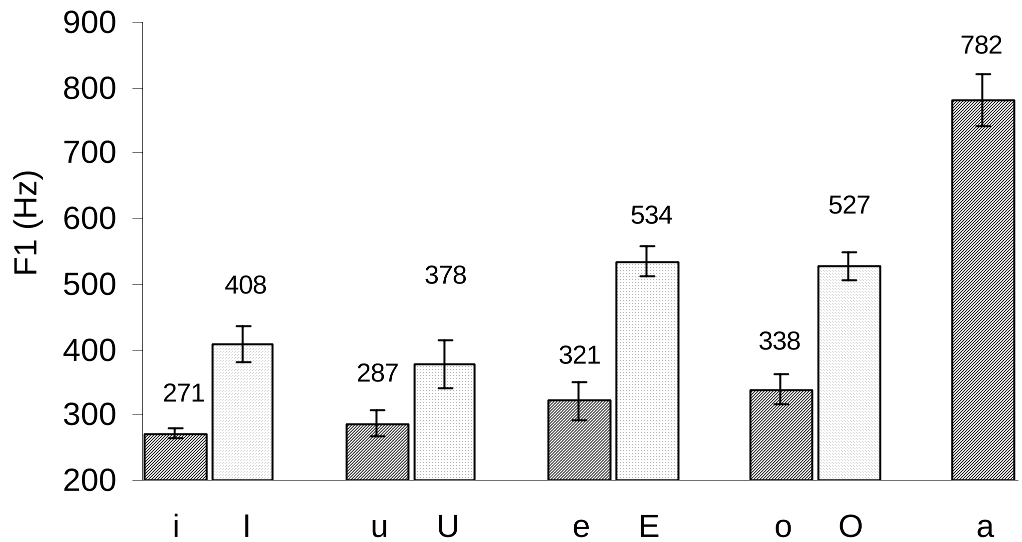


Figure 4.6. Mean F1 values at V1 midpoint for VCV words for Àkùrẹ́ speaker Ak2. The error bars mark one standard deviation.

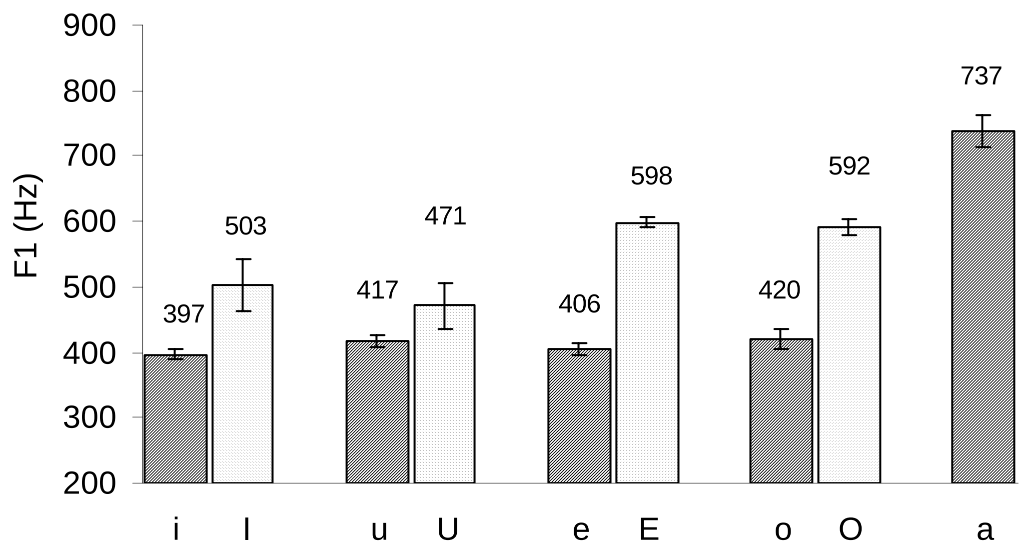


Figure 4.7. Mean F1 values at V1 midpoint for VCV words for Àkùrẹ́ speaker Ak3. The error bars mark one standard deviation.

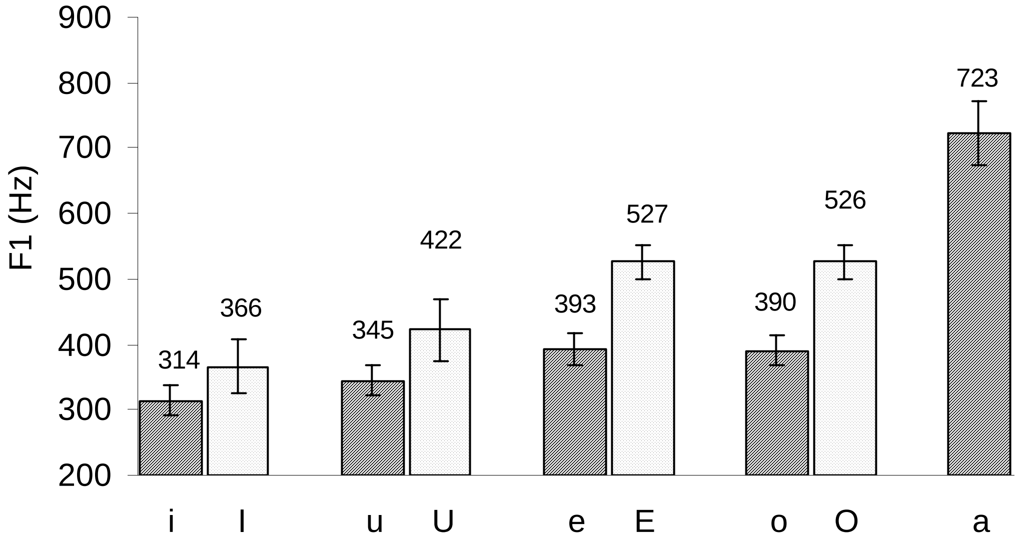


Figure 4.8. Mean F1 values at V1 midpoint for VCV words for Àkùrẹ́ speaker Ak6. The error bars mark one standard deviation.

4.2.1. Different phonemes: [−ATR] versus [+ATR] mid vowels in Àkùrẹ́

As expected, the ±ATR mid vowel pairs show clear differences—the pairs of e-ẹ, o-ọ (shown here as e-E, o-O) differ in both F1 and F2, though with F1 showing a clearer difference with no overlap. For [e] and [ẹ], a single-factor ANOVA shows a highly significant difference for all subjects for F1: for Ak2, the F1 difference was 213 Hz [$F(1,38) = 555$, $p < .001$]; for Ak3, the F1 difference was 192 Hz [$F(1,40) = 4821$, $p < .001$]; and for Ak6, the F1 difference was 134 Hz [$F(1,41) = 281$, $p < .001$]. Additional perspective is gained by focusing on the individual F1 values for all Àkùrẹ́ speakers.

Figure 4.9 shows the F1 of the midpoint of the first vowel in tokens of V_1CV_2, with V_1= [e, ẹ] for all V2 vowels. For all three Àkùrẹ́ speakers, the F1 values for [e] versus [ẹ] show no overlap, that is, the highest F1 value for [e] is lower than the lowest F1 value for [ẹ].

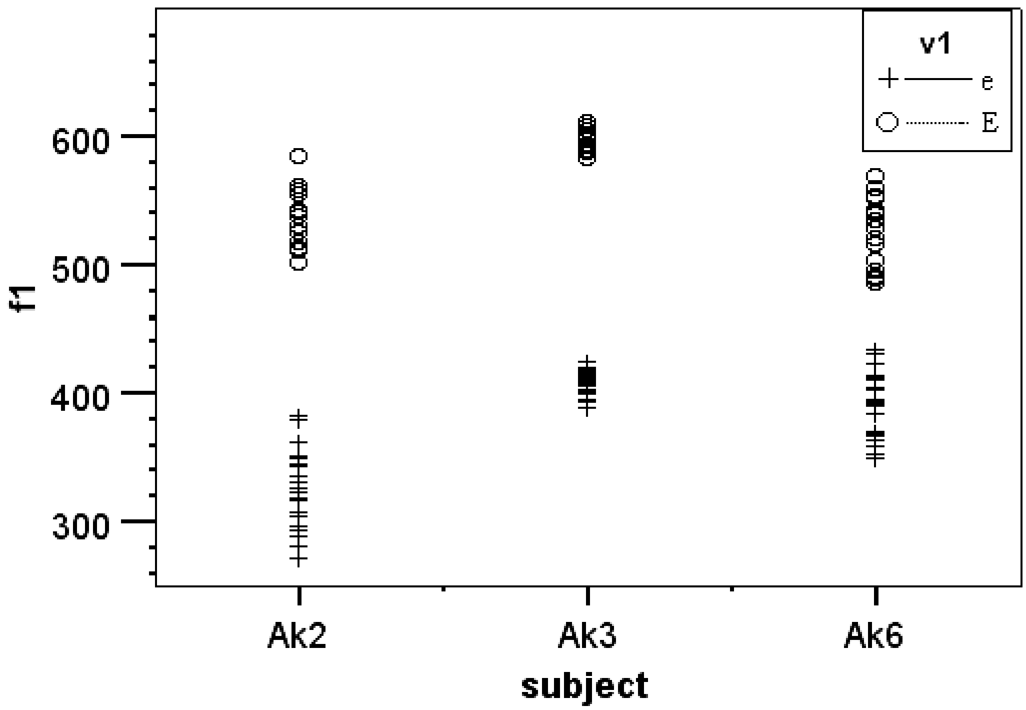


Figure 4.9. F1 values for midpoints of mid vowels [e, ẹ] (e, E, in graph) for Àkùrẹ́ speakers for V1 of VCV tokens for all V2s. The pluses (+) represent F1 values for tokens of [+ATR] [e]s at the mid point of V1s; the circles (○) represent the same for [−ATR] [ẹ].

For [o] and [ọ], F1 shows a highly significant difference for each Àkùrẹ́ speaker. For Ak2, the F1 difference was 189 Hz [$F(1,41)=787$, $p < .001$]; for Ak3, the F1 difference was 172 Hz [$F(1,41) = 1586$, $p < .001$]; and for Ak6, the F1 difference was 136 Hz [$F(1,41) = 336$, $p < .001$]. For all Àkùrẹ́ speakers, the F1 values for [o] versus [ọ] show no overlap, so that the highest F1 value for [o] is lower than the lowest F1 value for [ọ], as shown in Figure 4.10. In summary for the mid vowels in V1, the F1 differences are large and show no overlap, which is what we expect from phonemes.

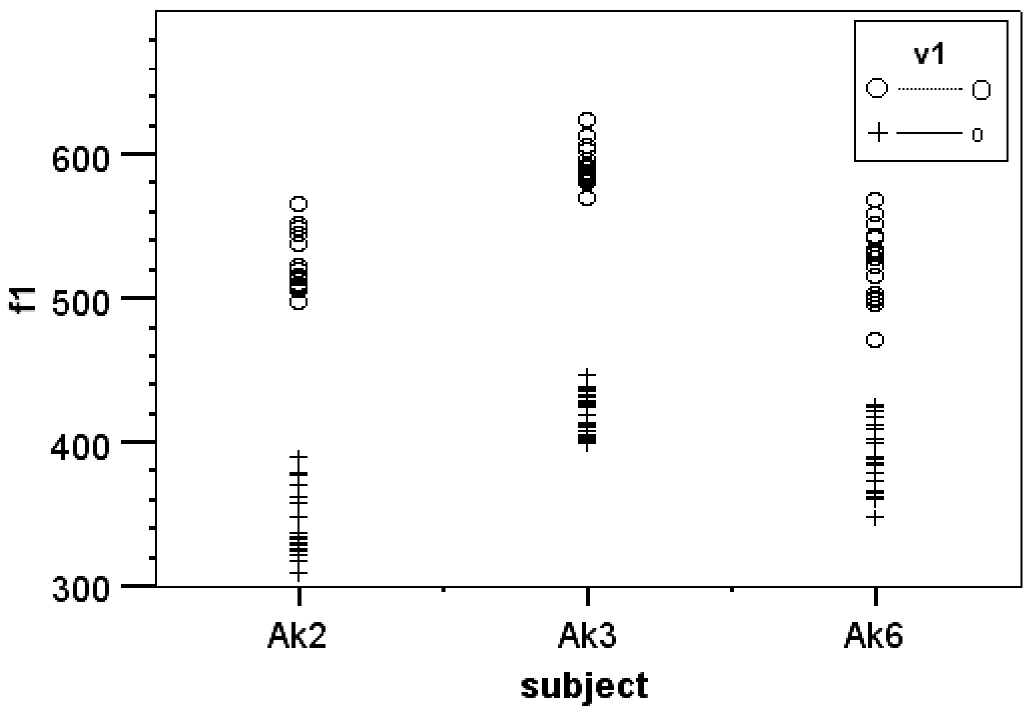


Figure 4.10. F1 values for midpoints of mid vowels [o, ọ] (o, O, in graph) for Àkùrẹ́ speakers for V1 of VCV tokens for all V2s. The pluses (+) represent F1 values for tokens of [+ATR] [o]s at the mid point of V1s; the circles (○) represent the same for [−ATR] [ọ].

Figure 4.11 shows the means for all labels for each Àkùrẹ́ speaker. As expected, the large difference between phonemes [e] and [ẹ] is maintained throughout the vowel. Similar results are found for [o] and [ọ], not shown.

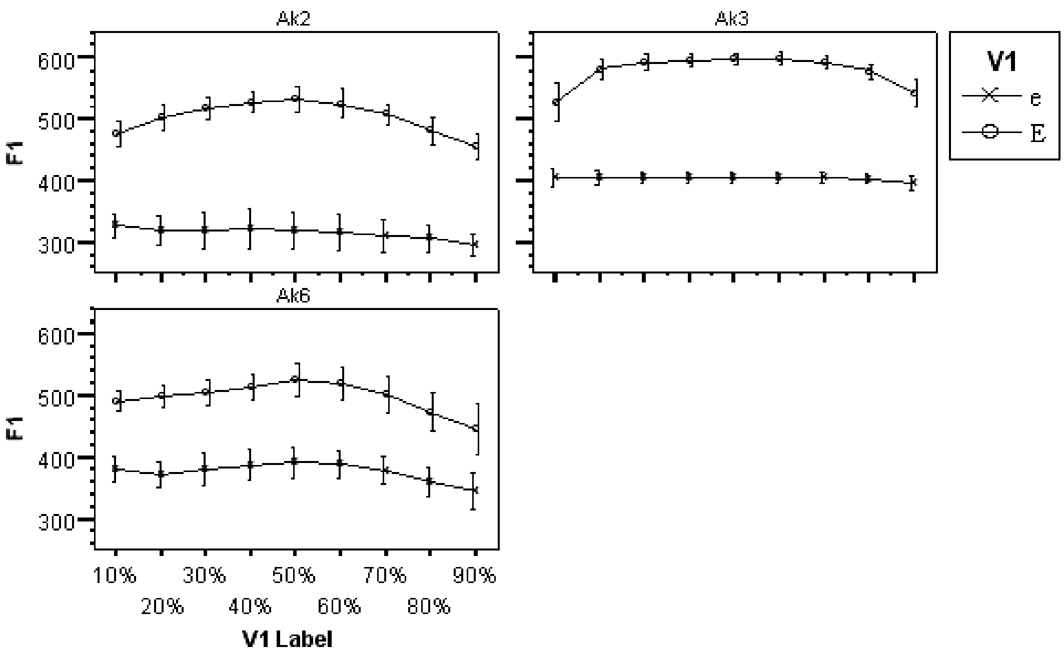


Figure 4.11. F1 means for all labels of mid vowels [e, ẹ] (e, E, in graph) for Àkùrẹ́ speakers for V1 of VCV tokens for all V2s. The x's represent F1 means for tokens of [+ATR] [e] at the mid point of V1; the circles (○) represent the same for [−ATR] [ẹ].

4.2.2. High vowel allophones in Àkùrẹ́

Next the high vowel allophones are examined. The comparison between [+ATR] and [−ATR] high vowels is a critical part of this study. The acoustic realization of the phonological harmony in Àkùrẹ́ is compared to the coarticulation in the same environments in Mọ̀bà and SY, the dialects without phonological high vowel harmony. *A priori*, these allophonic contrasts in Àkùrẹ́ need not be robust—they need not exist at all since they never carry the functional burden of contrastiveness alone. For example, in i*ke* versus *ikẹ*, the second vowel alone distinguishes the two words. Without the allophonic contrast, the Àkùrẹ́ vowel system would still be sufficiently contrastive, indeed it would be like the Mọ̀bà system.

Turning to the results, the Àkùrẹ́ high front vowel pair i-ị (i-I) shows a highly significant difference in F1 for all three Àkùrẹ́ speakers. For Ak2, the F1 difference was 137 Hz [$F(1,41)=546$, $p < .001$]; for Ak3, the F1 difference was 106 Hz

[F(1,41)=169, p < .001]; and for Ak6, the F1 difference was 42 Hz [F(1,41)=26, p < .001]. Figure 4.12 shows the F1 values for the high front pair of [i] and [ị]. For speakers Ak2 and Ak3, the pattern here resembles the pattern for the mid vowels with a gap between the phones. For Ak6, on the other hand, there is an overlap between the F1 of the [+ATR] and [−ATR] vowels. There may be two reasons for the apparent difference in the mid vowel pairs versus the high vowel pairs. Allophonic differences typically are often not as phonetically distinct as phonemic ones. In addition, the durations for high vowels are much shorter than for mid vowels, which could result in more coarticulation with neighboring segments.

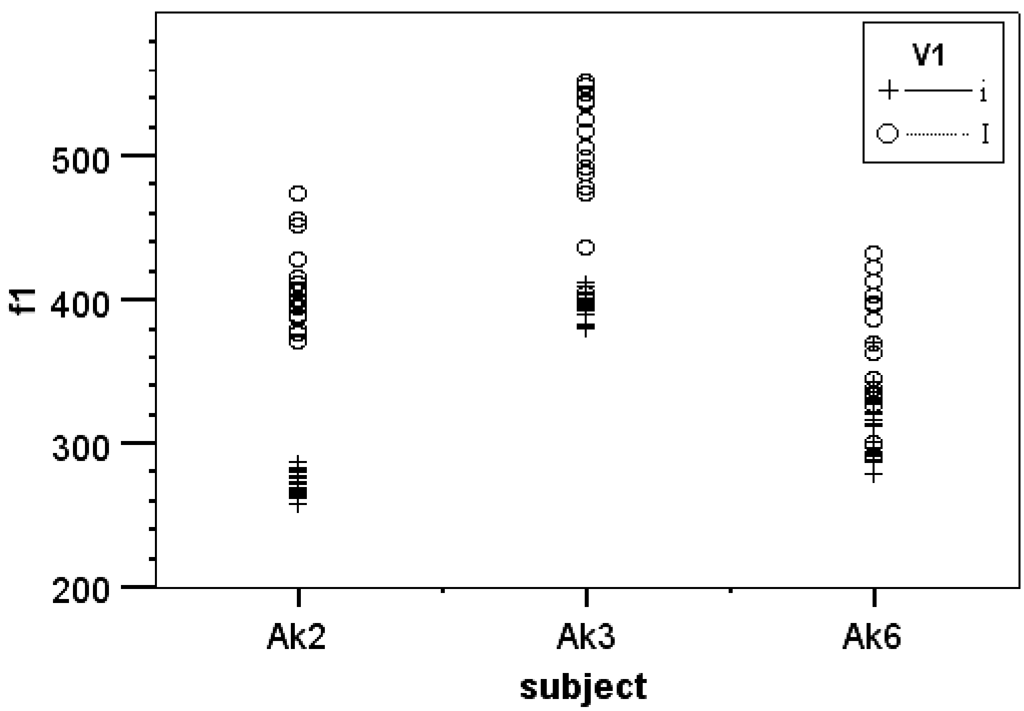


Figure 4.12. F1 values for midpoints of high front allophones [i], [ị] (i, I, in graph) for Àkùrẹ́ speakers for V1 of VCV tokens.

Similarly the allophonic high back pair, u-ụ (u-U), shows a highly significant difference in F1 for all three Àkùrẹ́ speakers. For Ak2, the F1 difference was 91 Hz [F(1,40)=106, p < .001]; for Ak3, the F1 difference was 54 Hz [F(1,37)=45, p < .001]; and for Ak6, the F1 difference was 76 Hz [F(1,40)=47, p < .001]. However, there is a large degree of overlap in F1 as shown in Figure 4.13. This figure does not resemble the corresponding figure for the contrasting mid vowels, shown above, presumably due to the allophonic nature of these vowels.

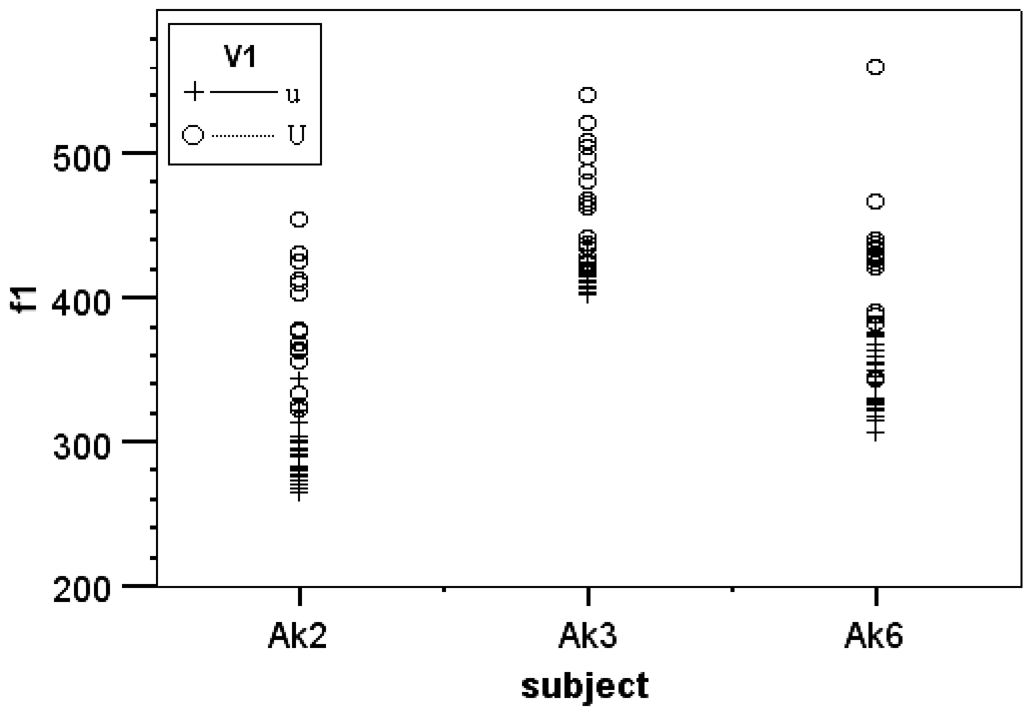


Figure 4.13. F1 values for midpoints of mid vowels [u, ụ] (u, U, in graph) for Àkùrẹ́ speakers for V1 of VCV tokens.

So for both i/I and u/U we have seen that while the allophone pairs have fairly large, and highly significant differences in the mean, they exhibit some overlap, except for Ak2 and Ak3 for /i/. That is, the target acoustic space for the two sets of vowels is not the same, but it isn't far enough apart to produce acoustically distinct outputs. Figure 4.14 shows the means for all labels for each Àkùrẹ́ speaker. For Ak2 and Ak3, the difference is large and maintained throughout the vowel. As seen in previous charts, considerable overlap is evident for speaker Ak6.

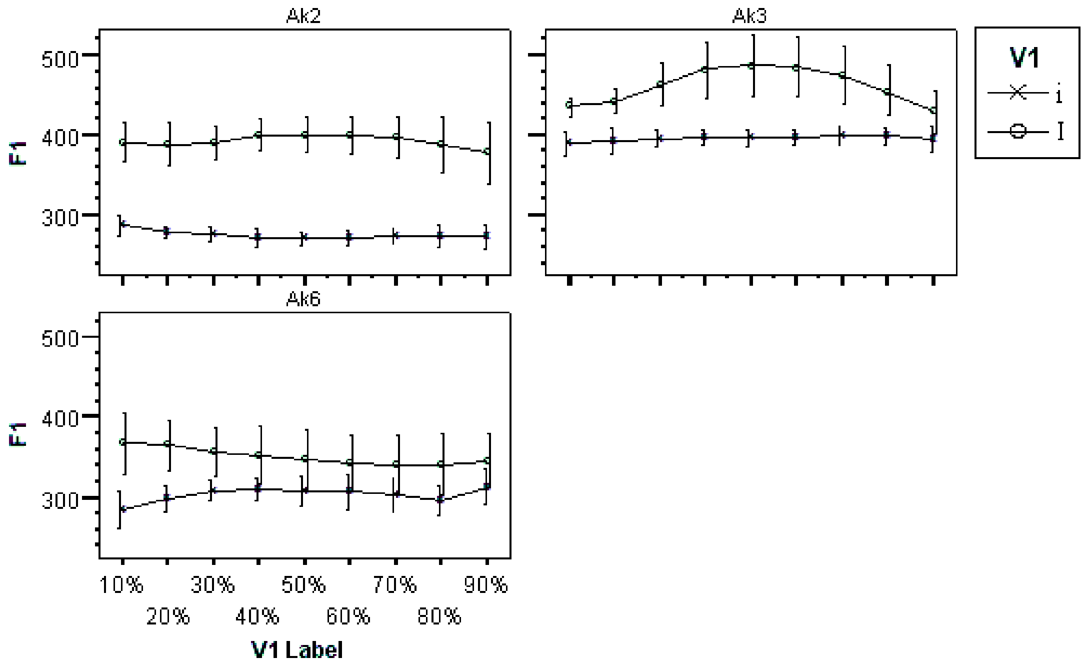


Figure 4.14. F1 means for all labels of high vowels [i, ị] (i, I, in graph) for Àkùrẹ́ speakers for V1 of VCV tokens for mid vowel V2s. The x's represent F1 means for tokens of [+ATR] [i]; the circles (○) represent the same for [−ATR] [ị].

4.2.3. Àkùrẹ́ coarticulation in [a]

For /a/ in Àkùrẹ́, and for all Yorùbá dialects, there is no claim of allophony. That is, there are not different allophones of /a/ in different phonological contexts. Nor do different allophones of /a/ trigger different phonological patterns. In this way, the Àkùrẹ́ /a/ is unlike /i/ and /u/ in Àkùrẹ́, but Àkùrẹ́ /a/ is like /i/, /u/, and /a/ in Mọ̀bà and SY. I examine here the tokens of /a/ in the same environments as seen for /i/ and /u/ above, that is, in V1 followed by [+ATR] vowels and [−ATR] vowels. Figure 4.15 shows the mean values for Àkùrẹ́ subjects for [a] in V1 position with the following vowels either [+ATR] or [−ATR].

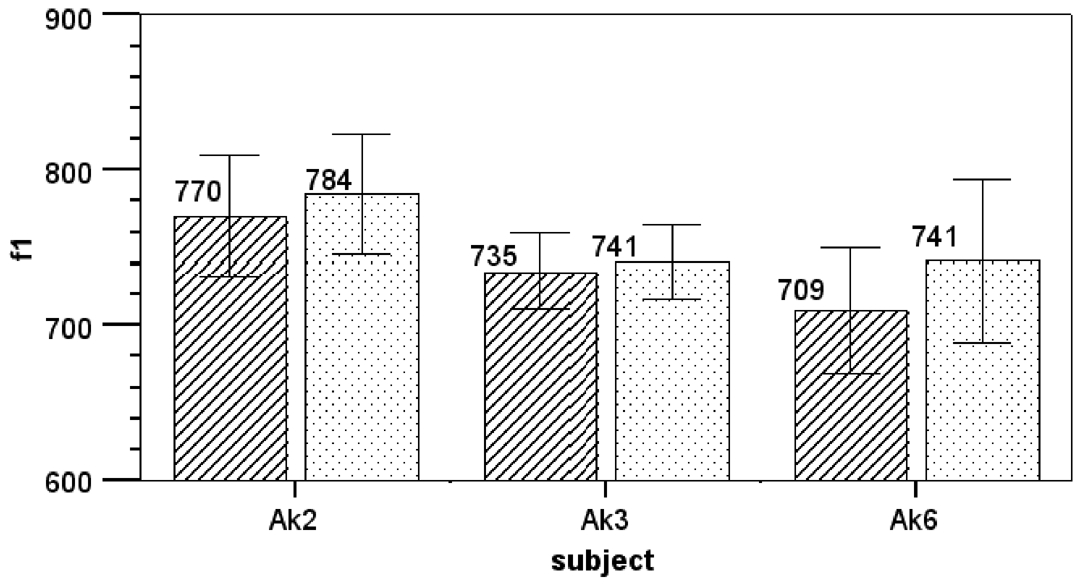


Figure 4.15. F1 mean values for midpoints of low vowels [a] for Àkùrẹ́ speakers for V1 of VCV tokens. The striped bars show [a] before [+ATR] vowels ([e, o, i, u]; the dotted bars show [a] before [−ATR] vowels [a, ẹ, ọ].

A two-way ANOVA with subject and ATR of V2 as factors shows that for all Àkùrẹ́ speakers combined, the F1 difference is not significant for the factor ATR of V2 [$F(1,71) = 2.98$, $p = .09$]. The differences between means for each subject are not significant, although there is a tendency for [a] followed by [−ATR] to have a higher F1 than [a] followed by [+ATR] vowels.

This in itself is not remarkable for these three subjects of Àkùrẹ́; however the tendency is evident throughout most of the speakers of this study as seen in Figure 4.16. In this figure, however, I have limited the V2 to mid vowels, in order to show the influence of the [ATR] of the V2, rather than the influence of height. A two-way ANOVA with subject and ATR of V2 as factors shows that for the eleven speakers for all three dialects, the F1 difference is significant for the factor ATR of V2 [$F(1,311) = 6.08$, $p = .014$]. This does not mean that this difference is present and significant for all three dialects. Further testing with more subjects or tokens could provide a more reliable finding. However, it does suggest that this coarticulatory pattern may be present in one or all dialects.

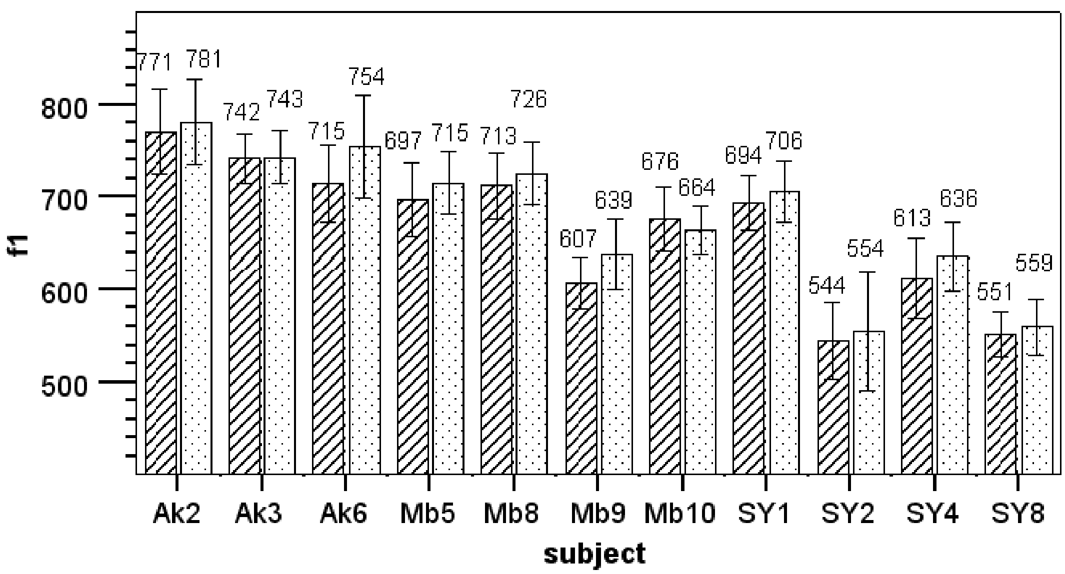


Figure 4.16. F1 mean values for midpoints of low vowels [a] for the speakers of all three dialects for V1 of VCV tokens with mid vowel V2s. The striped bars show [a] before [+ATR] mid vowels ([e, o]; the dotted bars show [a] before [−ATR] mid vowels [ẹ, ọ].

Returning to Àkùrẹ́ speakers, when examining the midpoint values for individual tokens, it is clear that no obvious pattern emerges from vowels of each set. Although there evidence of coarticulation, it would presumably be difficult for any learner to generalize a pattern.

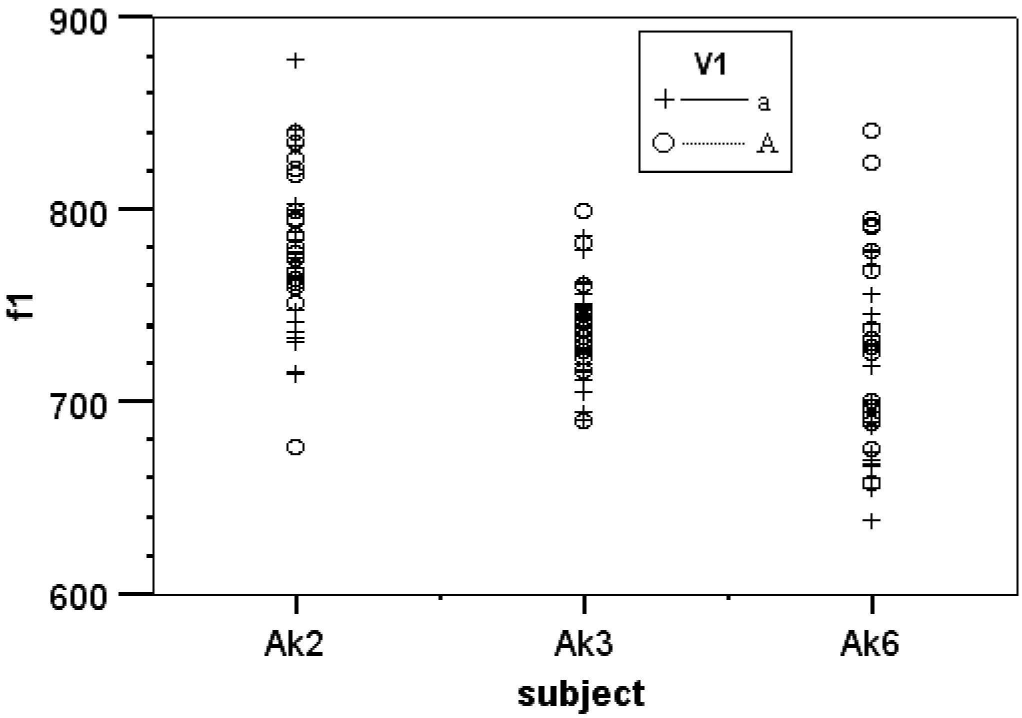


Figure 4.17. F1 values for midpoints of low vowels [a] for Àkùrẹ́ speakers for V1 of VCV tokens with '+' indicating a [+ATR] V2 and '○' indicating a [−ATR] V2.

When examining the F1 values throughout the vowel, in Figure 4.18, the coarticulation looks more pronounced. But the standard deviation and the previous figure indicates that there is considerable overlap at all points in the vowel.

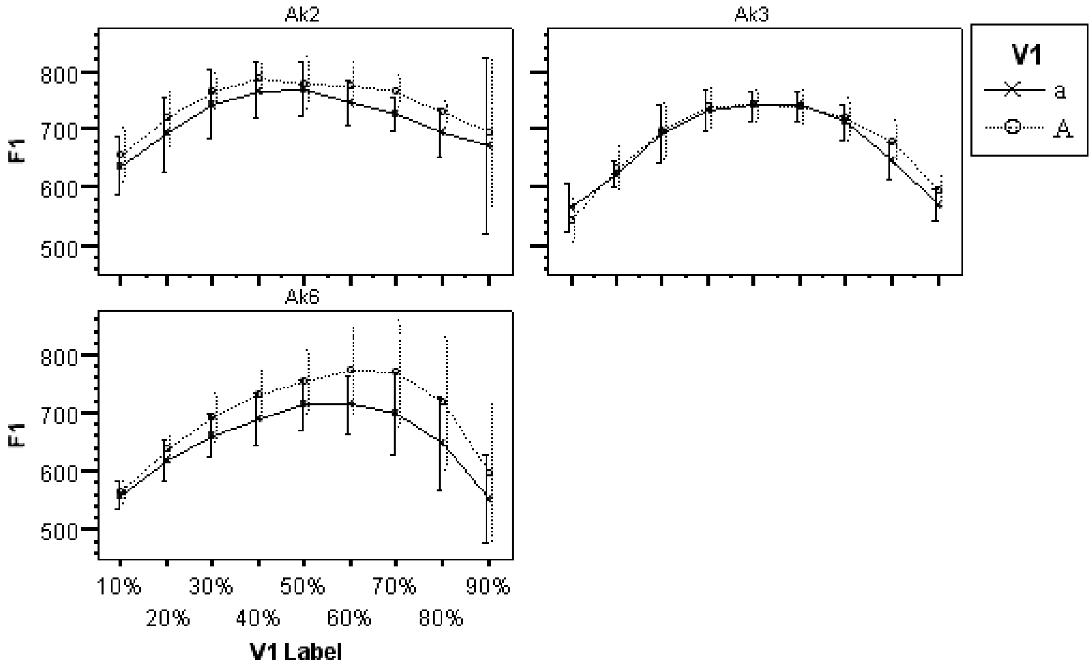


Figure 4.18. F1 means for all labels of low vowels [a] before [+ATR] mid vowels (labeled 'a' in graph) and before [−ATR] mid vowels (labeled 'A' in graph) for Àkùrẹ́ speakers for V1 of VCV tokens for mid vowel V2s.

4.2.4. Summary of Àkùrẹ́ vowel to vowel interaction

An examination of Àkùrẹ́ vowel interaction has shown a range of patterns. The phonemes, mid vowels, show highly significant differences and no overlap between tokens belonging to different phonemes. The low vowel /a/, which exhibits no allophony, does show evidence of coarticulation especially later in the vowel, such that [a] before [+ATR] vowels has a slightly lower F1 than [a] before [−ATR] vowels. This pattern for [a] is evident across speakers of all dialects. The allophonic pairs [i]/[ị] and [u]/[ụ] show highly significant differences in F1 for all Àkùrẹ́ speakers. However, while two of three subjects show no overlap for [i]/[ị], one speaker exhibits some overlap for [i]/[ị], and all speakers exhibit overlap for [u]/[ụ]. Thus, the behavior of allophonic pairs falls between that of phonemes and coarticulation.

4.2.5. Characteristics of Àkùrẹ́ harmony

The results of this part of the acoustic study corroborate the phonological evidence presented in Chapter 2, allowing us to specify the following characteristics of the Àkùrẹ́ high vowel harmony in (2), which will be compared with characteristics of coarticulation in Mọ̀bà and SY below.

(2) Characteristics of Àkùrẹ́ harmony

- targets — the targets affected by vowel harmony:
  - /i/ and /u/
- contexts — the contexts in which the vowel harmony occurs:
  - when the target is followed by a vowel [ẹ], [ọ], [a], [i̞], [ụ]
  - with an intervening consonant
- effect — the acoustic effect realized in the harmony:
  - raising of F1

4.3. Mọ̀bà

The picture for Mọ̀bà's vowels is simpler than that of Àkùrẹ́'s. Because Mọ̀bà has no allophonic variation, and notably none in the high vowels, these charts show seven vowels instead of the nine shown for Àkùrẹ́. The individual vowel charts for the four Mọ̀bà speakers are shown in Figures 4.19-4.22, with the four vowel charts shown together in Figure 4.23.

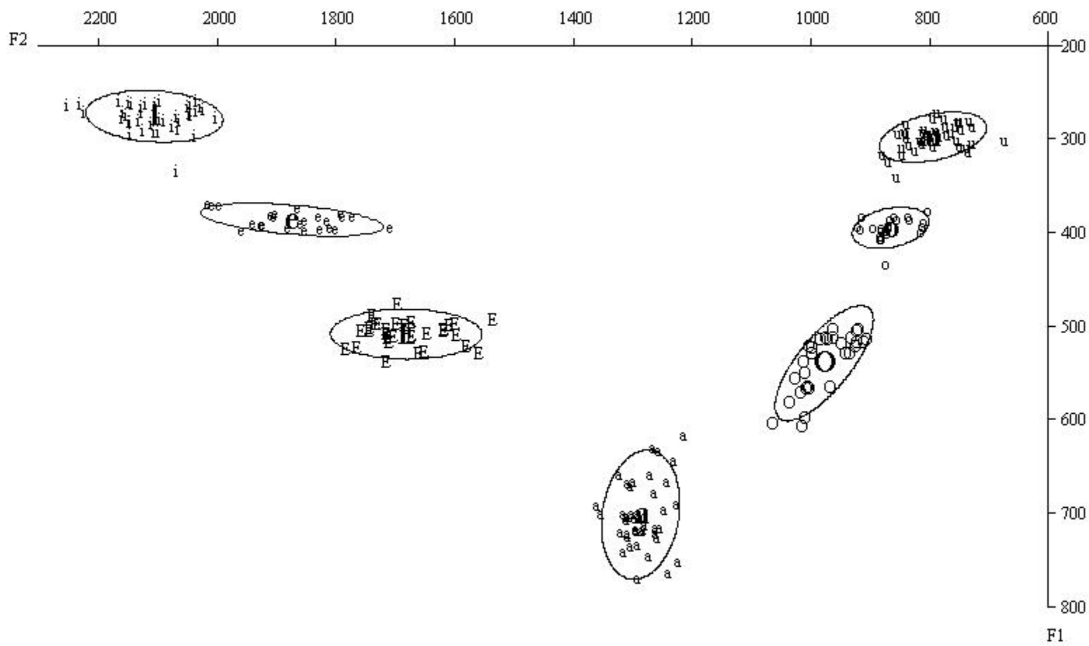


Figure 4.19. Vowel Chart (F1 versus F2) of Mọ̀bà Speaker Mb5 for the midpoint of V_1 in V_1CV_2. The smaller letters indicate the values for individual tokens; the large letters show the mean for the vowel, and the ellipses show one standard deviation from the mean for each vowel. The [+ATR] vowel are [i, e, o, u]; the [−ATR] vowels are [E, O, a].

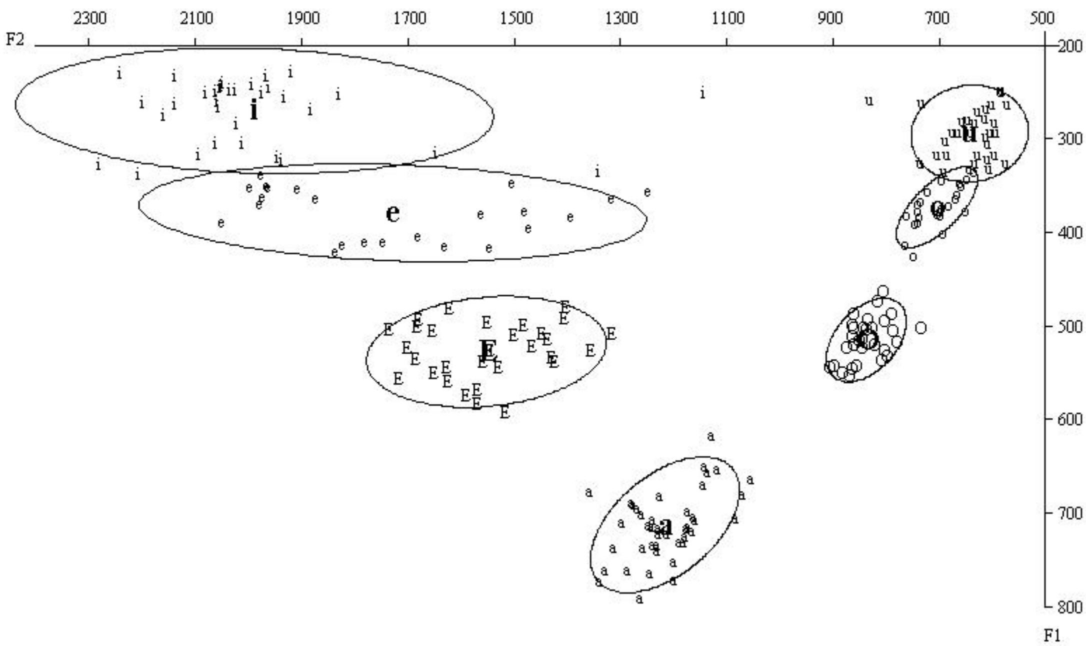


Figure 4.20. Vowel Chart (F1 versus F2) of Mọ̀bà Speaker Mb8 for the midpoint of V_1 in V_1CV_2. The smaller letters indicate the values for individual tokens; the large letters show the mean for the vowel, and the ellipses show one standard deviation from the mean for each vowel. The [+ATR] vowel are [i, e, o, u]; the [−ATR] vowels are [E, O, a].

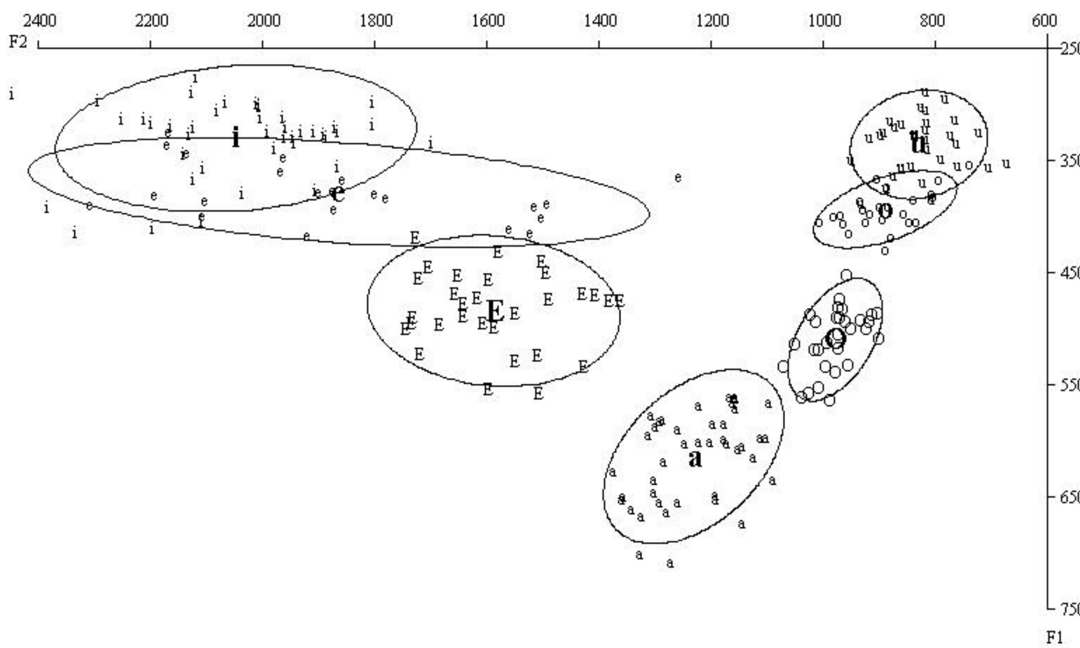


Figure 4.21. Vowel Chart (F1 versus F2) of Mọ̀bà Speaker Mb9 for the midpoint of V_1 in V_1CV_2. The smaller letters indicate the values for individual tokens; the large letters show the mean for the vowel, and the ellipses show one standard deviation from the mean for each vowel. The [+ATR] vowel are [i, e, o, u]; the [−ATR] vowels are [E, O, a].

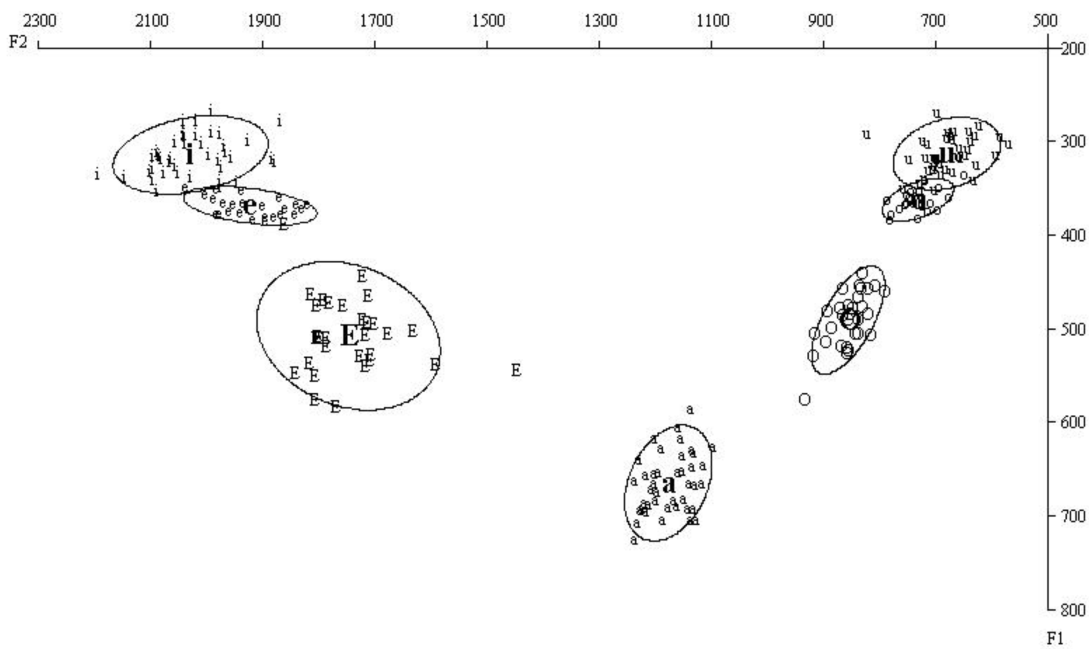


Figure 4.22. Vowel Chart (F1 versus F2) of Mọ̀bà Speaker Mb10 for the midpoint of V_1 in V_1CV_2. The smaller letters indicate the values for individual tokens; the large letters show the mean for the vowel, and the ellipses show one standard deviation from the mean for each vowel. The [+ATR] vowel are [i, e, o, u]; the [−ATR] vowels are [E, O, a].

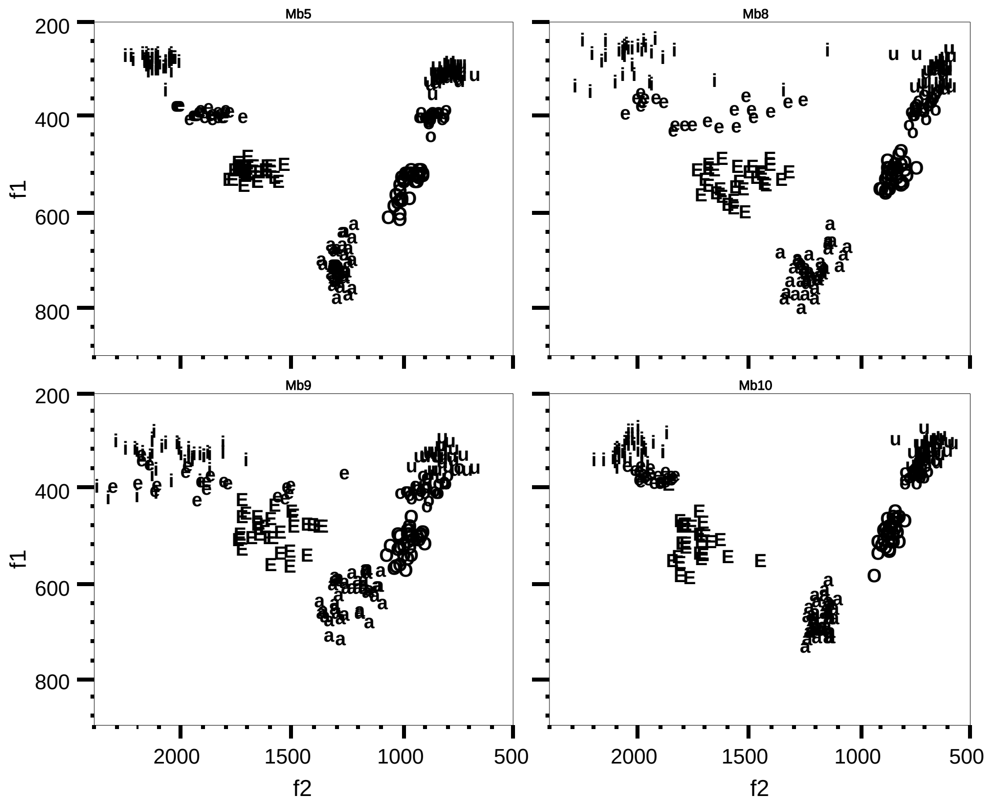


Figure 4.23. Vowel Charts (F1 versus F2) of Mọ̀bạ̀ Speakers Mb5, Mb8, Mb9, Mb10 for the midpoint V1 of VCV. [i, e, o, u] are the [+ATR] vowels; [E, O, a] are the [−ATR] vowels.

As with Àkùrẹ́, we see a large variation between the vowel charts of the four speakers. Mb5 and Mb10 show tight clusters of vowel tokens, with little overlap between vowels. The vowels of Mb8 and Mb9 are less tightly packed; they exhibit more variation especially in the F2 axis for [e] and [i]. (This is likely due to the rate of speaking; a look at vowel durations, in Figure 4.24, shows that Mb8 and Mb9 spoke much more quickly Mb5 and Mb10.)

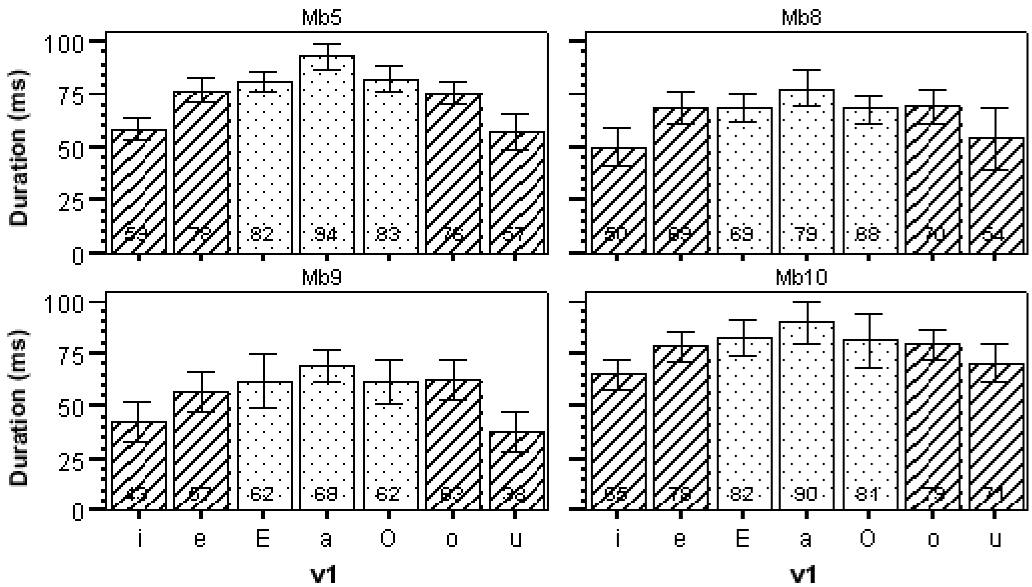


Figure 4.24. Mean V1 vowel durations for VCV words for four Mọ̀bà speakers. The error bars mark one standard deviation.

The F1 means for each Mọ̀bà speaker are shown in Figures 4.25-4.28.

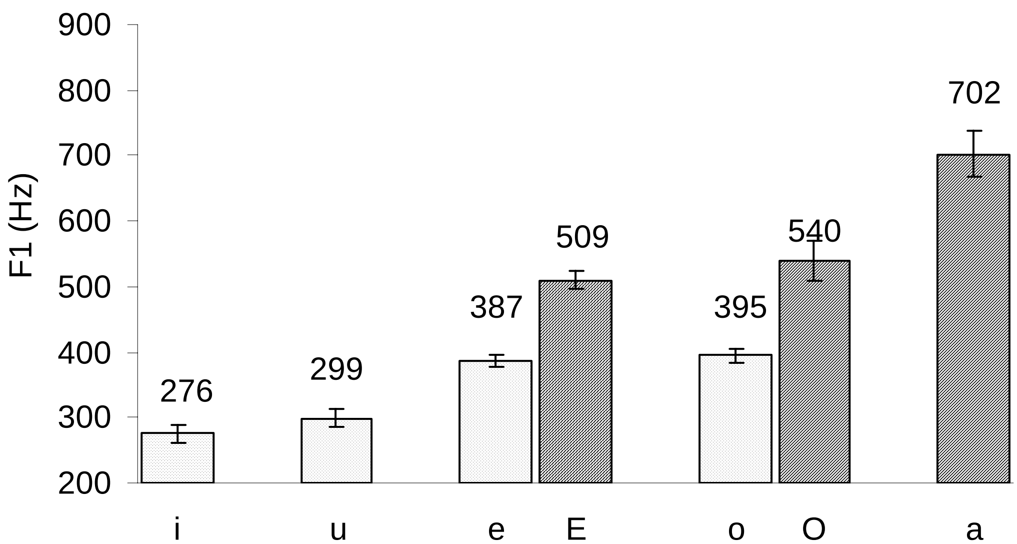


Figure 4.25. Mean F1 values at V1 midpoint for VCV words for Mọ̀bà speaker Mb5. The error bars mark one standard deviation.

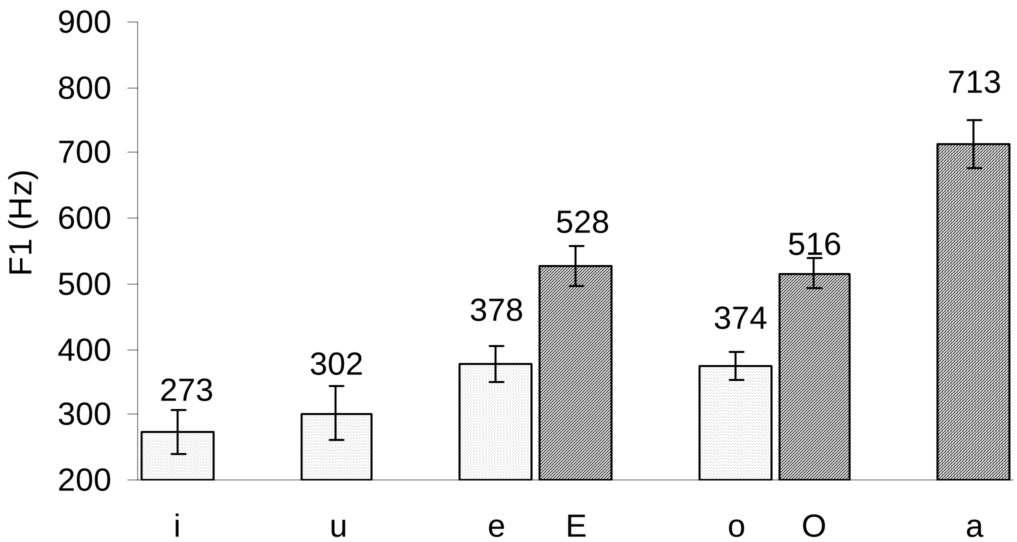


Figure 4.26. Mean F1 values at V1 midpoint for VCV words for Mọ̀bà speaker Mb8. The error bars mark one standard deviation.

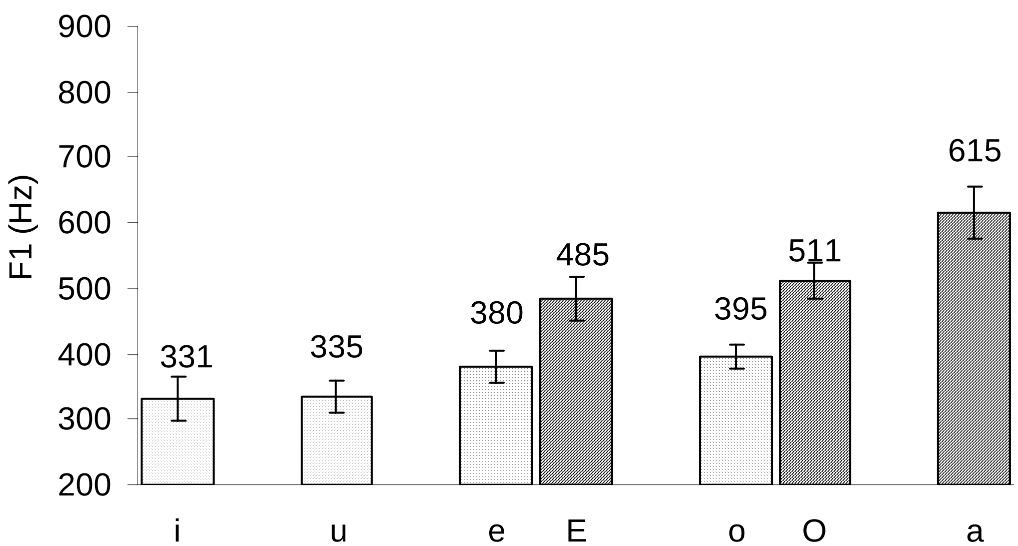


Figure 4.27. Mean F1 values at V1 midpoint for VCV words for Mọ̀bà speaker Mb9. The error bars mark one standard deviation.

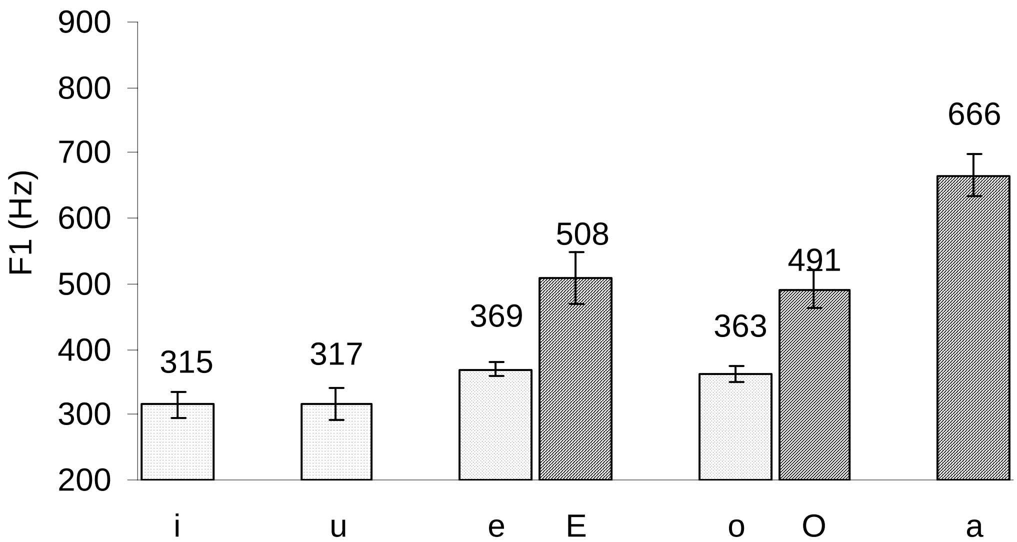


Figure 4.28. Mean F1 values at V1 midpoint for VCV words for Mọ̀bà speaker Mb10. The error bars mark one standard deviation.

4.3.1. Mọ̀bà—different phonemes, [−ATR] versus [+ATR] mid vowels

The ±ATR mid vowel pairs show clear differences— a single-factor ANOVA shows a highly significant difference for F1 for all subjects for both e/ẹ and o/ọ pairs. First the results for [e] versus [ẹ]: for Mb5, the F1 difference was 122 Hz [$F(1,70) = 1449$, $p < .001$]; for Mb8, the F1 difference was 150 Hz [$F(1,70) = 360$, $p < .001$]; for Mb9, the F1 difference was 105 Hz [$F(1,70) = 150$, $p < .001$]; and for Mb10, the F1 difference was 139 Hz [$F(1,70) = 270$, $p < .001$]. And for [o] versus [ọ]: for Mb5, the F1 difference was 145 Hz [$F(1,70) = 462$, $p < .001$]; for Mb8, the F1 difference was 142 Hz [$F(1,70) = 539$, $p < .001$]; for Mb9, the F1 difference was 116 Hz [$F(1,70) = 320$, $p < .001$]; and for Mb10, the F1 difference was 128 Hz [$F(1,70) = 404$, $p < .001$]. Additional perspective is gained by focusing on the individual F1 values for all Mọ̀bà speakers.

F1 values show no overlap between vowel tokens of [e] versus [ẹ] and [o] versus [ọ], but an examination of the individual F1 values shows that there is no large acoustic distance between [ẹ] and [e] vowels for speaker Mb9, as shown in Figures

4.29 and 4.30—the lowest F1 value for [ẹ] is close to the highest F1 value for [e], and similarly for [o] and [ọ]. However, if we examine closely the influence of the following vowel (V2) on V1 for speaker Mb9, in Figure 4.31, we see that at each V2 vowel height, there is a considerable F1 difference for the value of the V1. For example, while the F1 in [ẹ] of 'ẹki' nearly overlaps the F1 of the initial [e] of 'eke', it is not close to the F1 of the [e] of 'eki'. The same applies for the back V2s, with the [ẹ] of 'ẹku' not close to the [e] of 'eku'. So while there conceivably is confusion on the first vowel, this confusion would be resolved at the time of the second vowel.

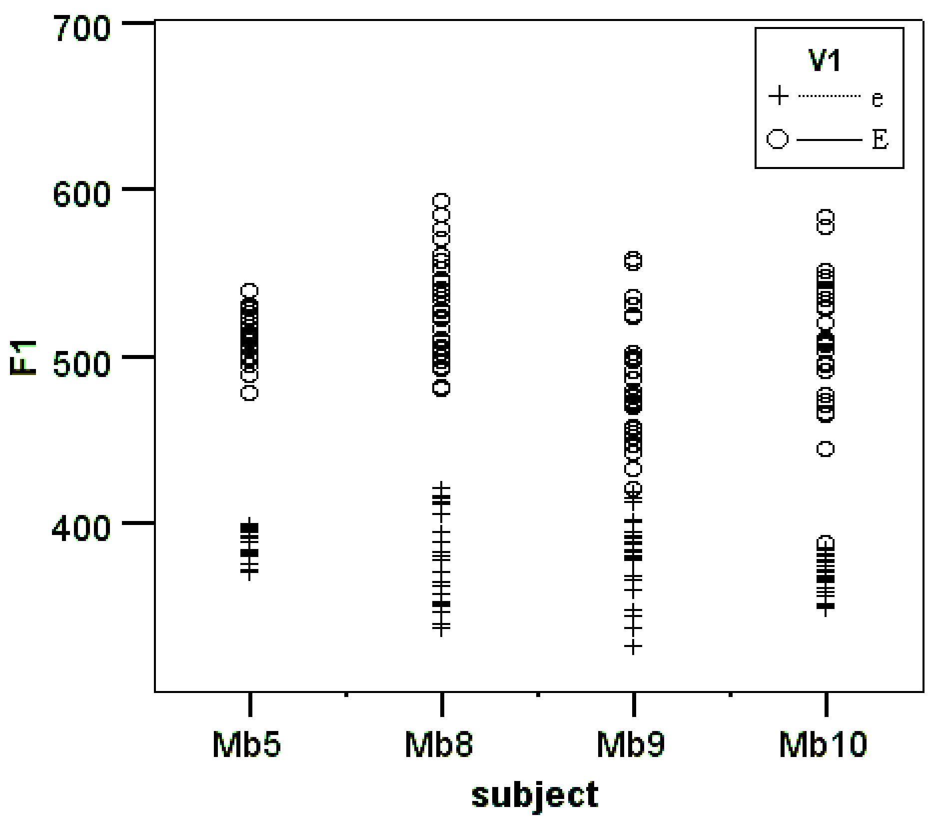


Figure 4.29. F1 values for midpoints of mid vowels [e, ẹ] (e, E, in graph) for Mọ̀bà speakers for V1 of VCV tokens.

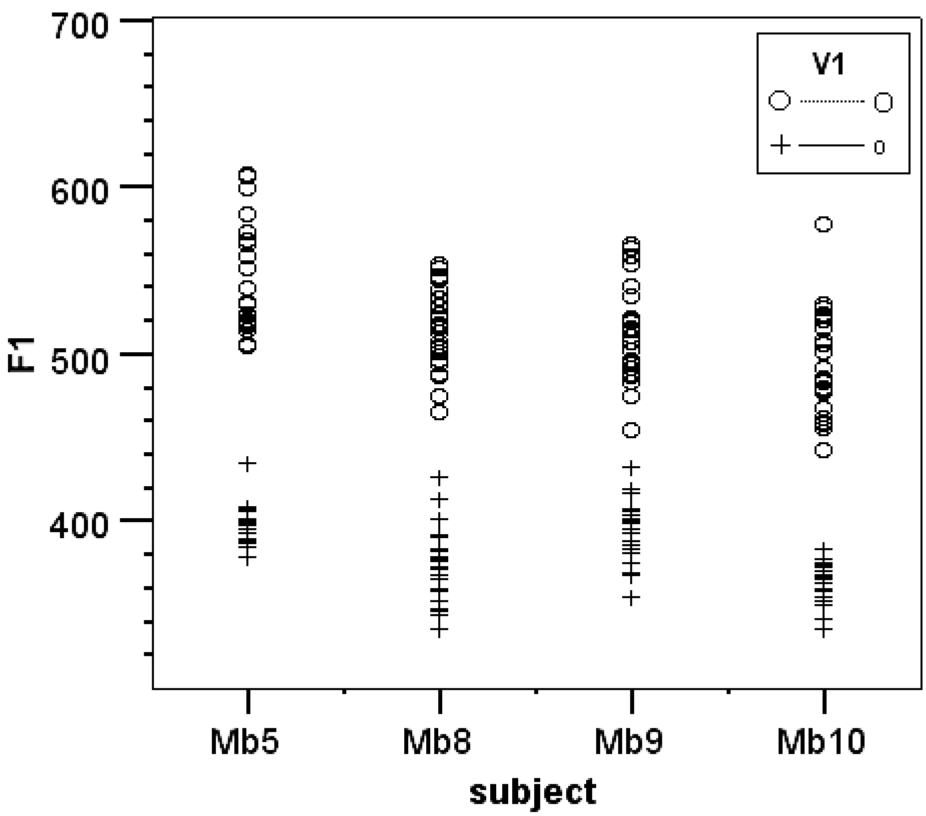


Figure 4.30. F1 values for midpoints of mid vowels [o, ọ] (o, O, in graph) for Mọ̀bà speakers for V1 of VCV tokens.

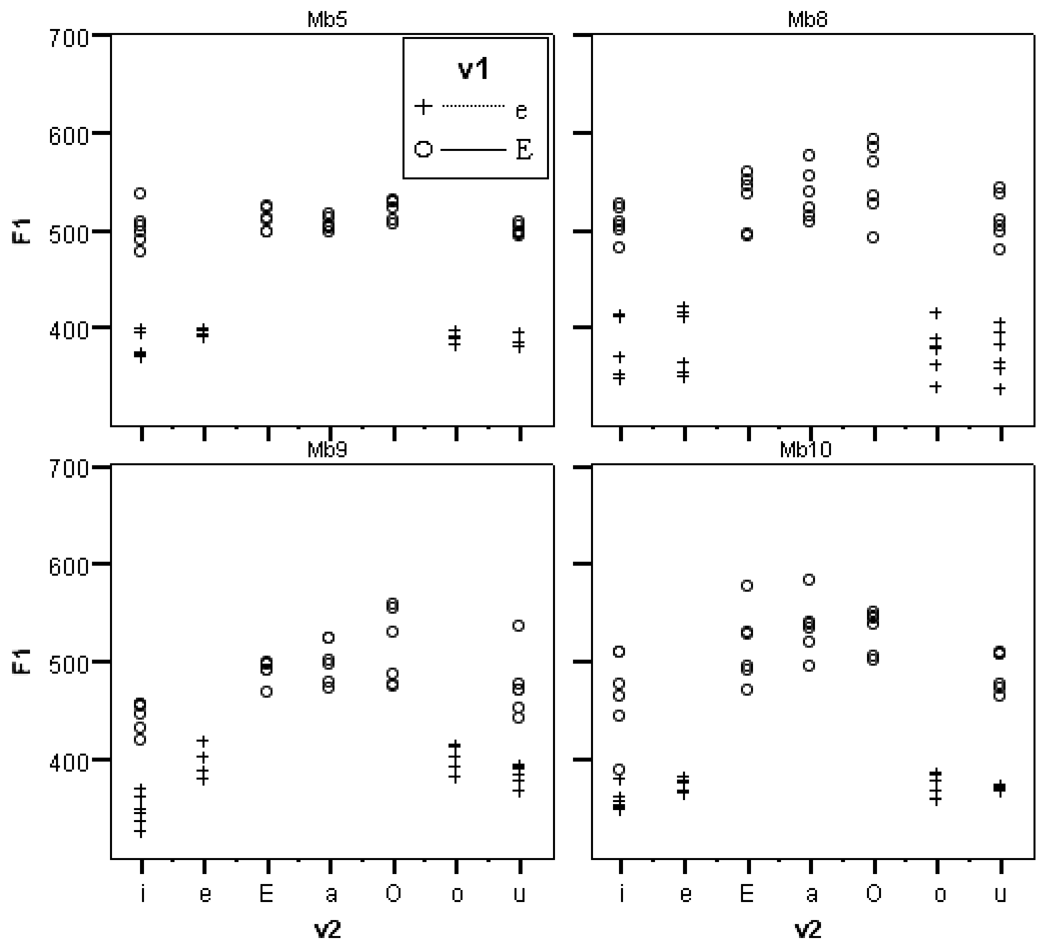


Figure 4.31. F1 values for midpoints of mid vowels [e, ẹ] (e, E, in graph) for Mọ̀bà speakers for V1 of VCV tokens.

Figure 4.32 shows the means for all labels for each Mọ̀bà speaker. As expected, the large difference between phonemes [e] and [ẹ] is maintained throughout the vowel. Similar results are found for [o] and [ọ], not shown.

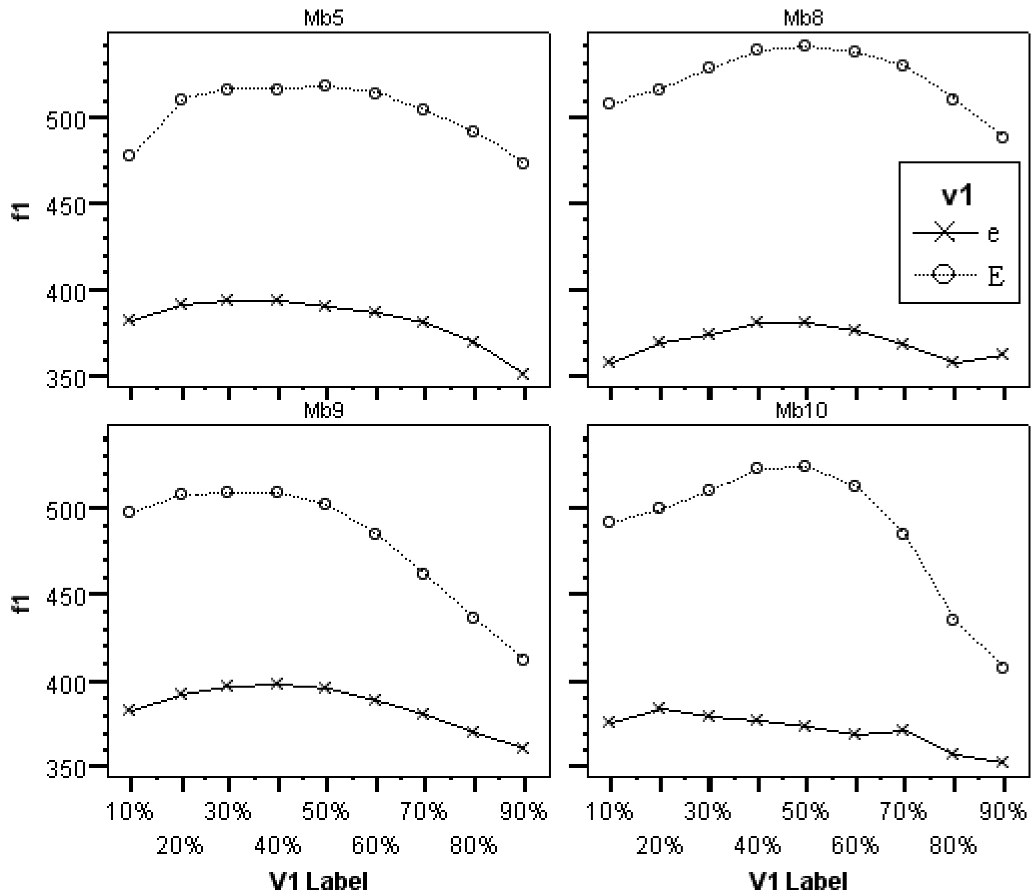


Figure 4.32. F1 means for all labels of mid vowels [e, ẹ] (e, E, in graph) for Mọ̀bà speakers for V1 of VCV tokens for mid vowelV2. The x's represent F1 means for tokens of [+ATR] [e] at the mid point of V1; the circles (○) represent the same for [−ATR] [ẹ] ('E' in graph).

4.3.2. Coarticulation in Mọ̀bà high vowels

We now look at how the tokens with [i] and [u] as V1 vary with differing vowels in V2; that is, we examine vowel to vowel coarticulation in Mọ̀bà high vowels, the same environment where we find harmony in Àkùrẹ́. We start by looking at the mean F1 values for [i] for four speakers of Mọ̀bà, in Figure 4.33. The dotted bars for each speaker indicate the mean F1 for [i] before the [+ATR] vowels (/i/,/u/,/e/, /o/); the striped bars indicate the mean F1 for [i] before [+ATR] vowels ([ẹ], [ọ], [a]).

Recall that this division does not correspond with an allophonic split—both of the cases are considered part of the phoneme /i/.

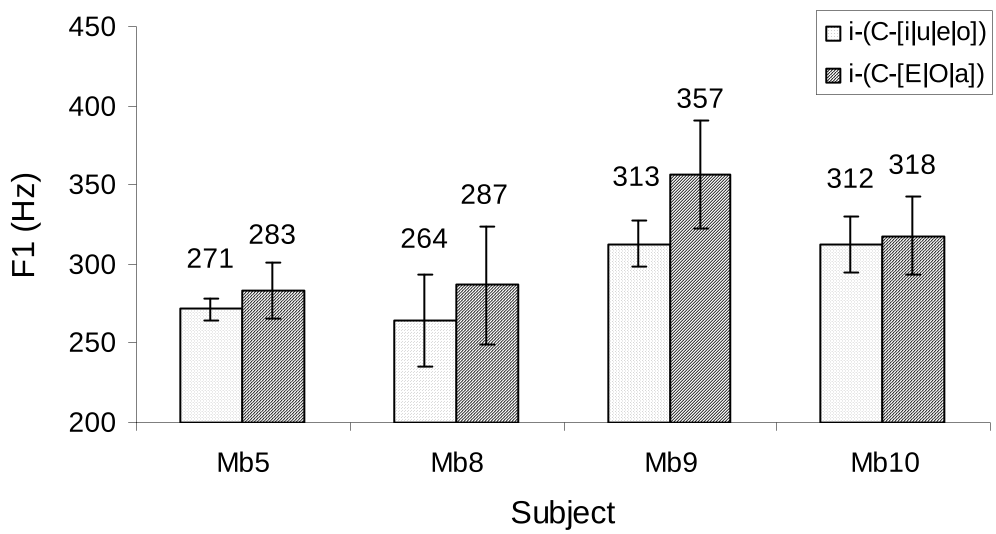


Figure 4.33. F1 means for midpoints of V1 [i] in iCV when followed by [+ATR] V2s (dotted bars) and by [−ATR] V2s (striped bars) for all Mọ̀bà speakers.

For each speaker, the average F1 is higher for [i]s before [−ATR] vowels. The direction here is the same as that found for F1 differences in Àkùrẹ́ harmony. For the Mọ̀bà speakers, the differences are highly significant for Mb9 [$F(1,40) = 32$, $p < .001$], significant for Mb5 [$F(1,40) = 8.9$, $p < .01$], slightly significant for Mb8 [$F(1,38) = 4.6$, $p < .05$], and not significant for Mb10 [$F(1,40) = .70$, $p = .408$]. When taken collectively, the F1 difference was highly significant [$F(1,164) = 14.8$, $p < .001$]. While the differences resemble in direction the Àkùrẹ́ data, one could suggest that the influence of the [low] and [high] vowels might be skewing the averages of the two groups in opposite directions. If it is only the peripheral vowels that contribute to the difference between the two groups, then it would not be fair to attribute the difference to [ATR]. For this reason, I include Figure 4.34, which shows the same as the previous figure, except that instead of two groups, we divide the [i] into groups depending on whether V2 is [high], [[+ATR] mid], [[−ATR] mid], or [low]. In this way, we can determine to what extent the low and high vowels contribute to the difference between the two groups in the previous figure.

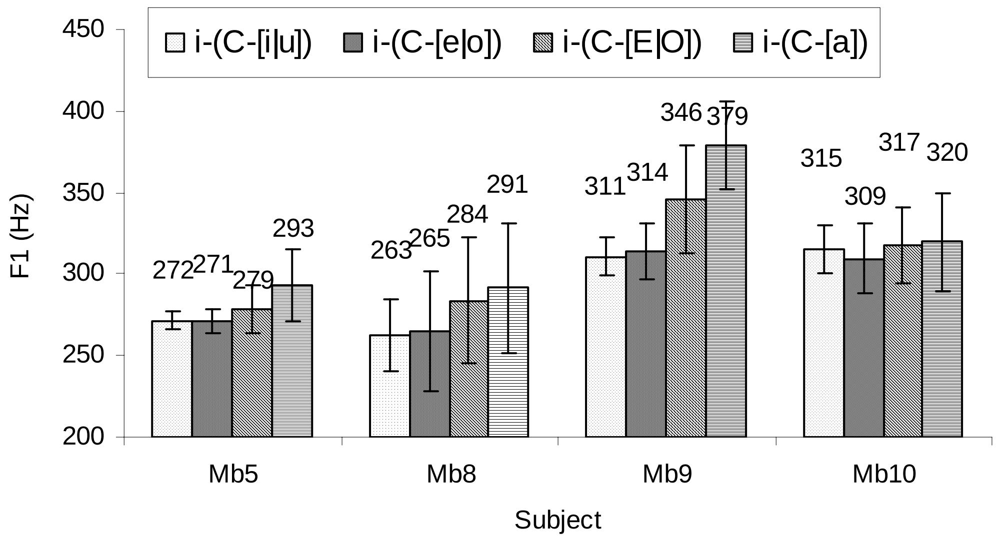


Figure 4.34. F1 means for midpoints of V1 [i] in iCV$_2$ for all Mọ̀bà speakers, where V$_2$ is [high], [mid +ATR], [mid −ATR], and [low], respectively, for the four bars for each speaker.

Indeed, [i] before [a] has a higher F1 than the other [−ATR] mid vowels, [ẹ] and [ọ], especially for Mb9 and Mb5. The [i] before high vowels, on the other hand, is about the same as the [i] before the other [+ATR] mid vowels, [e] and [o]. Looking at only the mid vowels as V2, that is, the middle two bars for each speaker, we see that the F1 differences are all in the same direction as predicted; however, the differences are not nearly as robust as we found for Figure 4.33, with only Mb9 having a statistically significant difference [$F(1,22) = 8.6$, $p < .01$]. When taken collectively, the F1 difference for all four speakers is slightly significant [$F(1,92) = 5.0$, $p < .05$]. When we consider the intervocalic consonant, however, the results are more interesting. In the next two figures, we see the same graph for tokens where the consonant is [k] and [b], respectively. For intervocalic consonant [k], in Figure 4.35, we see that [i] before [−ATR] vowels has a higher F1 than [i] before [+ATR] vowels. Although the number of tokens in this subset is small, the differences are significant for Mb5 [$F(1,10) = 20$, $p < .01$], and Mb9 [$F(1,10) = 14$, $p < .01$], though not for Mb8 [$F(1,9) = 1.0$, $p = .346$] and Mb10 [$F(1,10) = 4.7$, $p = .06$]. A two-way ANOVA with subject and ATR of V2 as factors shows that for all speakers combined, the F1 difference is highly significant for the factor ATR of V2 [$F(1,46) = 19$, $p < .001$].

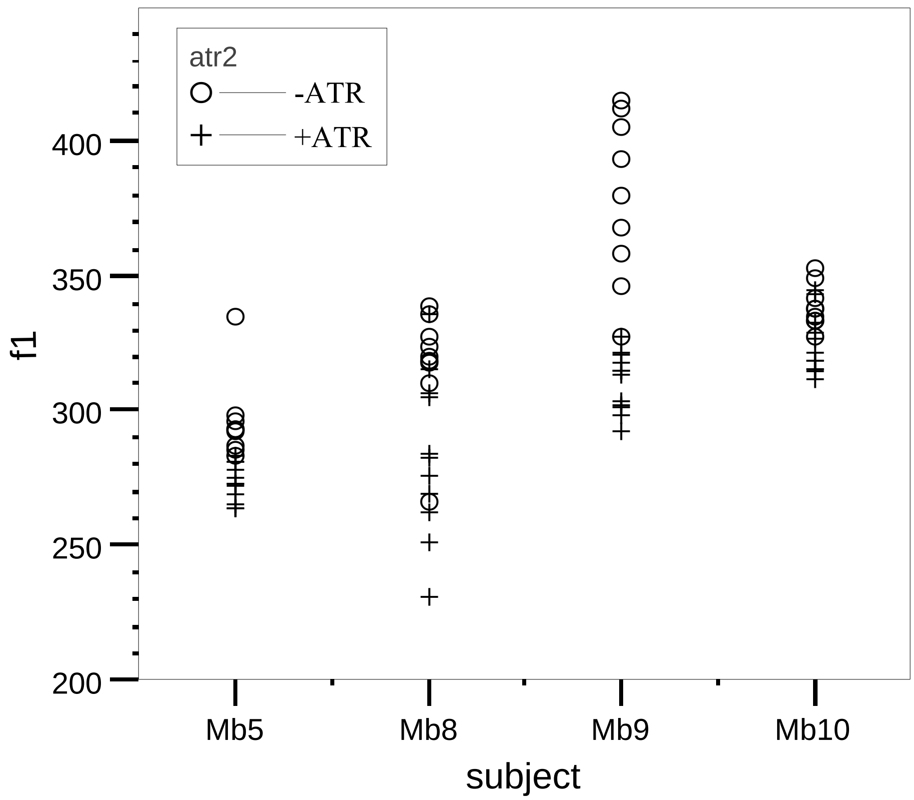


Figure 4.35. F1 values for midpoints of [i] in i-k-V_2, where '+' indicates a V2 of [+ATR] mid vowels [e, o], and 'O', [−ATR] mid vowels [ẹ, ọ], for four Mọ̀bà speakers.

For consonant [b], the difference is not as robust, as shown in Figure 4.36. The differences are not significant for any speaker: Mb5 [$F(1,10) = .4$, $p = .53$], Mb8 [$F(1,9) = 4.5$, $p = .06$], Mb9 [$F(1,10) = 1.7$, $p = .23$] and Mb10 [$F(1,10) = .12$, $p = .74$]. A two-way ANOVA with subject and ATR of V2 as factors shows that for all speakers combined, the F1 difference is not significant for the factor ATR of V2 [F(1,46) = 2.3, p = .134].

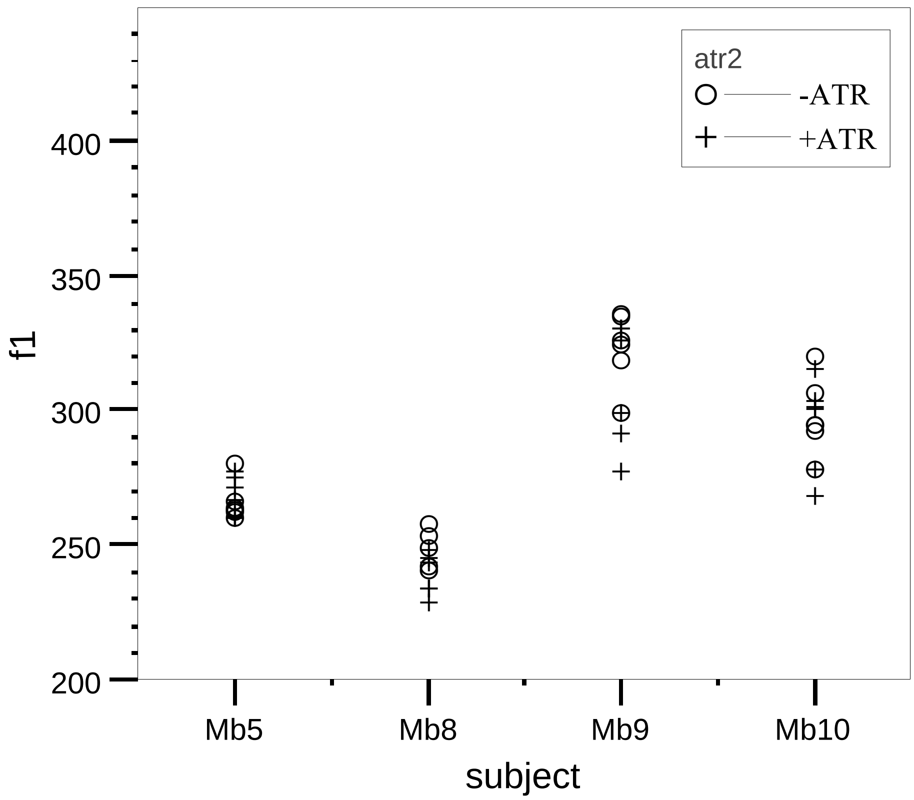


Figure 4.36. F1 values for midpoints of [i] in V-b-V_2, where '+' indicates a V2 of [+ATR mid] vowels [e, o], and 'O' indicates a V2 of [−ATR mid] vowels [ẹ, ọ] for four Mọ̀bà speakers.

Moving to the high vowel [u], Figure 4.37 shows the F1 means of [u] before [+ATR] vowels and before [−ATR] vowels for the four Mọ̀bà speakers. The difference here is not as robust as those found for [i] above.

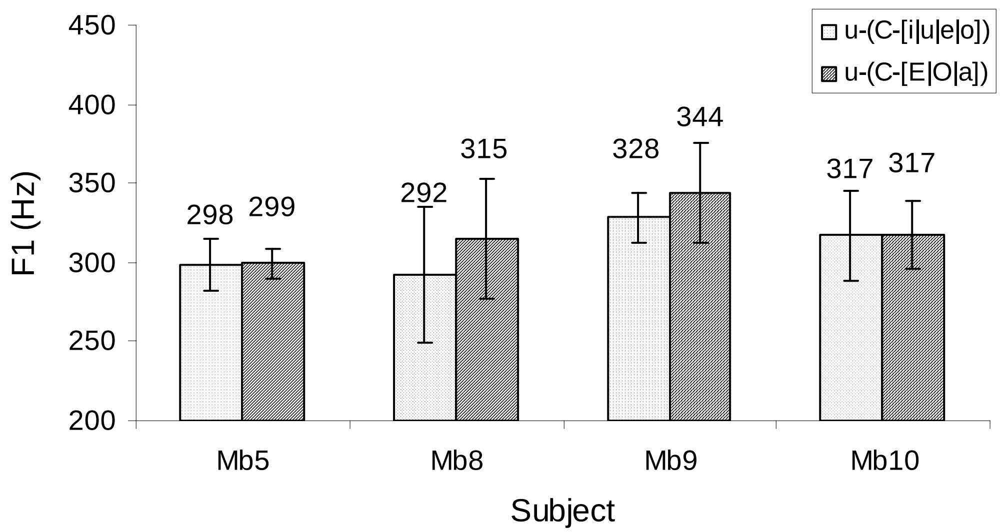


Figure 4.37. F1 means for midpoints of V1 [u] in uCV when followed by [+ATR] V2s (dotted bars) and by [−ATR] V2s (striped bars) for all Mọ̀bà speakers.

Splitting the V2 vowels into four groups according to V2 height and [ATR] in Figure 4.38, we can see that the pattern here is again not as pronounced as for [i]. For each speaker, the difference of the mean F1 of [u] before [+ATR] mid vowels [e] and [o] versus before [−ATR] mid vowels [ẹ] and [ọ] is not statistically significant.

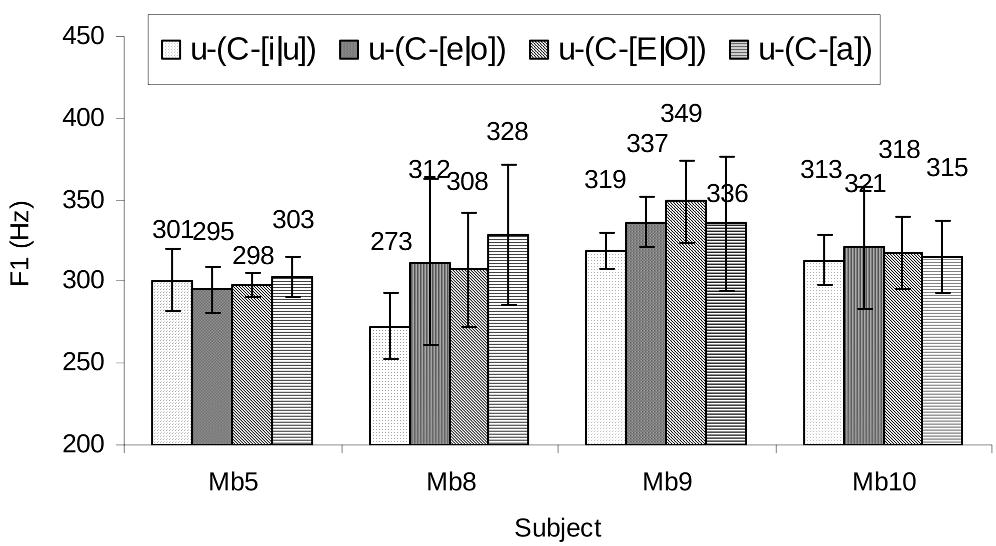


Figure 4.38. F1 means for midpoints of V1 [u] in uCV_2 for all Mọ̀bà speakers, where V_2 is [high], [mid +ATR], [mid −ATR], and [low], respectively, for the four bars for each speaker.

Likewise when the groups are divided by intervocalic consonant, as in Figures 4.39 and 4.40, the differences are not statistically significant.

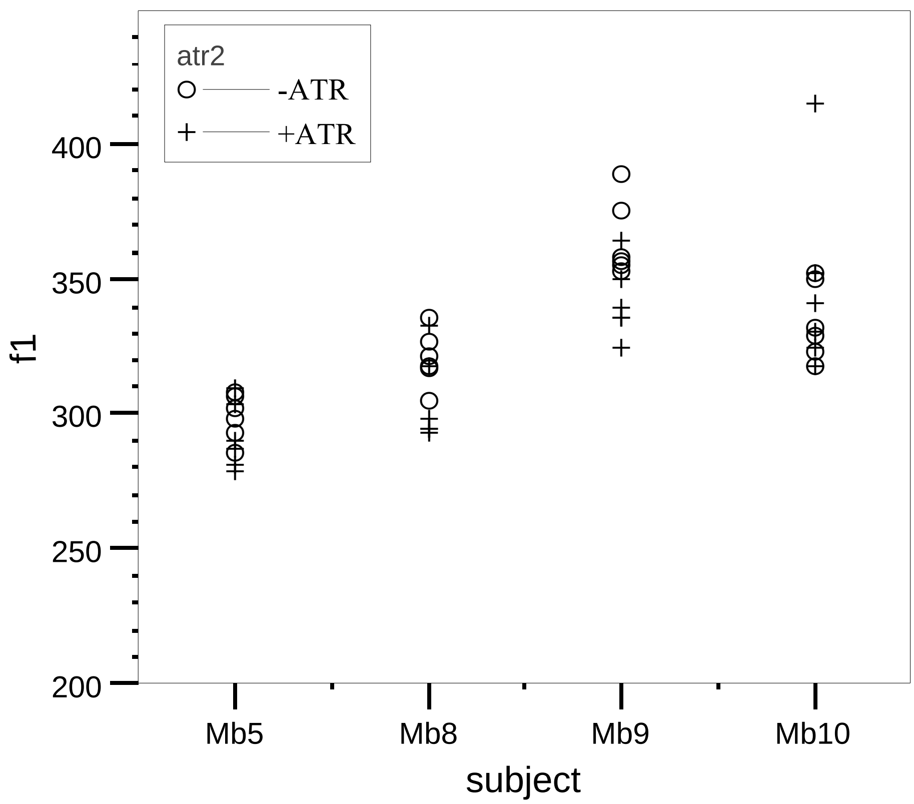


Figure 4.39. F1 values for midpoints of [u] in u-k-V_2, where '+' indicates a V2 of [+ATR] mid vowels [e, o], and 'O', [−ATR] mid vowels [ẹ, ọ], for four Mọ̀bà speakers.

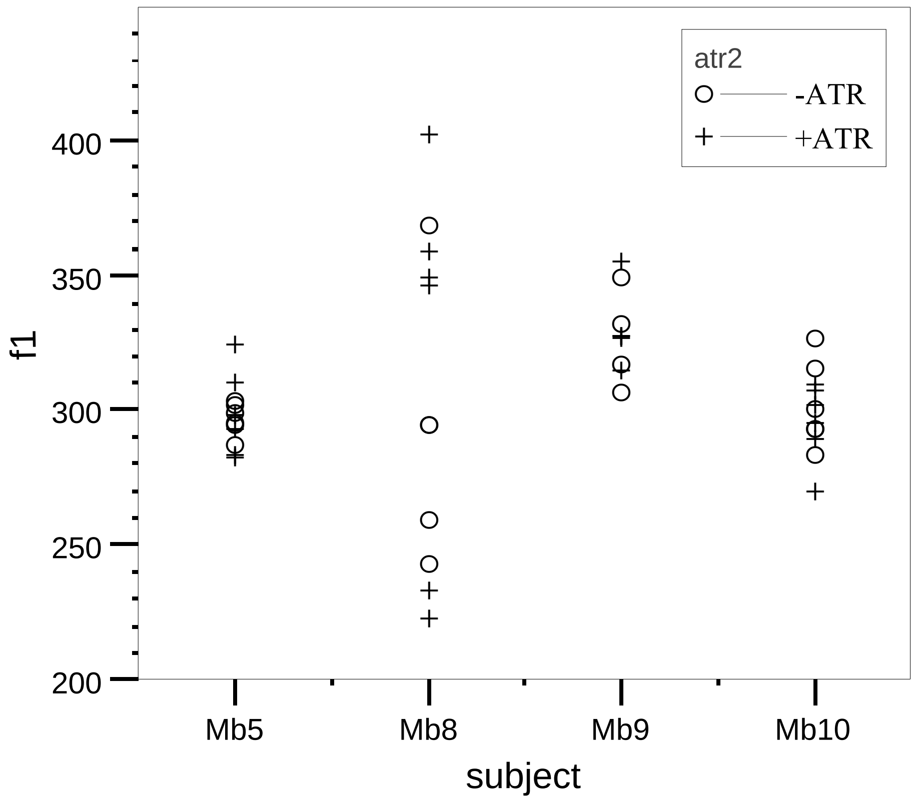


Figure 4.40. F1 values for midpoints of [u] in u-b-V_2, where '+' indicates a V2 of [+ATR] mid vowels [e, o], and 'O', [−ATR] mid vowels [ẹ, ọ], for four Mọ̀bà speakers.

In summary, Mọ̀bà [i] exhibits some coarticulation resembling the patterns of [i] in Àkùrẹ́ vowel harmony, with [i] having a raised F1 when followed by [−ATR] vowels when the intervocalic consonant is [k]. The same pattern is exhibited with consonant [b], but the results are not statistically significant. For [u], the pattern for this data is less robust, if it exists at all.

Figure 4.41 shows the means for all labels for each Mọ̀bà speaker for /i/ followed by mid vowels. All speakers show a difference between the two sets throughout the vowel. For Mb8 and Mb9—the fastest of the four speakers—the difference is quite large.

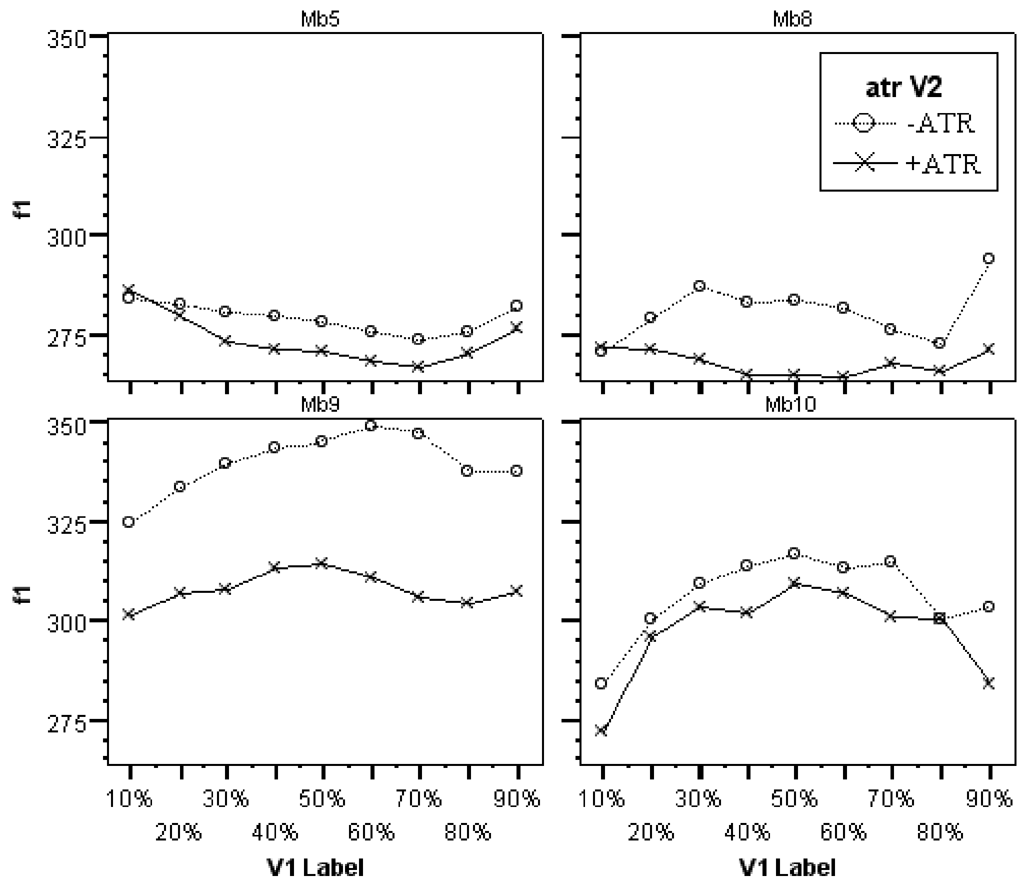


Figure 4.41. F1 mean values at various points throughout V1 for iCeo versus iCEO for four Mọ̀bà speakers.

4.3.2.1. Other intervocalic consonants

In light of the different influence of the two intervocalic consonants used in the main experiment, I investigated how other consonants interact with vowels. Do other consonants show coarticulation patterns resembling harmony, like [k], or are they more like [b] in that the patterns are not as evident? In a separate study with Mọ̀bà speaker Mb10, I examined the effect of several intervocalic consonants in VCV tokens, where V1 was /i/, V2 was /e/ or /E/, and C was /d, g, gb, k, t/. Otherwise the methodology was the same as for the previous experiment. The coarticulation pattern reported so far for [k] above, is seen quite clearly for most of the consonants, as shown in Figure 4.42.

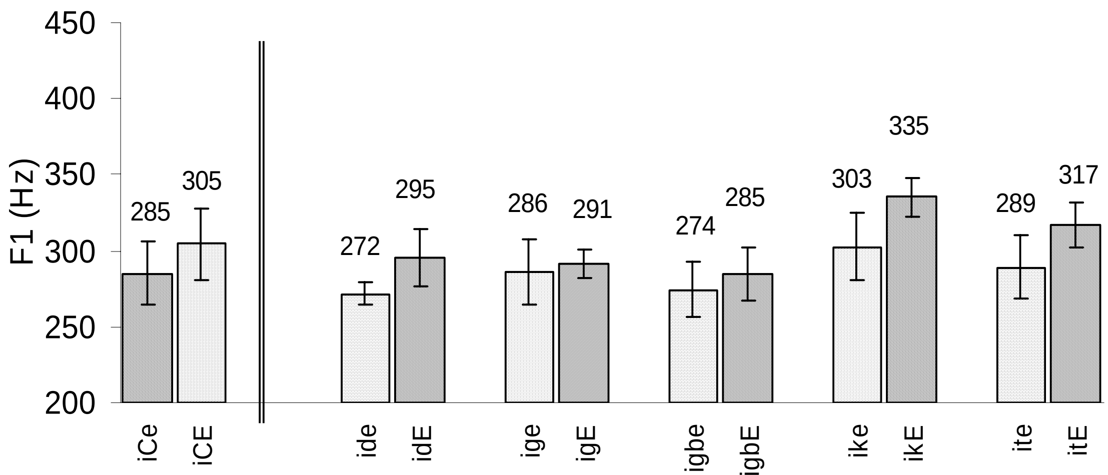


Figure 4.42. F1 Averages at V1 midpoint for Mọ̀bà speaker Mb10, for different consonants. For each bar n=6 tokens. Error bars mark one standard deviation.

The [i] before [+ATR] [e] had a lower mean F1 value than did the [i] before [−ATR] [ẹ] for each consonant. A one-way ANOVA was performed for each consonant group with F1 as the dependent variable. Slightly significant effects were found for V2 for consonants [d], [F(1, 10)=8.4, p = .016]; [k], [F(1, 10)=9.7, p = .011]; and [t] [F(1, 10)=6.8, p = .026]. To measure the overall effect, a two-way ANOVA was used with F1 of V1 as the dependent variable and C and V2 as factors. The F1 of [i] in i̲Ce versus i̲CE showed a highly significant effect for V2 [F(1, 20), p < .001]. This experiment has shown that for /i/, there is a strong coarticulatory effect in the same direction as the harmony pattern found in Àkùrẹ́.

In addition, it is clear that the identity of the intervocalic consonant influences the degree of coarticulation with some consonants allowing more vowel-to-vowel coarticulation than others. This is not surprising, as many studies have shown that the consonant plays a large role in the degree of vowel-to-vowel coarticulation. Recasens (1987) shows that the degree of coarticulation across a consonant depends on the type of consonant: a more articulatorily stable consonant—that is, one which exhibits less variation within the consonant in different environments—allows less coarticulation between flanking vowels. Recasens compared related sonorant consonant pairs in VCV tokens in Spanish and Catalan, for example, Spanish alveolar lateral [l] versus Catalan dark alveolar lateral [ɫ]. Looking at variation in F2 during the consonant as a measure of consonant stability in different contexts, he found that the Catalan [ɫ] varied much less due to vowel context than the Spanish [l]; that is, the [ɫ] was more highly constrained in its articulation. He also found that the amount of vowel-to-vowel coarticulation across the Catalan [ɫ] was lower than it was across the Spanish [l]. That is, the more highly constrained consonant also restricted the vowel-to-vowel coarticulation. Similar results were obtained with [r] and [ɾ], found in both languages: the trill [r] varied less due to vowel context and also restricted vowel-to-vowel coarticulation across it. These findings are corroborated by data from Choi and Keating (1991) in their study of VCV coarticulation in Slavic languages. Russian and Bulgarian have both plain and palatalized consonants underlyingly. With two types of consonants, each consonant should be relatively constrained in its articulation so that perceptual distinctiveness is maintained (although, unlike in the previous study, in this study there was no independent measure of how constrained the consonants are). Indeed, the vowel-to-vowel coarticulation in both Slavic languages is relatively small (in F2) compared with a language like English which has only one set of consonants. These studies show that if the position of the tongue body is constrained, then it will not be able to accommodate the previous vowel nor anticipate the following vowel. The coarticulation effect does not simply cross the consonant to get to the vowel, but rather involves the consonant. While no attempt was made to determine the relative variability of consonants in this study, the results of this acoustic study corroborate the findings that different consonants allow a varying degree of coarticulation.

It is clear from the results of the study with speaker Mb10, that the choice of only two consonants for the larger study was not ideal, and that the inclusion of at least one other consonant in the study would have made the results more general, and more favorable to supporting the claim of similarity between harmony and coarticulation. In light of this, for future studies, I recommend a wider range of intervocalic consonants. A related question not pursued here is how coarticulation of /u/ would have fared in the same experiment.

4.3.3. Mọ̀bà coarticulation in [a]

As in Àkùrẹ́, the Mọ̀bà low vowel /a/ is realized only as [a], there is no allophony. An examination of coarticulation of [a] is shown in Figure 4.43, where the striped bars indicate the mean F1 of [a] before [+ATR] mid vowels, and the dotted bars show the same for [a] before [−ATR] mid vowels. The graph shows a tendency for [a] before [−ATR] mid vowels to have a greater F1 than [a] before [+ATR] mid vowels, except for speaker Mb10. A two-way ANOVA with subject and ATR of V2 as factors shows that for all Mọ̀bà speakers combined, the F1 difference is not significant for the factor ATR of V2 [$F(1,95) = 3.33$, $p = .07$]. Only the differences in Mb9 are slightly significant [$F(1,23) = 5.51$, $p < .05$].

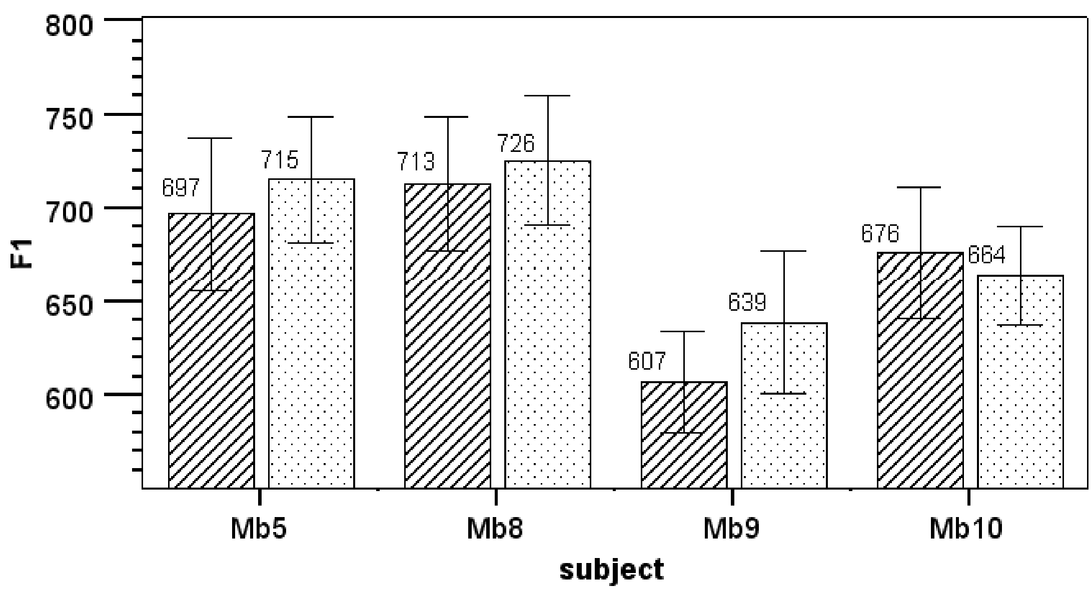


Figure 4.43. F1 mean values at various points throughout V1 for aCeo (striped bars) versus aCEO (dotted bars) for four Mọ̀bà speakers.

Figure 4.44 shows the means for all labels for each Mọ̀bà speaker. Except for Mb10, all Mọ̀bà speakers exhibit patterns similar to those of Àkùrẹ́ speakers.

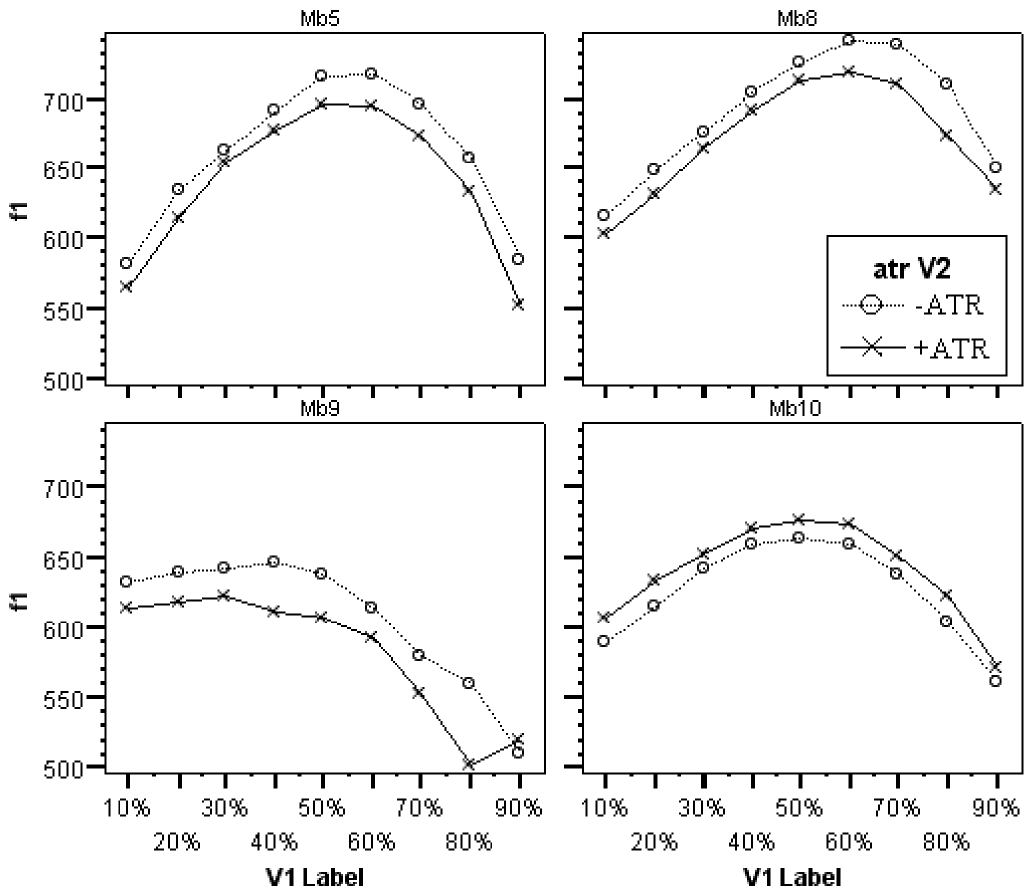


Figure 4.44. F1 means for all labels of low vowels [a] before [+ATR] mid vowels (labeled ‘a’ in graph) and before [−ATR] mid vowels (labeled ‘A’ in graph) for Mọ̀bà speakers for V1 of VCV tokens for mid vowel V2s.

4.3.4. Mọ̀bà summary

Mọ̀bà mid vowel phonemes exhibit clear differences with no overlap, as found for Àkùrẹ́ mid vowels. For the low vowel set examined, subtle coarticulatory effects may be in evidence, although further study is necessary to test this claim. The coarticulatory effects found in the high vowel sets was much more robust and statistically significant for front vowels. The patterns found in Mọ̀bà /i/ are those predicted in this study, particularly when considering the additional study using a greater number of consonants. The pattern for /u/ showed tendencies in the predicted direction, but was not found to be statistically significant in this study.

4.4. SY

In this section, we look at the results for SY, which, like Mọ̀bà, does not exhibit vowel harmony in high vowels. The vowel charts for the four SY speakers are shown in Figures 4.45-4.48, with the four vowel charts shown together in Figure 4.49. Because SY has no allophonic variation in the high vowels and no high back vowel /u/ in word-initial position, these vowel charts show only six vowels.

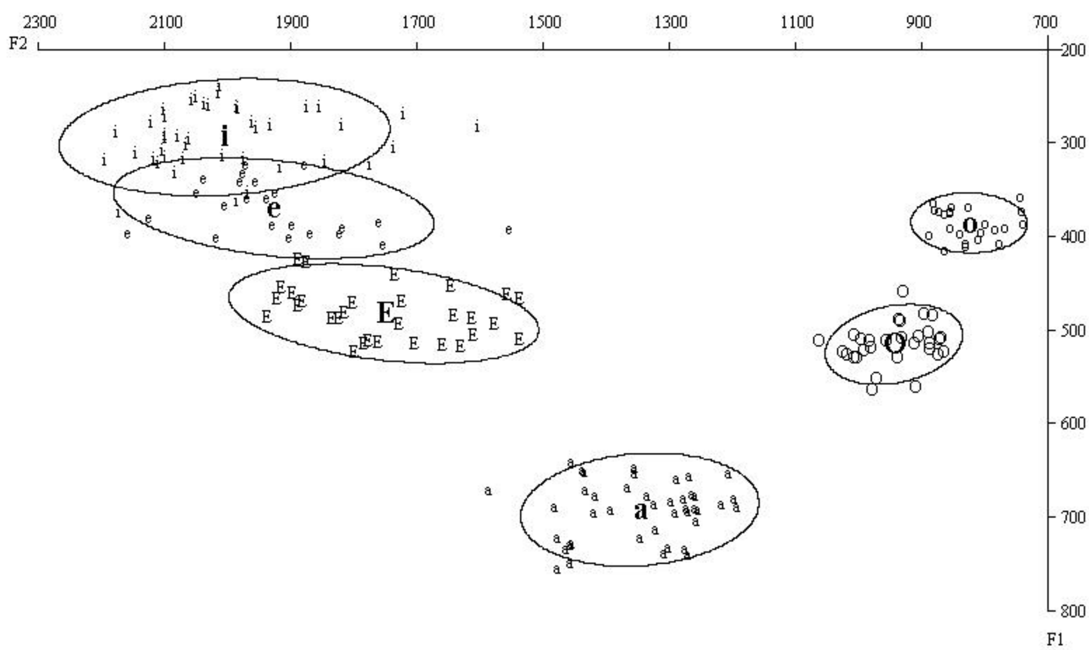


Figure 4.45. Vowel Chart (F1 versus F2) of SY Speaker SY1 for the midpoint of V_1 in V_1CV_2. The small letters indicate the values for individual tokens; the large letters show the mean for the vowel, and the ellipses show one standard deviation from the mean for each vowel. The [+ATR] vowels are [i, e, o, u]; the [−ATR] vowels are [E, O, a].

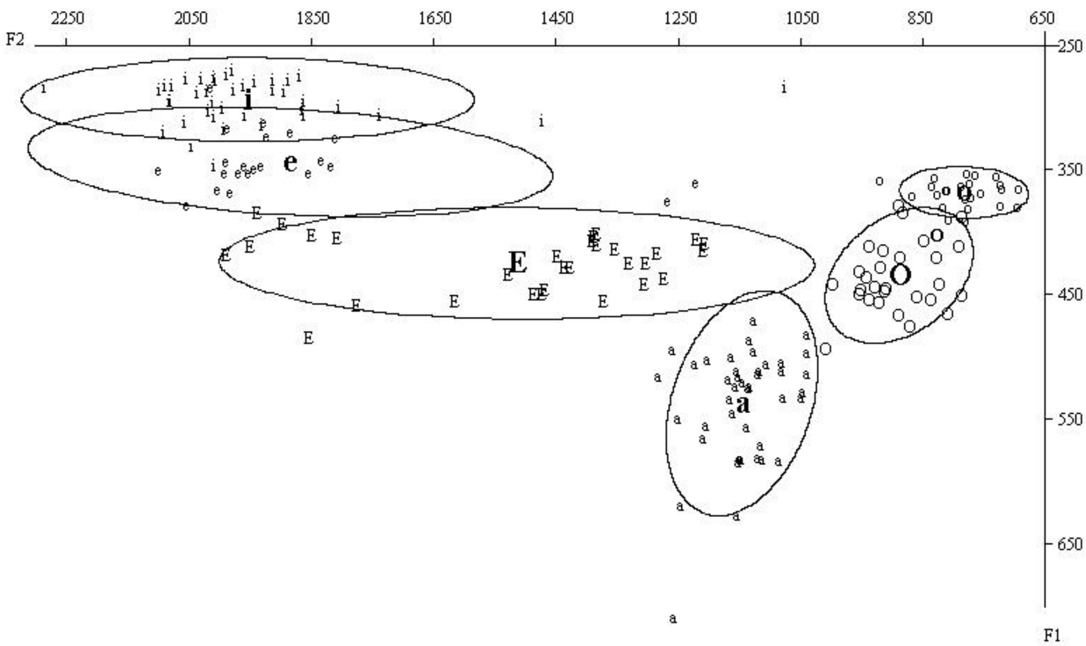


Figure 4.46. Vowel Chart (F1 versus F2) of SY Speaker SY2 for the midpoint of V_1 in V_1CV_2. The small letters indicate the values for individual tokens; the large letters show the mean for the vowel, and the ellipses show one standard deviation from the mean for each vowel. The [+ATR] vowels are [i, e, o, u]; the [−ATR] vowels are [E, O, a].

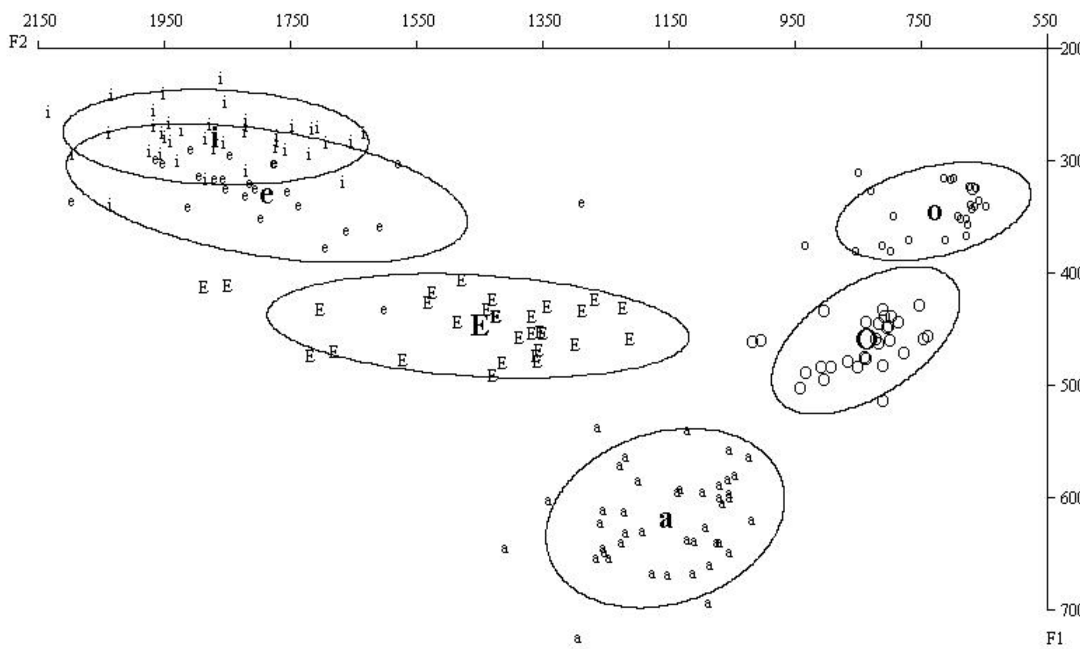


Figure 4.47. Vowel Chart (F1 versus F2) of SY Speaker SY4 for the midpoint of V_1 in V_1CV_2.The smaller letters indicate the values for individual tokens; the large letters show the mean for the vowel, and the ellipses show one standard deviation from the mean for each vowel. The [+ATR] vowels are [i, e, o, u]; the [−ATR] vowels are [E, O, a].

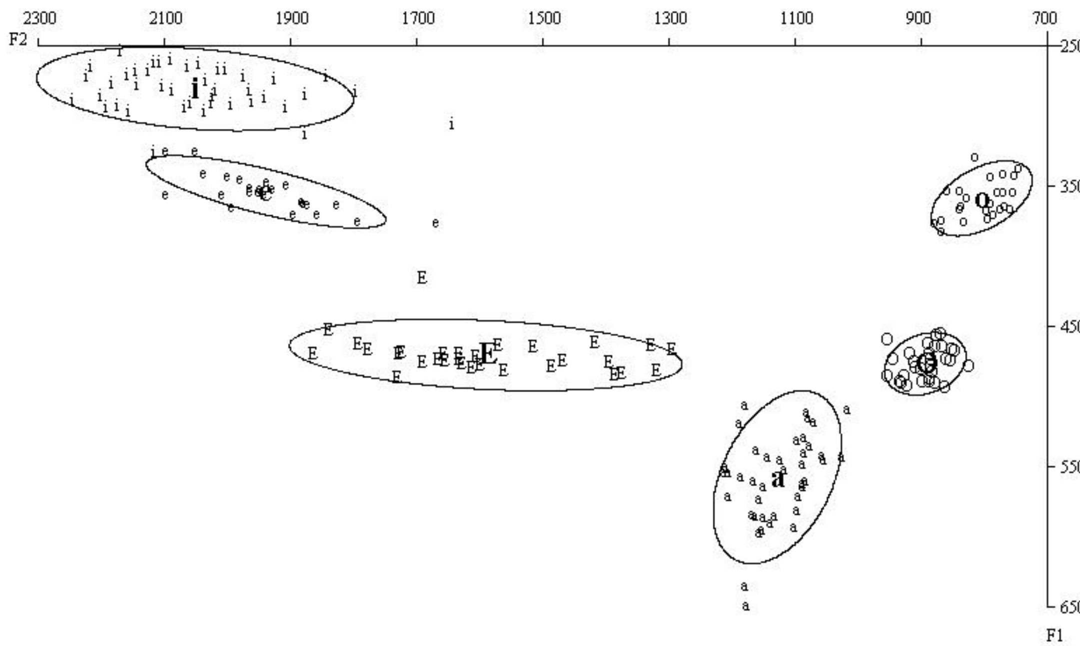


Figure 4.48. Vowel Chart (F1 versus F2) of SY Speaker SY8 for the midpoint of V_1 in V_1CV_2. The smaller letters indicate the values for individual tokens; the large letters show the mean for the vowel, and the ellipses show one standard deviation from the mean for each vowel. The [+ATR] vowels are [i, e, o, u]; the [−ATR] vowels are [E, O, a].

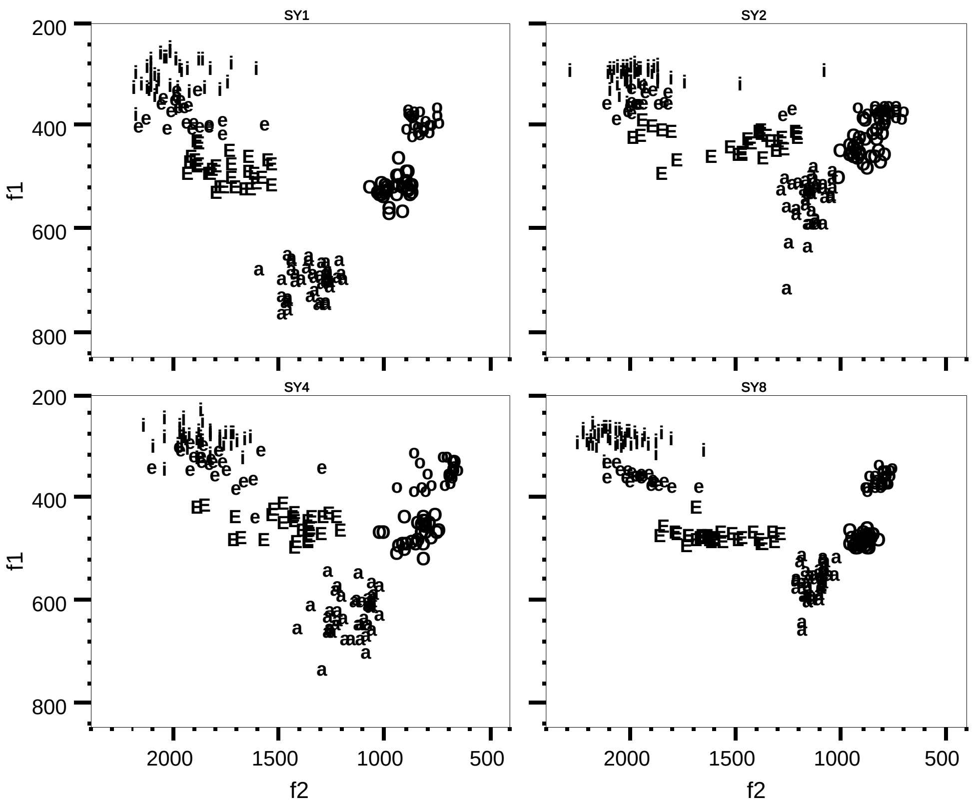


Figure 4.49. Vowel Charts (F1 versus F2) of SY Speakers SY1, SY2, SY4, SY8 for the midpoint V1 of VCV. [i, e, o, u] are the [+ATR] vowels; [E, O, a] are the [−ATR] vowels.

The F1 means for each SY speaker are shown in Figure 4.50−Figure 4.53.

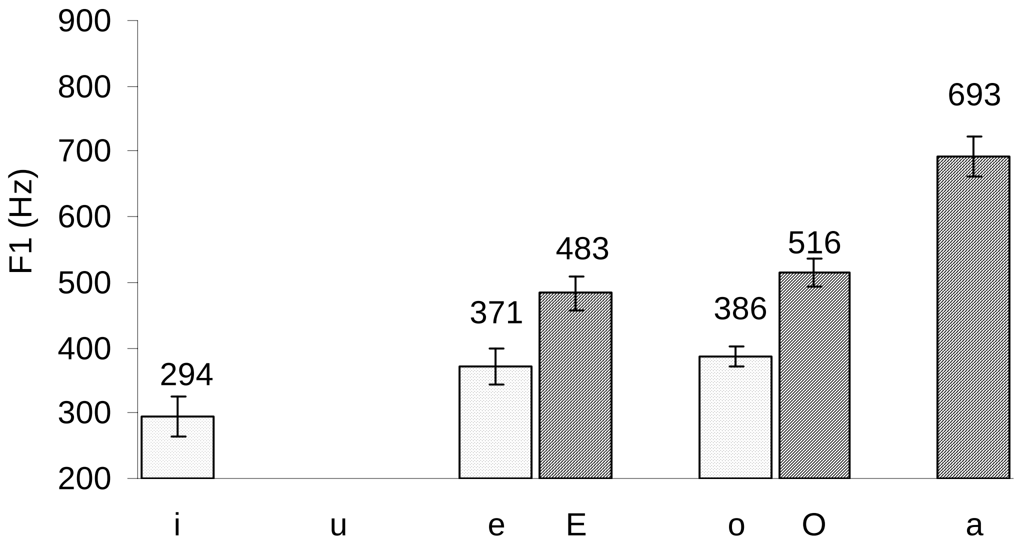


Figure 4.50. Mean F1 values at V1 midpoint for VCV words for SY speaker SY1. The error bars mark one standard deviation.

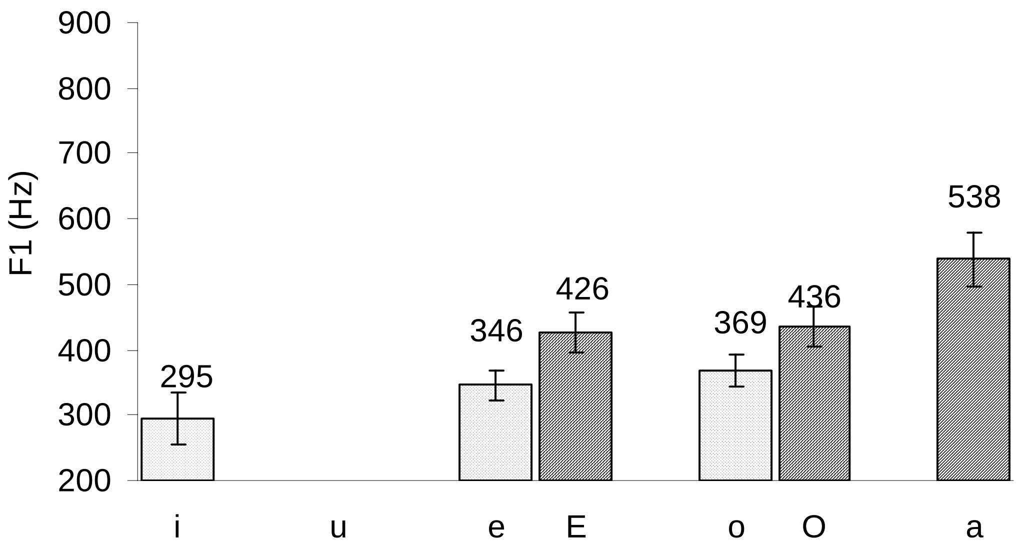


Figure 4.51. Mean F1 values at V1 midpoint for VCV words for SY speaker SY2. The error bars mark one standard deviation.

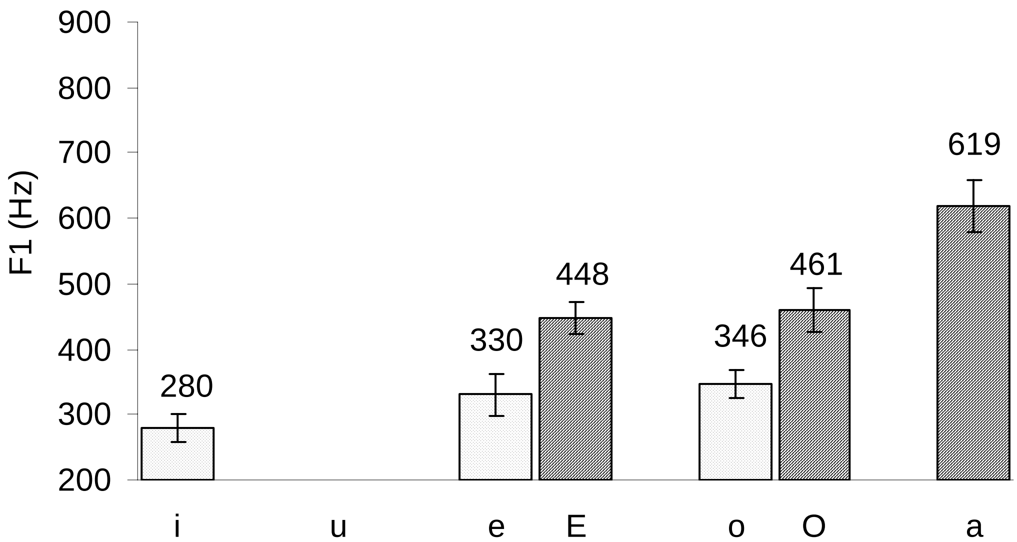


Figure 4.52. Mean F1 values at V1 midpoint for VCV words for SY speaker SY4. The error bars mark one standard deviation.

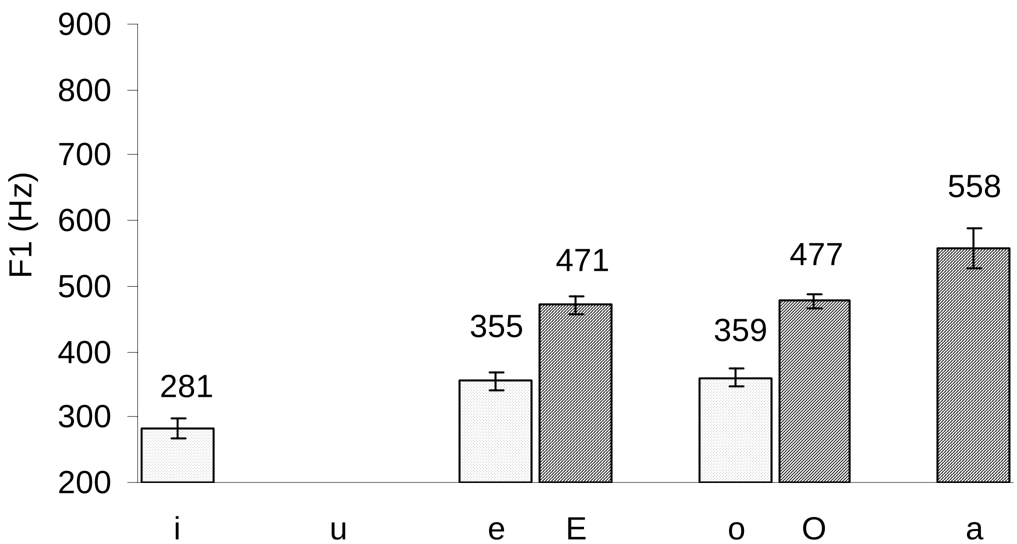


Figure 4.53. Mean F1 values at V1 midpoint for VCV words for SY speaker SY8. The error bars mark one standard deviation.

4.4.1. Different phonemes, [−ATR] versus [+ATR] mid

The ±ATR mid vowel pairs show clear differences—the pairs of /e/-/ẹ/ and /o/-/ọ/ show highly significant differences in F1 for all speakers. First the results for [e] versus [ẹ]: for SY1, the F1 difference was 112 Hz [$F(1,52) = 230$, $p < .001$]; for SY2, the F1 difference was 80 Hz [$F(1,51) = 164$, $p < .001$]; for SY4, the F1 difference was

118 Hz [F(1,52) = 248, $p < .001$]; and for SY8, the F1 difference was 116 Hz [F(1,52) = 1060, $p < .001$]. And for [o] versus [ọ]: for SY1, the F1 difference was 130 Hz [F(1,51) = 565, $p < .001$]; for SY2, the F1 difference was 67 Hz [F(1,52) = 123, $p < .001$]; for SY8, the F1 difference was 115 Hz [F(1,70) = 208, $p < .001$]; and for SY8, the F1 difference was 118 Hz [F(1,52) = 1203, $p < .001$].

Even though F1 values show no overlap between vowel tokens of [e] versus [ẹ], an examination of the individual F1 values shows that some F1 values of [e] and [ẹ] are very close to each other for SY1, SY2, and to some degree SY4, as shown in Figure 4.54.

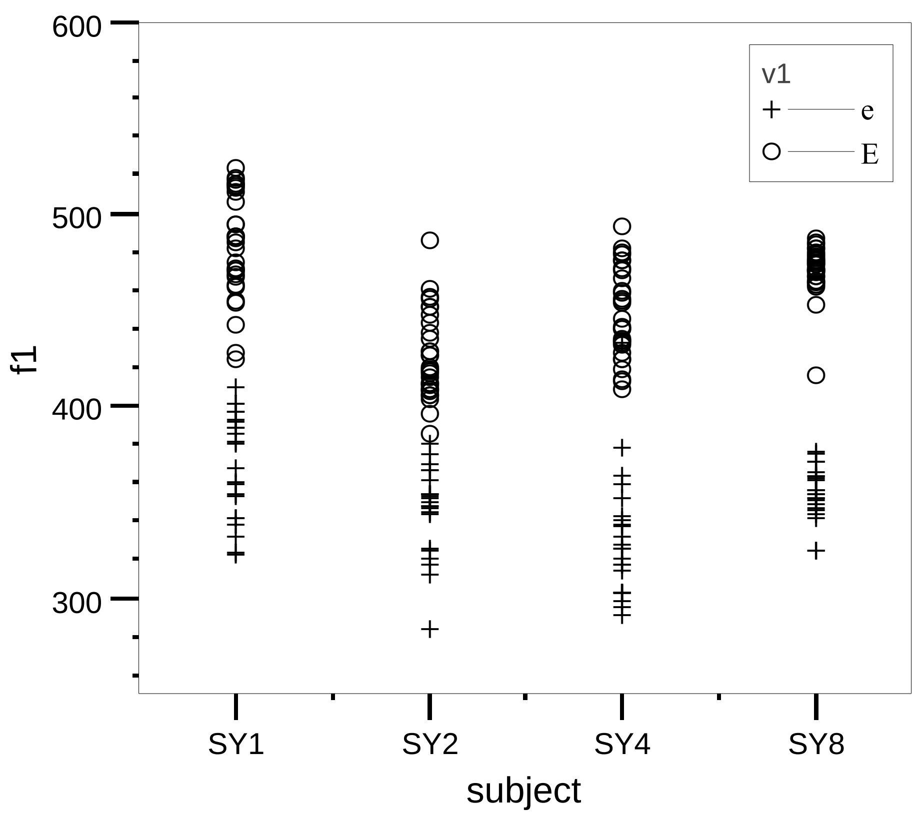


Figure 4.54. F1 values for midpoints of mid vowels [e, ẹ] (e, E, in graph) for SY speakers for V1 of VCV tokens.

As we did for Mọ̀bà, if we examine the influence of the following vowels (V2) on V1 for each speaker, Figure 4.55, then the contrast between F1 values for [e] and [ẹ] appears somewhat more evident, however some overlap still remains.

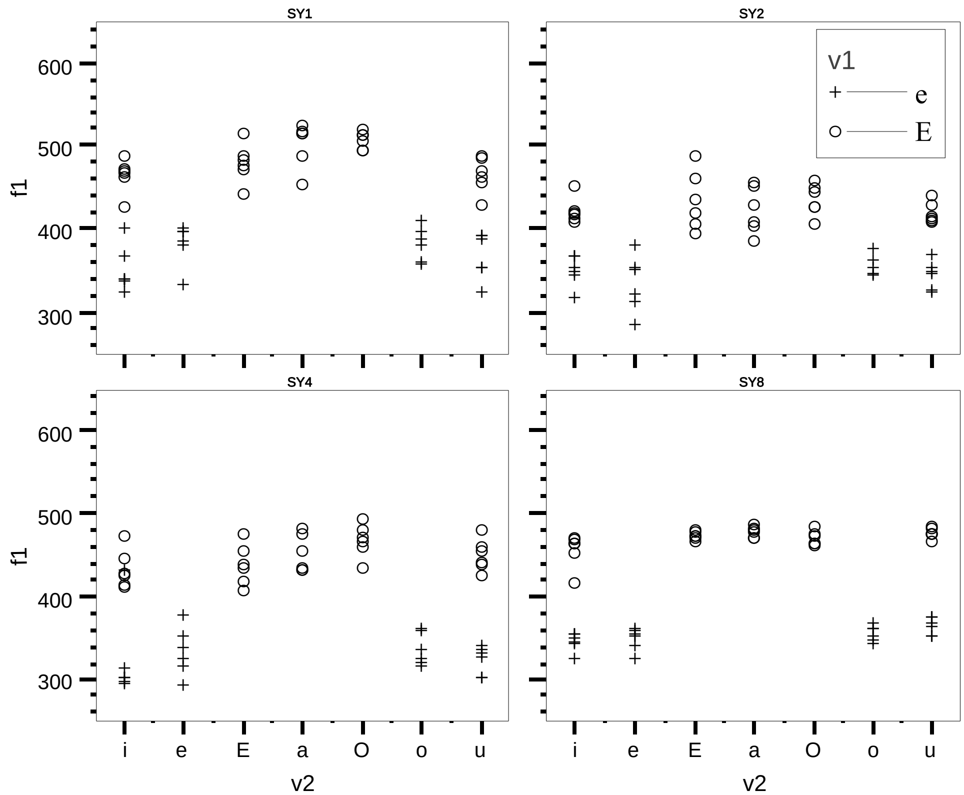


Figure 4.55. F1 values for midpoints of mid vowels [e, ẹ] (e, E, in graph) for SY speakers for V1 of V-C-V tokens.

Once we consider the intervocalic consonant, as in Figure 4.56 for [k] and Figure 4.57 for [b], the contrast is much more robust.

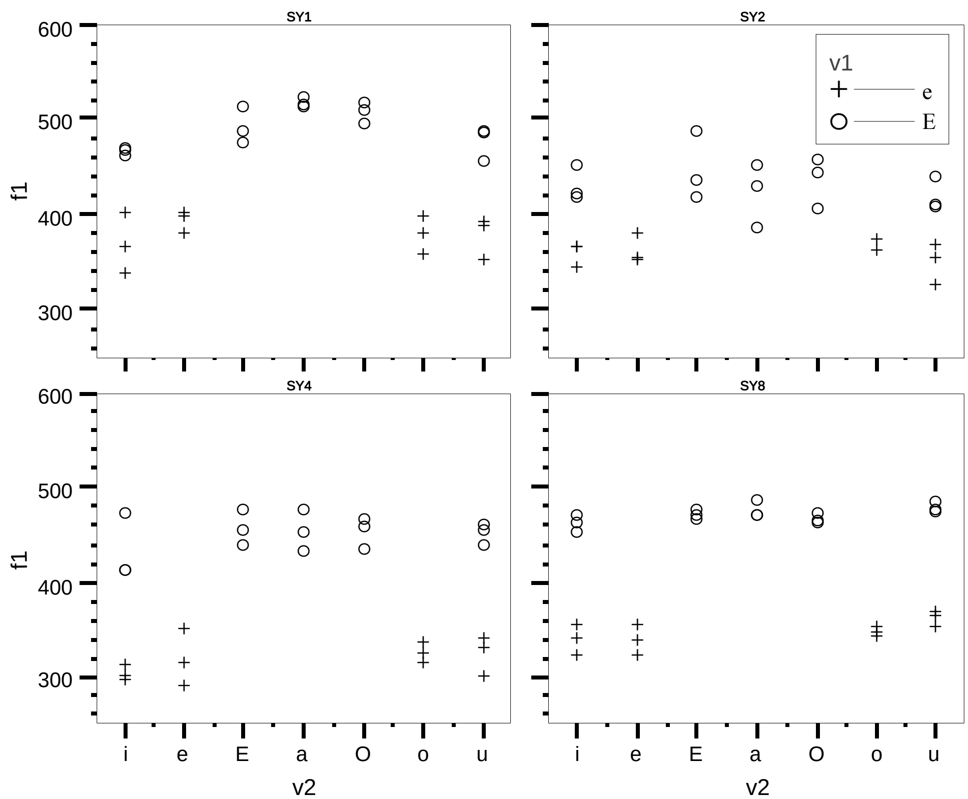


Figure 4.56. F1 values for midpoints of mid vowels [e, ẹ] (e, E, in graph) for SY speakers for V1 of V-k-V tokens.

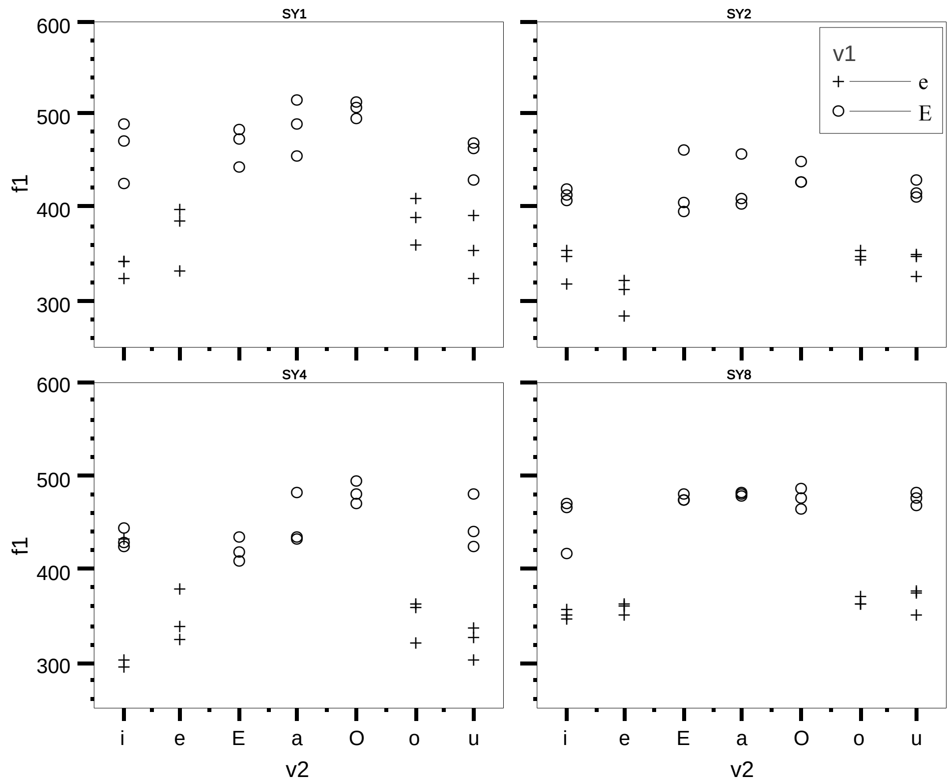


Figure 4.57. F1 values for midpoints of mid vowels [e, ẹ] (e, E, in graph) for SY speakers for V1 of V-b-V tokens.

For [o] versus [ọ], the F1 values are far apart for each speaker, except SY2, as shown in Figure 4.58. When this vowel pair is viewed by vowel and consonant (not shown here) the contrast is more clear.

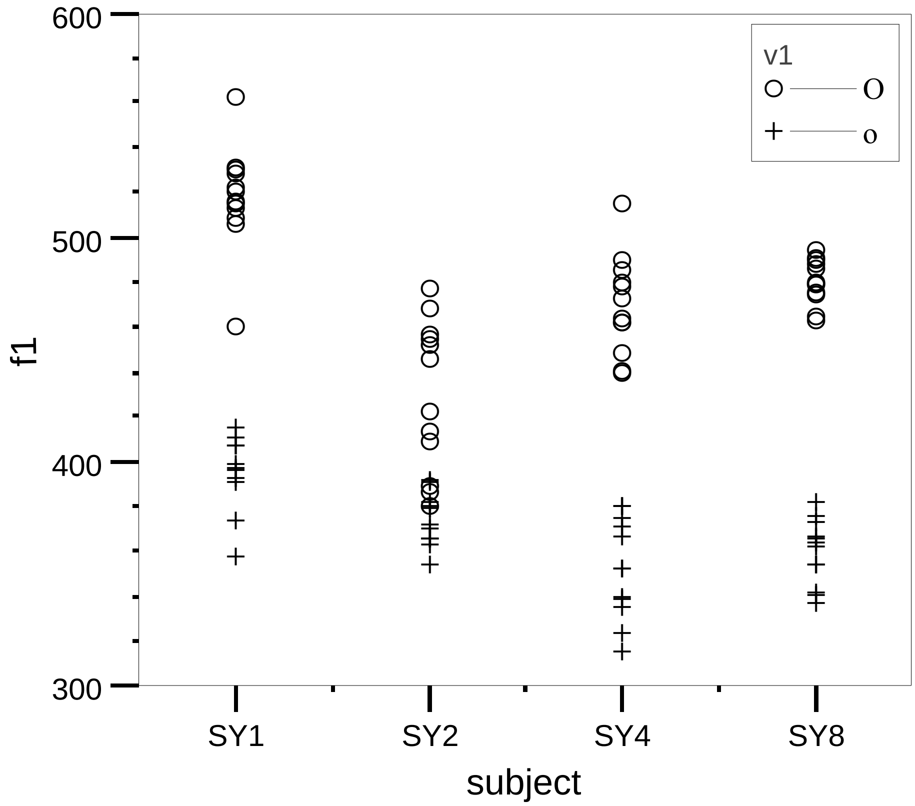


Figure 4.58. F1 values for midpoints of mid vowels [o, ọ] (o, O, in graph) for SY speakers for V1 of VCV tokens.

4.4.2. Coarticulation in SY high vowels

Now we examine vowel to vowel coarticulation in /i/, the only high vowel in word initial position. Figure 4.59 shows the mean F1 values for [i] when followed by [+ATR] and [−ATR] vowels for four speakers of SY. The dotted bars for each speaker indicate the mean F1 for [i] before the [+ATR] vowels (/i/,/u/,/e/, /o/); the striped bars indicate the mean F1 for [i] before [+ATR] vowels ([ẹ], [ọ], [a]).

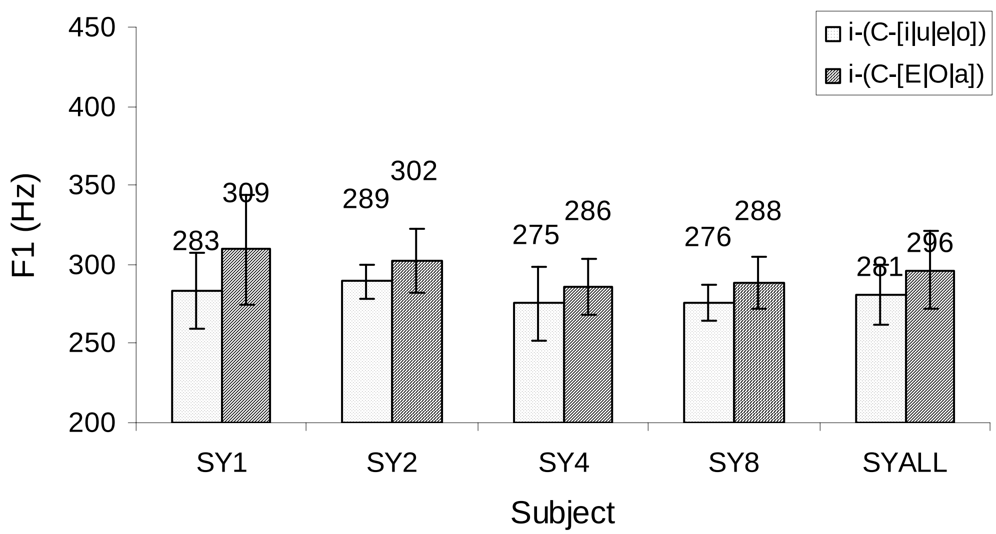


Figure 4.59. F1 means for midpoints of V1 [i] in iCV when followed by [+ATR] V2s (dotted bars) and by [−ATR] V2s (striped bars) for all SY speakers.

For each speaker, the average F1 is higher for [i]s before [−ATR] vowels. The direction here is the same as that found for F1 differences in Àkùrẹ́ harmony and in Mọ̀bà coarticulation. For the SY speakers, the differences are significant for SY1 [$F(1,40) = 8.1$, $p < .01$], slightly significant for SY2 [$F(1,39) = 6.6$, $p = .014$], not significant for SY4 [$F(1,38) = 2.4$, $p = .13$], and significant for SY8 [$F(1,40) = 8.3$, $p < .01$]. When taken collectively, the F1 difference was highly significant [$F(1,163) = 20.0$, $p < .001$].

For a more conservative comparison, Figure 4.60 shows the same graph arranged by V2 height, as we did for Mọ̀bà, so that the influence of the peripheral vowels could be evaluated.

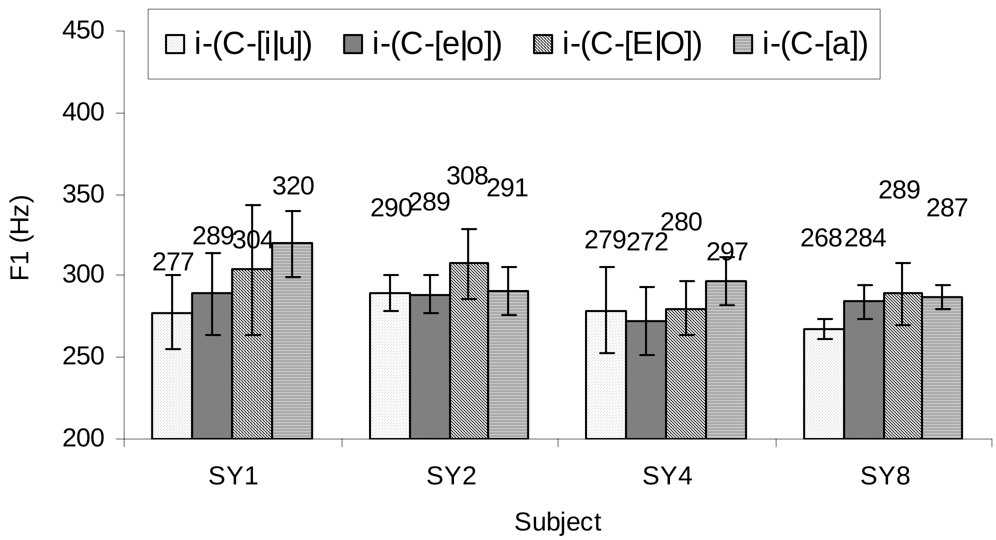


Figure 4.60. F1 means for midpoints of V1 [i] in iCV_2 for all SY speakers, where V_2 is [high], [mid +ATR], [mid −ATR], and [low], respectively, for the four bars for each speaker.

The peripheral vowels do not show as strong an effect as for Mọ̀bà, with only SY1 and SY4 having [a]s with the highest F1 of the four vowel height environments, and only SY1 and SY8 have [i]s with the highest F1. Looking at only the mid vowels as V2—the middle two bars for each speaker—we see that the F1 differences are all in the predicted direction; however, the differences are not nearly as robust as we found for Figure 4.59, with only SY2 having a slightly significant difference [$F(1,21) = 6.8$, $p = .016$]. When taken collectively, the F1 difference for all four SY speakers is slightly significant [$F(1,92) = 5.5$, $p = .02$]. A three factor ANOVA with subject, intervocalic consonants, and ATR of V2 as factors shows that for all speakers combined, the F1 difference is significant for the factor ATR of V2 [$F(1,93)$, $p=.002$]. In summary, SY [i]—like Mọ̀bà [i]—exhibits some coarticulation resembling the patterns of [i] in Àkùrẹ́ vowel harmony.

4.4.3. Coarticulation in SY /a/

The behavior of /a/ in SY is similar to /a/ in Àkùrẹ́ and Mọ̀bà. There is a slight tendency for the mid point of V1 [a] to have a greater F1 before [−ATR] mid vowels. A two-way ANOVA with subject and ATR of V2 as factors shows that for all SY speakers combined, the F1 difference is not significant for the factor ATR of V2 [$F(1,95) = 2.88$, $p = .09$].

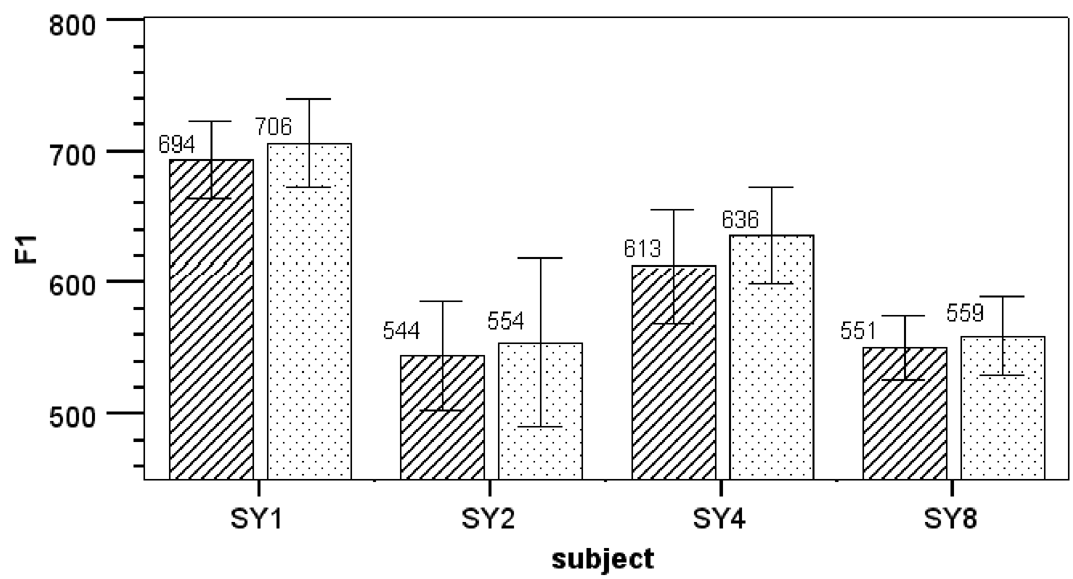


Figure 4.61. F1 mean values at various points throughout V1 for aCeo versus aCEO for four SY speakers.

4.4.4. SY summary

SY mid vowel phonemes exhibit clear differences with no overlap, as found for Àkùrẹ́ and SY mid vowels. For the low vowel set examined, small coarticulatory effects may be in evidence, though more study is required to verify this. The coarticulatory effects found in the high vowel set was much more robust and statistically significant than for the low vowel sets. The patterns found in SY high vowels are those predicted in this study.

4.5. Characteristics in Mọ̀bà and SY coarticulation compared to Àkùrẹ́

Having examined the coarticulatory effects in high vowels of Mọ̀bà and SY, these effects can be compared with the high vowel harmony in Àkùrẹ́ from (2), which is repeated here.

(2) Characteristics of Àkùrẹ́ harmony [Repeated]

- targets — the targets affected by vowel harmony:
  - /i/ and /u/
- contexts — the contexts in which the vowel harmony occurs:
  - when the target is followed by a vowel [ẹ], [ọ], [a], [ị], [ụ]
  - with an intervening consonant
- effect — the acoustic effect realized in the harmony:
  - raising of F1

In Mọ̀bà and SY coarticulation we have seen similar characteristics, though as expected, the effects are not as strong as seen in Àkùrẹ́. We have observed an F1 increase in /i/ vowel when the following vowel is [ẹ], [ọ], and [a], with an intervening consonant. Figure 4.62 shows the difference between vowel harmony for the high front vowel phoneme /i/, for Àkùrẹ́ speakers on the left, and vowel-to-vowel coarticulation, for Mọ̀bà and SY speakers on the right. The vowel /u/ shows a weaker effect than /i/. For other potential target vowels, the only one that can be followed by all the vowels above is /a/. The vowel /a/ showed some coarticulatory effects due to the following vowel, though not as strong as /i/. Although not investigated further, I expect that the coarticulatory pattern on the low vowel /a/ exhibits patterns resembling vowel harmony, since /a/ also participates in vowel harmony, for example, in Akan (Stewart 1967, Clements 1981) and Degema (Fulop et al.1998).

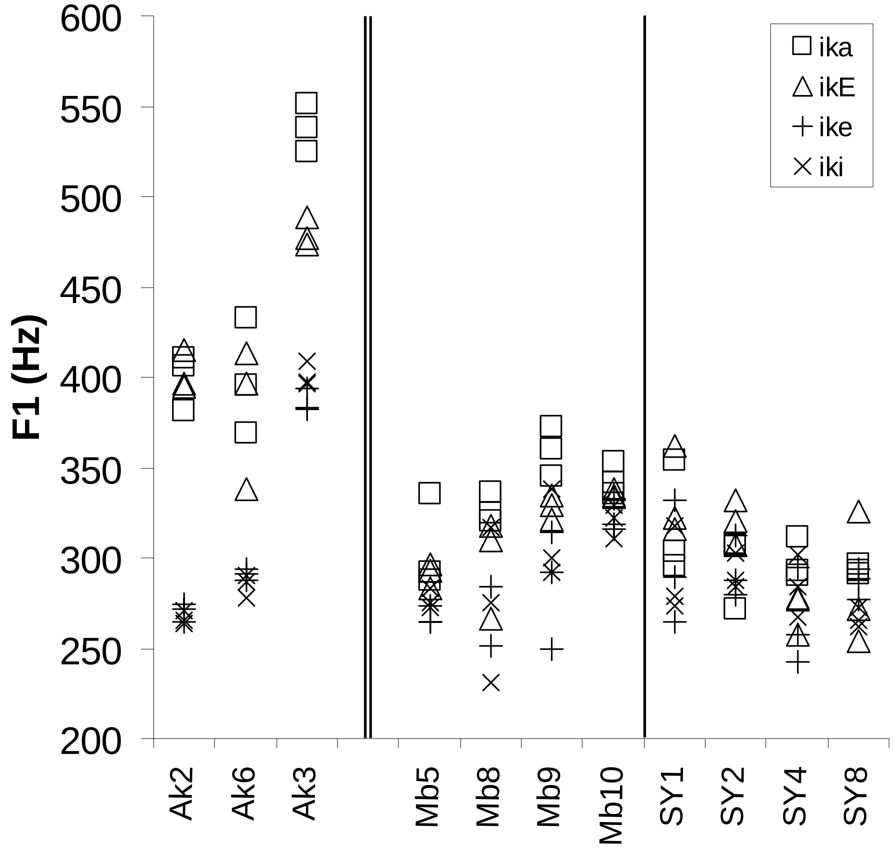


Figure 4.62. F1 values of /i/ measured in the middle of the first vowel (V1) in individual tokens of four VCV contexts: *ika*, *ikẹ*, *ike*, *iki* in speakers of Àkùrẹ́, Mọ̀bà, SY.

I expected that the harmony patterns in Àkùrẹ́ would be weaker in SY and Mọ̀bà, both in magnitude of F1 difference and in the level of overlap. What was not expected was that /u/ would not show a statistically significant level of difference before [+ATR] versus [−ATR] vowels. Although the patterns in SY and Mọ̀bà are relatively salient for /i/, they are not for /u/. This leaves the dilemma of explaining how it is that /i/ and /u/ pattern together as targets for harmony, since they do not both exhibit harmony-like patterns in these data. Further studies examining /u/ with different intervocalic consonants and a greater number of tokens may shed light on this issue. This issue is taken up further in the next chapter.

CHAPTER FIVE
DISCUSSION

5.1. Overview

The results of this study strongly support Ohala's (1994:491) claim that "vowel harmony [...] is a fossilized remnant of an earlier phonetic process involving vowel-to-vowel assimilation." In this chapter, I summarize the results showing that coarticulation in Mọ̀bà and Standard Yorùbá is similar to the high vowel harmony in Àkùrẹ́ Yorùbá (§5.2). The results are consistent with predictions for /i/ made in chapter 3, however /u/ does not pattern as expected (§5.2.2). The data support the conclusion that the phonology in question has its origin in the coarticulation. I propose a scenario where a language without harmony develops into a language with harmony due to the effects of coarticulation (§5.3). In order to test this scenario, I apply a decision tree model and determine that the coarticulation data contain the information necessary for a learner to infer a harmony pattern without resorting to pre-existing features (§5.4). Grammar change in general is briefly discussed (§5.5). I revisit the historical development of high vowel harmony in Yorùbá in §5.6. The physical origins of coarticulation are discussed in §5.7. Finally, in §5.8, I present conclusions and directions for further study.

5.2. Summary of results

In chapter four, I showed that the acoustic and distributional characteristics of the Àkùrẹ́ Yorùbá high vowel harmony pattern are evident to some degree in the coarticulatory patterns of Mọ̀bà and Standard Yorùbá (SY), particularly for /i/, though not so much for /u/. In Àkùrẹ́, the alternation of [±ATR] high vowels is a typical phonological pattern in the following ways: for Àkùrẹ́ speakers, the difference between the F1 mean of [+ATR] [i] and [−ATR] [ị] is large, between 54 and 91 Hz, with very little overlap; the harmony spreads leftward phonologically from the high vowels to preceding vowels; and some Àkùrẹ́ speakers are aware of the difference between the two allophones of /i/. For SY and Mọ̀bà, the effects are not phonological. Notably, and as predicted, the coarticulatory effects in SY and Mọ̀bà are much smaller than the harmony effects in Àkùrẹ́. For the SY and Mọ̀bà speakers, the difference in the F1 mean between [i] before [+ATR] vowels and [i] before [−ATR] vowels is between 6 and 43 Hz with considerable overlap. Additionally, in SY and Mọ̀bà, the vowel patterns do not trigger additional harmony to their left. Finally, Yorùbá speakers do not recognize the /i/ or the /u/ as having allophones. Consistent with this, the literature on Mọ̀bà and SY phonology—extensive for SY—does not contain claims that these vowels have allophones or undergo harmony. In sum, Mọ̀bà and SY exhibit coarticulatory effects, while Àkùrẹ́ exhibits a phonological effect. Figure 5.1 illustrates this point, showing F1 values at the midpoint of the initial vowel /i/ for individual tokens of ikV, where V2 is {i, e, ẹ, a} for speakers of Àkùrẹ́, Mọ̀bà, and SY.

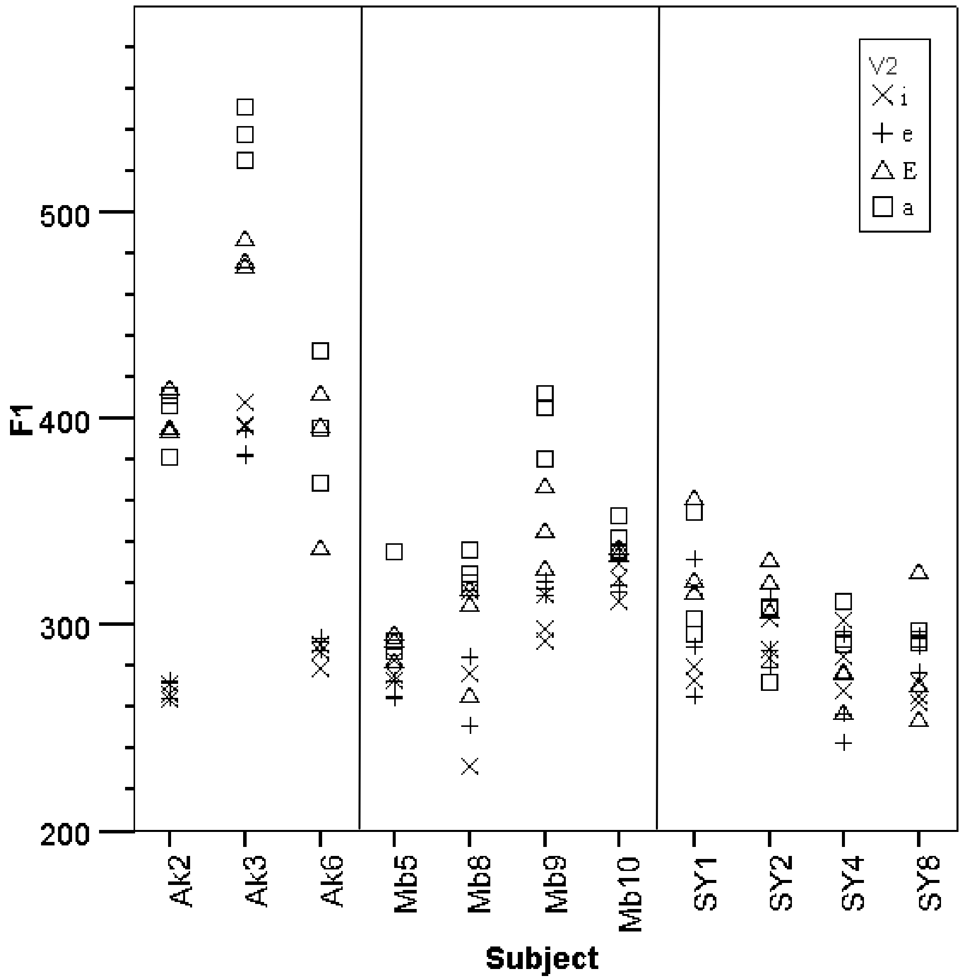


Figure 5.1. F1 values of /i/ measured in the middle of the first vowel in individual tokens in four VkV contexts: *ika*, *ikẹ*, *ike*, *iki* in speakers of Àkùrẹ́, Mọ̀bà, and SY. V2, the identity of the second vowel, is labeled “i”, “e”, “E”, “a” for [i], [e], [ẹ], [a], respectively.

Although there is a clear difference between the phonological behavior of high vowels in Àkùrẹ́ and the phonetic behavior of the corresponding vowels in the other two dialects, the similarities are striking in terms of effect and context. In other words, characteristics of the phonological pattern are present in the phonetic pattern. For each speaker of the three Yorùbá dialects, the mean F1 value for /i/ before the vowels [ẹ, ọ, a] was greater than the mean F1 value of [i] before the vowels [i, u, e, o], though the quality of the difference varied between Àkùrẹ́ and the other dialects, as noted above. We see in §5.4, below, that the categorization of the vowels into /ẹ, ọ, a/ and

/i, u, e, o/ can be extracted from the phonetic data, both in Àkùrẹ́ and the other dialects. Even after eliminating the outermost vowels /i/, /u/, and /a/ and thus examining a more constrained set of V2 vowels—the mid vowels—the similarity between Àkùrẹ́ and the other dialects in terms of F1 was maintained. Finally, in both the harmonizing and coarticulatory environments, intervening consonants do not block the effects. A close examination of one Mọ̀bà speaker, Mb10, shows that coarticulation occurs with different intervening consonants, just as is found with harmony in Àkùrẹ́. An interesting topic for future research would be to look closely at coarticulation and harmony across all intervocalic consonants for Àkùrẹ́ and Mọ̀bà.

For the purposes of this study, interactions within different domains were not studied. In fact, this study does not address the question of how the domain of the vowel harmony is determined, largely because this issue would have expanded the study greatly. It is not necessarily the case that domain of harmony can be extracted from the patterns of vowel-to-vowel coarticulation. Further study could shed light on how vowel-to-vowel coarticulation is realized in different levels of the prosodic hierarchy, and in what ways the domain of vowel harmony differs from this, following such studies such as Fougeron and Keating (1997) and Keating et al. (1999).

5.2.1. Predicted outcomes for /i/

The results for /i/ fit the predicted pattern discussed in Chapter 3, repeated here as Figure 5.2. That is, the F1 of the initial vowel is realized in the direction of the F1 in the second vowel.

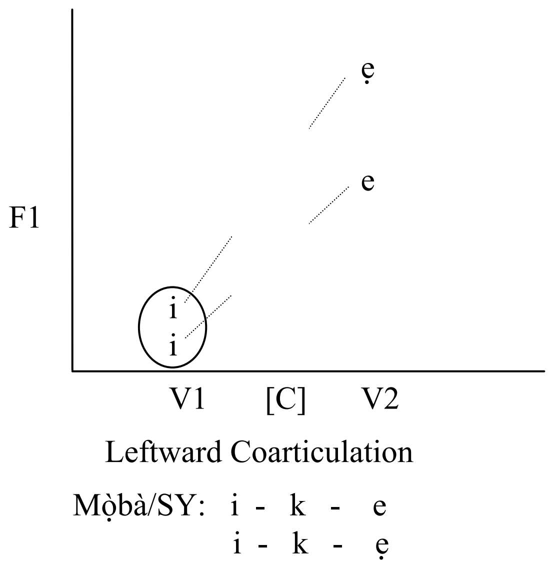


Figure 5.2. Schematic of predicted F1 value for vowels in Mọ̀bà and SY tokens /*ike*/ and /*ikẹ*/. (Repeated from Figure 3.4).

A corresponding graph is shown in Figure 5.3, this time with data from SY speaker SY1. The results closely resemble the predicted pattern.

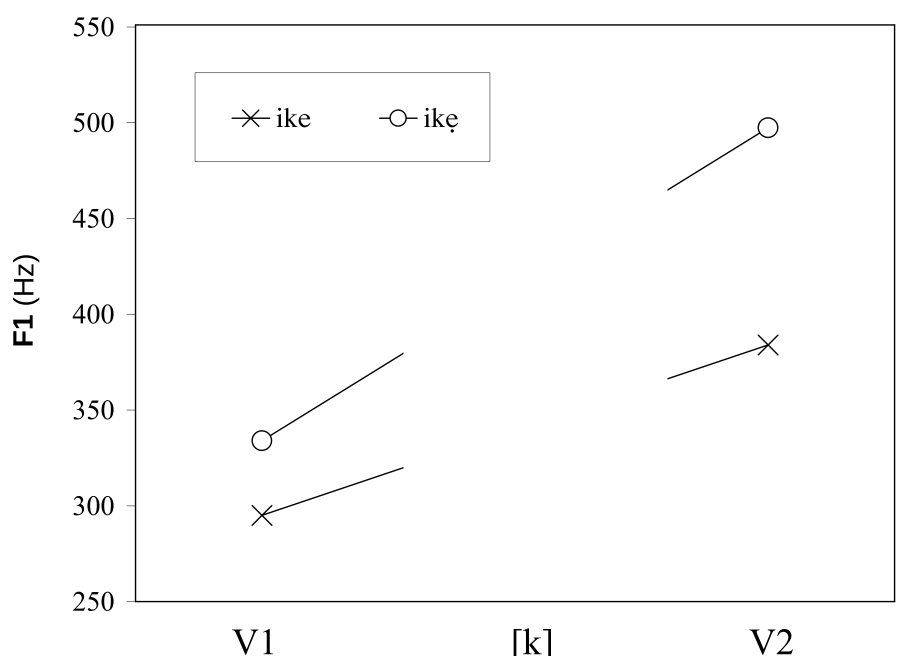


Figure 5.3. Actual F1 means from SY1 tokens for the middle of V1 and V2 vowels for 3 tokens each of /ike/ and /ikẹ/. (The lines do not represent measured values.)

In Mọ̀bà and SY, the predictions are met for V1 /i/: the coarticulation is in the direction of V2, just as in Àkùrẹ́, where the harmony in V1 spreads from V2.

5.2.2. Predicted outcomes for /u/

While the parallel between harmony and coarticulation is clear for /i/, the same cannot be said for /u/. For Àkùrẹ́, the phonological behavior of /u/ is consistent with allophony, with /u/ participating in harmony with spreading to the left. The acoustic pattern of /u/ in Àkùrẹ́ also indicates allophony, though not as clearly as for /i/. The data indicate that the differences between /u/ allophones is statistically different; however, for /u/ a great deal of overlap occurs in the acoustic realization of the allophones, especially when the intervocalic consonant is /b/. Further, in Mọ̀bà—recall that there is no initial /u/ in SY—while the difference between /u/ before [+ATR] versus /u/ before [−ATR] vowels is in the expected direction, it is not statistically significant in the data set examined here.

The prediction made in chapter 3 was that both high vowels of Mọ̀bà, and the one high vowel, /i/, of SY, would exhibit coarticulatory effects resembling vowel harmony in Àkùrẹ́. /u/ does not act as predicted, a problem for my analysis. There is an apparent asymmetry between the phonological pattern and the phonetic pattern on which I claim it is based. Three types of explanations may address this problem. The first is methodological. It may be that the coarticulatory pattern in /u/ is simply not as robust as it is for /i/, and that more data are required before the pattern emerges. There is some evidence that the vowel space in the back of the mouth is more restricted (Abigail Cohn, personal communication). This explanation can be examined in further study, and indeed, without further examination, no conclusions can be drawn. Since relatively few tokens of /u/ were measured, one could examine many more tokens for each consonant, while in addition testing a greater number of intervocalic consonants.

A second explanation for the asymmetry questions the assumption that Mọ̀bà vowels are like the vowels of Proto-Yorùbá with respect to coarticulation. It may be that the Mọ̀bà vowels have changed enough so that coarticulatory effects are no longer robust enough to show the patterns effect which once were present in Proto-Yorùbá. It is difficult to see how this explanation could be tested, except by examining other Yorùbá dialects to seek the effect. If another Yorùbá dialect—or indeed any language—exhibited the desired coarticulatory patterns in both /i/ and /u/, this would count as evidence in favor of the argument presented here. A third account for the apparent asymmetry is that in fact there is an asymmetry, and that as a part of the phonologization process, the pattern is generalized from some subset of possible contexts to two target vowels, /i/ and /u/, or even more target vowels, /ẹ/, /ọ/, /e/, /o/, /i/, and /u/, since the mid vowels already exhibit harmony. I return to this account in the next section, in which I present an explanation of how coarticulation can lead to vowel harmony.

5.3. Phonologization of coarticulation

Although the results above show that coarticulation resembles vowel harmony, a question remains as to how a learner of Proto-Yorùbá might infer a vowel harmony given the input presented to her. In what follows, I show that the vowel harmony pattern can be extracted from the coarticulatory patterns of the input with minimal assumptions. If coarticulation is to lead to vowel harmony, a learner must be able to do the following: identify two categories of vowel from the data she is exposed to; identify the physical properties of the two categories; and then, determine what context correlates with each category.

5.3.1. Adult learning of grammars

I turn briefly to some research into what kinds of grammars can be learned. For each of the studies mentioned, adults were exposed to elements of a made up language for a short period of time, sometimes with a feedback session, before being tested on what they learned. In spite of the difference from a normal language learning environment, the results are interesting, especially considering how the short exposure. Peperkamp and Dupoux's (2004) experiment with adults learning phoneme categories in an artificial language suggests that unnatural allophonic groupings are learned as well as natural ones. The patterns were learned quite quickly, after only 35 minutes of exposure. Pycha et al. (2003) looked at how adults learn harmony-like patterns in words from made up languages. They found that a harmony pattern and a disharmony pattern were equally learnable after about twenty minutes of exposure and feedback. In the first condition, a suffix's vowel agreed in the feature [BACK] with the root vowel (/ɛ/ occurred after /i, ɪ, æ/ and /ʌ/ occurred after /u, ʊ, a/); in the second condition, a suffix's vowel disagreed—that is, disharmony—with the feature [BACK] with the root vowel. In a third condition, the two suffix vowels corresponded to an arbitrary set of vowels which did not form a natural class. Specifically, /ɛ/ occurred after /i, æ, ʊ/ and /ʌ/ occurred after /ɪ, u, a/. In this condition, they did worse, suggesting that the lack of a natural class made the task more difficult. It is obviously not clear if the same effect would occur with children learning a phonological pattern triggered by an unnatural class, since the adults could have acquired the natural class through experience as opposed to having innate notions of natural class or features. Newport and Aslin (2004) found that adult learners were able to use statistical learning to learn patterns of non-adjacent vowels (and non-adjacent consonants) after about 20 minutes of exposure. This finding is relevant to vowel harmony, where interacting vowels are non-adjacent. The motivation for the research here is certainly in part to see if patterns typically absent in languages are absent because they are not readily learnable. Indeed, Newport and Aslin (2004) found that adults were not able to learn patterns involving non-adjacent **syllables**—as opposed to segments—and in fact, non-adjacent syllables do not typically interact in phonology.

5.3.2. Window model

Keating's (1990) window model of coarticulation is a fitting model for discussing the transition from coarticulation to phonology. In order to account for variation in a segment's acoustic realization, at least due in part to coarticulation, Keating proposes a range of target values—a window—associated with the production of any segment. For any segment, the same target is present in different contexts. By interpolation through windows of neighboring segments, the realization of a segment changes depending on context. In Figure 5.4, the ellipse represents the target window for /i/ in an abstract acoustic space, each "i" being an exemplar of a spoken /i/ token, following the terminology of an exemplar model such as one elaborated in Pierrehumbert (2001b).[1]

[1] While these models are not equivalent, combining them here is helpful for expository purposes.

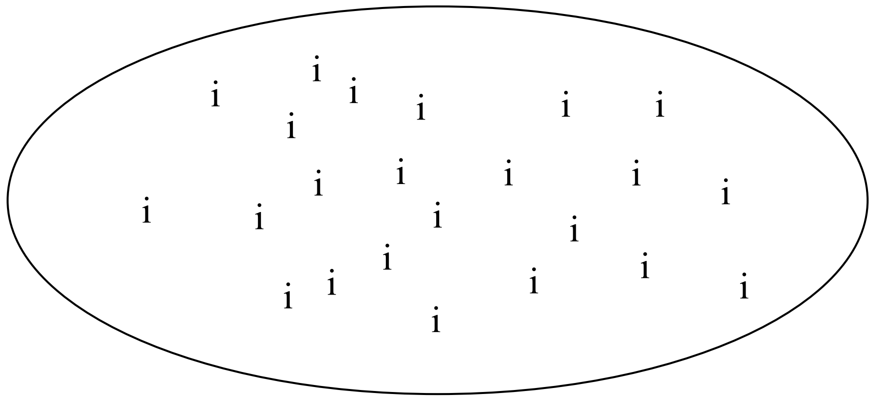


Figure 5.4. Window model of [i] in Proto-Yorùbá (Mọ̀bà like language). The ellipse represents the window for [i], in an abstract acoustic space. The letter ‘i’ represents different instances of [i], which have been arbitrarily placed for expository purposes.

Target window sizes vary from language to language. It has been noted by Bradlow (1995), Lindau and Wood (1977b), and Manuel (1990) that the same vowels have different acoustic realizations in different languages. Manuel’s (1990) study of three related languages shows that a vowel space with more vowel phonemes is more likely to exhibit a more tightly packed window. Specifically, looking at Figure 5.5, based on Manual, the distribution of instances of /A/ is smaller for Sotho, with 7 vowels, than for Ndebele, with 5 vowels.

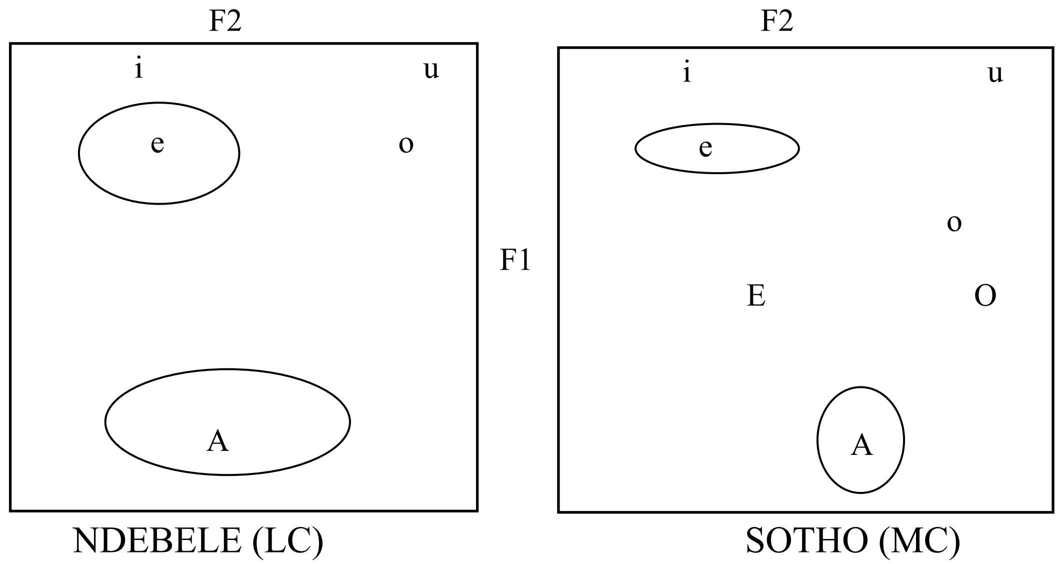


Figure 5.5. Schematization of differences in coarticulation of /e/ and /A/ (by unspecified vowels) in Ndebele and Sotho, related Bantu languages with differing vowel systems, following Manual (1990).

If a target window can vary from language to language, then it follows that a target window can change over time within the same language. A larger window, which reflects more variation in general, would conceivably lead to more context dependent coarticulation. If a window becomes larger over time, it is quite likely that context effects would become more prominent, since the larger the acoustic difference the more prominent the auditory difference. Contrast this with the languages in Choi and Keating's (1991) study of Slavic languages, where they found measurable but slight coarticulation in Russian, Bulgarian and Polish. In these languages, a smaller window would likely prevent the context dependent coarticulation from reaching an auditorily prominent level sufficient for differentiation.

5.3.3. Phonologization scenario

In this section I present a hypothesized scenario for how coarticulation might lead to vowel harmony, in which I have speculated freely. I follow Hyman (1976), discussed in Chapter 1, except that the development of vowel harmony in Yorùbá replaces Pekinese. I present three stages of development and an explanation of how the first stage might lead to the third. At stage A, which I argue to be like Proto-Yorùbá, no vowel harmony exists in the high vowels, and seven oral vowels are present. As is universally the case, some vowel-to-vowel coarticulation is present. This language is like Mọ̀bà with respect to its vowel phonology, that is, harmony does not extend to high vowels. In stage A, the coarticulation is not attended to, and does not become part of the grammar. Learners of the language infer what was the simplest pattern that models the data to some degree of precision. The simplest solution to the learner—that is, the solution composed of the simplest phonological grammar—would be to posit that all tokens close to a particular set of acoustic properties are /i/s, for example, and all vowels close to another set of acoustic properties are /u/s, etc., as shown in (1).[2]

(1) Characterization of simplest model

/i/ is realized as [i] in all contexts

/u/ is realized as [u] in all contexts

Before discussing the transition stage B, we look at stage C, an Àkùrẹ́-like language where high vowel harmony is now present. Each high vowel now has two allophones. The windows for the high front vowels is shown in Figure 5.6.

[2] What makes one vowel token acoustically similar to another is a problem that is not addressed here, but it is clear that the lower formants are an important component. An additional problem not addressed here is that of normalization of acoustic input due to differing characteristics of the speaker.

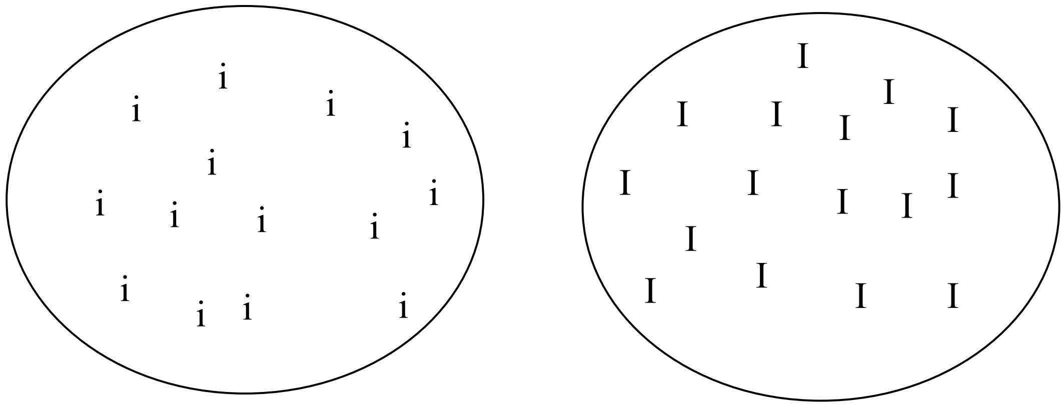


Figure 5.6. Windows of [i] and [ị] in Proto-Àkùrẹ́. One ellipse represents the window for [i], the other for [ị], shown as “I”, in an abstract acoustic space. The tokens have been arbitrarily placed for expository purposes.

Learners at stage C, being exposed to input like in Figure 5.6, would posit two allophones for two different contexts. The grammar at stage C would posit that /i/ has one realization before one set of vowels, and another realization before another, without regard to intervening consonant, as shown in (2). The grammatical pattern of stage C is inherently more complex than that of stage A but attains a greater precision of capturing the data. This is essentially high vowel harmony applied to /i/.

(2) Characterization of high vowel harmony (for /i/)[3]

/i/ is realized as [i] before one set of vowels

/i/ is realized as [I] before the complementary set of vowels

Of course, arriving at stage C is not inevitable—for many or all other dialects at that time in Yorùbá development stage C did not occur, and has likely not occurred since.

Stage B[4] is the intermediate stage. At stage B, learners attend to the effects of coarticulation and infer a different pattern, and thus a different grammar. It is not clear

[3] The separate problem of whether and how the allophones are considered by the learner to be part of the same phoneme is not addressed here.

[4] Stage A and C here are analogous to Hyman’s stage I and II, respectively. Stage B is an intermediate stage that does not have an analogue in Hyman’s discussion.

if stage B is very different from stage A. As I have no way of determining whether Mọ̀bà and SY are closer to stage A or stage B, I must assume that they are similar to stage B and I thus examine them for evidence of harmony.

It is one thing for a linguist to look at a set of data *ex post facto* to see if a generalization is found, it is quite another thing for a learner to make the more difficult calculation of determining criteria for splitting a data set given an input like Figure 5.7, which resembles Figure 5.4, except that the data are real, and some context is noted. In this figure, the F1 and F2 of the tokens of V1 /i/ are labeled with the word they came from, so the label includes the identity of the following consonant and vowel.

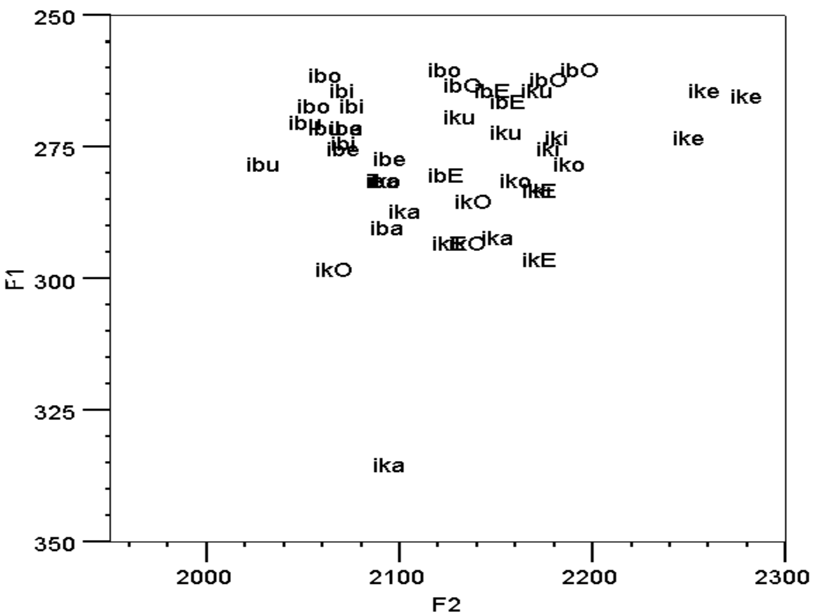


Figure 5.7. Mb5 iCV tokens. F1 and F2 values.

Assuming that the learner does posit two realizations of [i] depending on context, how would the learner determine which contexts belonged to which set? Such a determination must presumably come from the data to which the learner is exposed. Looking at Figure 5.7, it is not clear how to divide the data into two.

Figure 5.8 shows the mean F1 averages for /i/ and /u/ for four Mọ̀bà speakers for every CV context in the experiment. With the means arranged in descending order, there is no clear point at which to divide the higher from the lower F1 means.

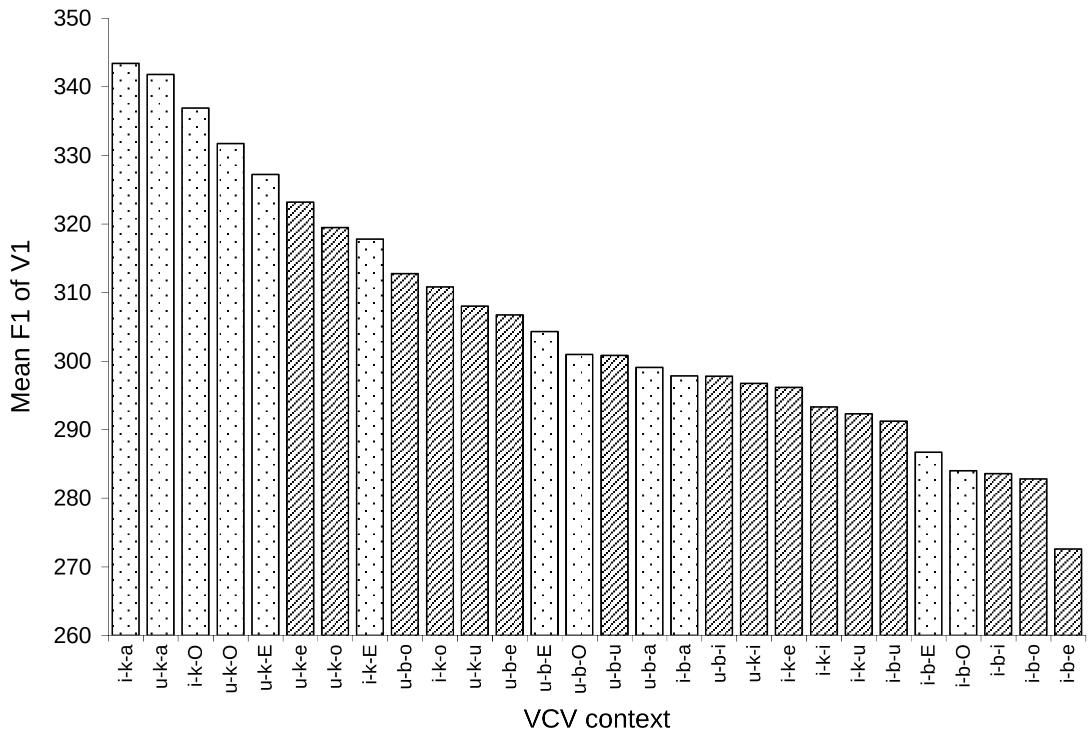


Figure 5.8. F1 means for midpoint of V1 in VCV tokens in different contexts for four Mọ̀bà speakers. The bars are arranged from greatest to smallest mean. Dotted bars indicate [−ATR] V2s, striped bars indicate [+ATR] V2s.

If a phonological split were based purely on this data without regard to symmetry or simplicity, we may expect to see a division into two or more groups such as in (3), as if a line were drawn at a point somewhere in the continuum of contexts shown in Figure 5.8. However, such a split is not attested.

(3) Hypothetical Phonological Split

High F1: ika, uka, ikO, ukO, ukE, uke, uko, … ube

Low F1: ubẹ, ubọ, ubu, uba, iba, ubi, uki, ike, … ibe.

Contrast this with the corresponding graph for the Àkùrẹ́ speakers, shown in Figure 5.9, where the split between the phonologically different contexts is clear, with [−ATR] high vowels having a categorically higher F1 than [+ATR] high vowels.

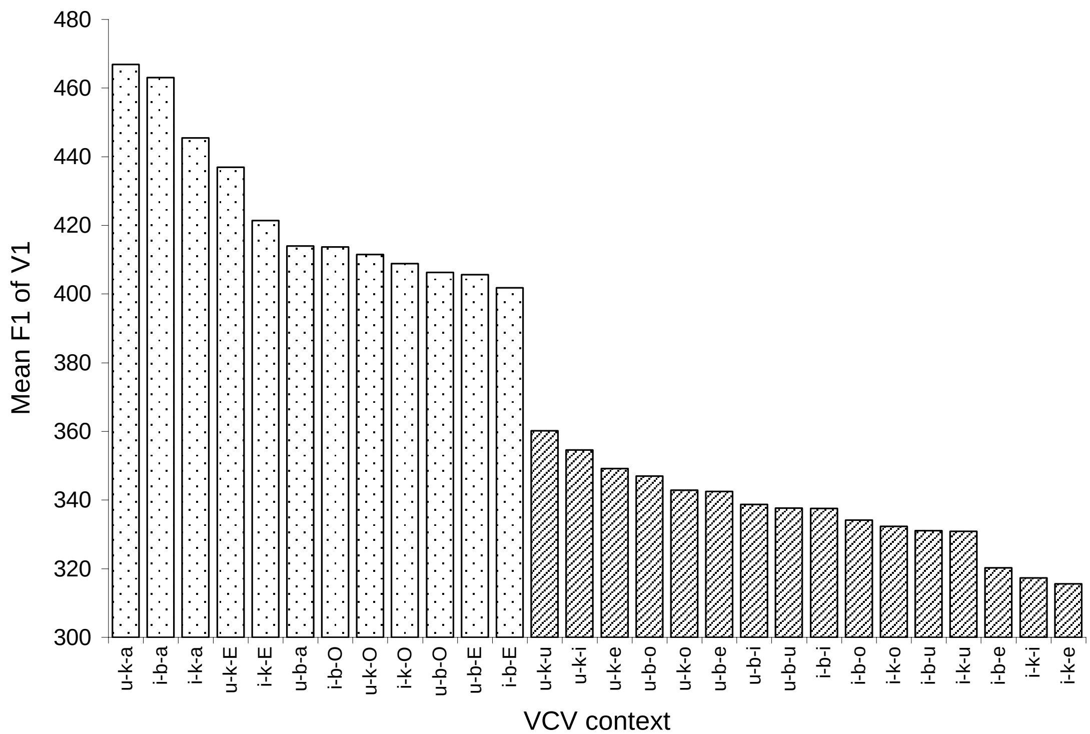


Figure 5.9. F1 means for midpoint of V1 in VCV tokens in different contexts for three Àkùrẹ́ speakers. Dotted bars indicate [−ATR] V1s, striped bars indicate [+ATR] V1s.

The asymmetry exists between the phonetic pattern and the phonological pattern. This type of asymmetry is addressed by Hayes (1999), who discusses the asymmetry between the phonetics and phonology of obstruent voicing across languages. Using an aerodynamic model (Keating 1984), Hayes determines that the degree of difficulty in the voicing of stops varies depending on both the identity of the stop and its context. However, the pattern in the phonetics was not mirrored in the phonology. If a phonological constraint banning voiced stops were motivated purely by aerodynamic factors, then it might look like the constraint in (4), since these are the contexts where the model predicts that voicing a stop is most difficult.

(4) Hypothetical Phonological Constraint (from Hayes 1999)

No post-obstruent voiced stops, no [d, g] in initial position, and no [g] after oral sonorants.

However, such a complicated—yet aerodynamically motivated—constraint is not found. The conditions in phonological constraints do not mirror closely the patterns in phonetics. Instead, one finds constraints such as (5), illustrating that phonological constraints are more symmetrical than their phonetic counterparts.

(5) Existing Phonological Constraints (from Hayes 1999)

No voiced obstruent word-finally (Polish)

No voiced obstruent after another obstruent (Latin)

No voiced obstruent geminates (Japanese)

No voiced velar obstruents (Dutch)

Hayes (1999:274) argues that learners create phonological constraints from phonetically grounded experience in articulation and perception, and that in going from the phonetics to phonology "learners execute a trade-off between phonetic accuracy and formal simplicity." While I have not examined articulation or perception, the acoustics of this study are obviously linked to perception. Hayes's claims are supported by the data from this study. The **phonological** realization of F1 is symmetrical, as shown in Figure 5.9. For /i/ and /u/ alike in Àkùrẹ́, the F1 differs largely dependent on whether the following vowel is [+ATR] or [−ATR]. However, the realization of F1 in high vowels in Mọ̀bà, shown in Figure 5.8, is not symmetrical; it varies greatly depending on the context, both the consonantal and vowel contexts, in fact.

Of course, a learner could conceivably posit an elaborate grammar where /i/ has a different acoustic realization before each consonant and before each different vowel that follows a consonant, or any number of complex patterns. While this grammar, shown in (6), would be more precise in modeling the data than the previous models in (1) and (2), it is also unequivocally more complex, barring any additional existing mental structure.

(6) Characterization of complex grammar

/i/ is realized as $[i]^x$ before ka

/i/ is realized as $[i]^y$ before ba

/i/ is realized as $[i]^z$ before ko, etc…

Why learners choose a simpler model in (2) over a more complex model in (6) is beyond the scope of this study.[5] However, we know that phonological rules tend to be more general than models such as (6)—and so we know that precision is not the primary factor in determining a learner's grammar. On the other hand, we cannot claim that solely because (2) is simpler than (6), that it is preferable, for then the learner would never posit the grammar in (2) over the simpler grammar in (1). Since we know that models such as (2) (or an equivalent) are posited, then we know that simplicity, like precision, is not the primary factor.

---

[5] Pierrehumbert (2001a) argues that phonological constraints are more general because speakers must have similar grammars; more fine grained constraints would not hold over numerous speakers and so could not be readily learned by all learners.

It is easy to see why some grammatical patterns would not be posited from the data. For example, a generalization such as in (7) could not be made because the acoustic data do not support the generalization.

(7) Unlikely generalization

/i/ → $[\mathrm{i}]_{\mathrm{HighF1}}$ / ____ C [a, ẹ, i]

/i/ → $[\mathrm{i}]_{\mathrm{LowF1}}$ / ____ C [ọ, u, e, o]

Similarly, the learner would never posit the generalization in (8), an anti-harmony system, for the same reason. An observant learner could not make the generalization that the vowels {[i], [u], [o], [e]} trigger F1 raising,

(8) Anti-harmony generalization

/i/ → $[\mathrm{i}]_{\mathrm{HighF1}}$ / ____ C [i, u, e, o]

/i/ → $[\mathrm{i}]_{\mathrm{LowF1}}$ / ____ C [a, ẹ, ọ]

While certain generalizations can be ruled out, there remain several that could fit the data. For example, the rule in (9) fits the data well, but then so does the rule in (10).

(9) /i/ → $[\mathrm{i}]_{\mathrm{HighF1}}$ / ____ C [a, ẹ, ọ]

/i/ → $[\mathrm{i}]_{\mathrm{LowF1}}$ / ____ C [i, u, e, o]

(10) /i/ → $[\mathrm{i}]_{\mathrm{HighF1}}$ / ____ C [a]

/i/ → $[\mathrm{i}]_{\mathrm{LowF1}}$ / ____ C [ẹ, ọ, e, o, i, u]

In the next section, I introduce an algorithm which I use to determine how to split the data set.

5.4. Decision tree model

A learning mechanism was employed using a simple decision tree learning algorithm in order to determine how the data might reasonably be split up, and thus whether the data support one grammar over another. The input to the algorithm consists of acoustic data and a list of questions. In the acoustic data, each record represents one instance of a vowel token; a few sample records with a header are shown in (11).

(11)	C	V2	F1
	b	i	293
	b	e	234
	b	e	252
	b	ẹ	271
	k	e	279
	k	ọ	393
	k	u	267

Each line represents an iCV token from the main experiment of this study. The first field of the record is the intervening consonant, either /b/ or /k/; the second field is the second vowel, one of /i, e, ẹ, a, ọ, o, u/; and the third field is the F1 value of the initial vowel /i/ at the middle of the vowel. The second input into the program is a list of so called questions. An example of a set of questions, shown in (12), helps to explain them.

(12)	CONSONANT	b			
	VOWEL2	ẹ	a	ọ	
	VOWEL2	i			
	VOWEL2	i	e	ẹ	a

For each data point in (11), one can ask one of these questions in (12): either the data point has the attribute stated in the question or it does not. For example, for the first question, {CONSONANT b}, some data points contain a consonant /b/, all other data points do not. Similarly, for the next question, either the data point contains a V2 with an /ẹ/, /a/, or /ọ/, or it does not. Every question divides the data into two sets, although one set may be empty in which case the question is effectively ignored. Each question has a complementary question which states the same thing in a different way. So if the data contain only /b/s and /k/s, it would be redundant to include the question {CONSONANT k}, if the question {CONSONANT b} were included.

The decision tree learning algorithm determines which one of the given questions best divides up the data, under certain assumptions. A metric is used to determine the relative strength of each split. Each question splits the data into two groups, called x and y. For each question, the value of the split is determined by the formula in (13), which measures the sum of the squares of the distance of each point in a set from the set's mean.

(13) $$\sum_{i=1}^{n}(x_i - \overline{x})^2 + \sum_{i=1}^{n}(y_i - \overline{y})^2$$

For a question that splits the group well, the value will be lower. From all the questions, the one that produces the lowest value is the best split of the data. After the first pass, the remaining two groups can be recursively split in the same manner until some criteria are met.[6] I am primarily interested in the first question found in the data, which is intended to model how a learner might split up the data. There are no doubt better algorithms to approximate more closely how a learner might group data, for

[6] The algorithm is considered greedy; that is, at any point, the best split at that level is the one that produces the lowest value at that point. An algorithm that is not greedy would recursively exhaust all possible lower splits to determine the lowest value for the current split. A greedy algorithm is simpler, much less computationally intensive, and suffices for the purposes of this preliminary examination into decision tree learning.

example, one that uses a perceptual based scale such as mels, instead of the acoustic scale in Hz., but the current method is adequate for our purposes here.

For the first trial, /i/ in Àkùrẹ́ provides a good starting point. The data for Àkùrẹ́ /i/ are categorical and are thus clearly split into two sets, as shown in Figure 5.9 above. The problem that is modeled here is, given the different realizations of /i/, what is the best way to split it into two? Note that the model does not say whether a split into two, or more, parts is warranted, but rather, if there is to be a split, which question best models the split. All the data for /i/ in the three speakers of Àkùrẹ́ were combined as input into the program.

The choice of which questions to use constrains the outcome of the modeling—the algorithm can only choose the best question of the ones provided. Initially, the questions input into the program were those shown in (14).

(14)	VOWEL2	i				
	VOWEL2	e				
	VOWEL2	ẹ				
	VOWEL2	a				
	VOWEL2	ọ				
	VOWEL2	o				
	VOWEL2	u				
	VOWEL2	ẹ	ọ	a		# −ATR
	VOWEL2	i	u			# HIGH
	VOWEL2	i	e			# +ATR, −BACK
	VOWEL2	u	o			# +ATR, +BACK
	VOWEL2	a	ọ	o	u	# BACK
	VOWEL2	i	e	ẹ	a	# FRONT + a

The questions in this initial group include references to the second vowel. The vowel questions are also limited to individual vowels and sets of vowels typically associated with natural classes based on common features, which, where relevant, are indicated as a comment at the end of each line, following the “#”. In a view of phonology proposed by Chomsky and Halle (1968:400) in their epilogue and expanded in works on feature geometry (see §1.2 for discussion and references), the innate features and feature hierarchy limit the types of phonological patterns that a learner may acquire. The first group of questions models this assumption, because I only include the vowel sets that form natural classes, and these natural classes—in a feature geometry view—are a reflex of the feature geometry. For example, in the first pass, the set {a, e, u} is not present precisely because there is no feature that defines that set. My claim is that a pre-existing feature structure is not required, and so a different group of questions is used in the second pass. This list of questions consists of all possible combinations of vowels, a much less constrained set of questions compared with the first set. The reason for starting with the constrained set is to see first how the data split occurs with the constraints of features, and then to try the same data with an unconstrained set of questions, a set with all possible combinations of questions. If both sets of questions result in the expected pattern, then by Occam's Razor the more general case would suffice to model the phonologization. That is, it would not be necessary to posit *a priori* features and feature structure in order to model the learning of this vowel harmony.

First, I present the results with the feature constrained list of questions. Of the entire 126 tokens, the mean F1 was 369.1 Hz. According to the algorithm, the best split was with the question [−ATR]: of the 54 data points with V2 of /ẹ, ọ, a/, the mean F1 of 425.2 Hz; and of the 72 data points with V2 of /e, o, i, u/, the mean F1 of 327.0 Hz. In other words, of the tokens of /i/, when constrained by questions using natural classes, the data split into those before [−ATR] vowels and those before [+ATR] vowels. This is not surprising, since the Àkùrẹ́ data fall very clearly into two groups, as we have seen.

A second run of the algorithm employs the relatively unconstrained set of the 64 questions consisting of all possible combinations of V2, of which a subset is shown in (15).

(15)	VOWEL2	i			
	VOWEL2	e			
	VOWEL2	ẹ			
	VOWEL2	a			
	VOWEL2	ọ			
	VOWEL2	o			
	VOWEL2	u			
	VOWEL2	i	u		
	VOWEL2	i	a		
	VOWEL2	i	ọ		
	VOWEL2	i	u	e	
	VOWEL2	i	u	ẹ	[. . .]

The results were the same as for the constrained set—the best initial split was with the question splitting [−ATR] and [+ATR] vowels. Similar results are found for /u/ in Àkùrẹ́ for both sets of questions. Again this is not surprising, for the same reasons mentioned before. However, it is both noteworthy and obvious that according to this model, Àkùrẹ́ learners can acquire the harmony directly from language data without a pre-existing notion of features.

Next, the same procedure is applied to the other dialects, where high vowel harmony does not occur. The data for /i/ in the six speakers of Mọ̀bà and SY were combined. The initial questions were those in (14), the questions using natural classes. Of the entire 325 tokens, the mean F1 was 293.2 Hz. The best split was found with the question [−ATR], with 140 data points with V2 of /ẹ, ọ, a/, mean F1 of 304.1 Hz, and 185 data points with V2 of /e, o, i, u/, mean F1 of 285.0 Hz. In other words, of the tokens of /i/, when constrained by questions using vowel features, the data points split into those before [−ATR] vowels and those before [+ATR] vowels. While this is the same result as for Àkùrẹ́, it is more interesting because in this case, the difference between the two sets is much smaller with a great deal of overlap. The second unconstrained set of questions produced the same results—the best initial split was with the question splitting [−ATR] and [+ATR] vowels. This is an important finding, suggesting that, in this case, pre-existing abstract features were not necessary for the learner to extract a phonological pattern from a coarticulatory pattern.

I have shown that the phonetic patterns of coarticulation in /i/ contain patterns found in the phonology. However, as was mentioned earlier, the /u/ does not pattern as nicely. When the 168 data points for /u/ for four Mọ̀bà speakers are combined and entered into the program with all of the V2 questions, the [ATR] split is not the initial split; in fact the question that produces the best split is the high vowels, /i/ and /u/. As discussed above, it is not clear why the /u/ does not pattern as clearly as /i/, whether it is something about the Mọ̀bà /u/ which is reduced, or there is something inherent in /u/s in general that leads to this. When the data points for /i/ and /u/ for Mọ̀bà speakers are combined, the best split is the vowels /i/, /u/, and /e/, which still does not support the hypothesis of this section. However, when that question is removed, the best split is the [ATR] question, suggesting that further study may provide results consistent with the results of /i/ alone.

A final run of the decision tree learning program, again using Mọ̀bà and SY /i/, included a third set of questions. This time, the intervocalic consonants were included in the questions in addition to all possible vowel questions. The best initial split was the consonant, that is, the split was between those tokens with intervocalic /k/ versus those with /b/. When /i/ preceded a /k/, the F1 average for 162 tokens was 302.9 Hz, while the F1 average before a /b/ for 163 tokens was 283.7 Hz. At first glance, this result implies that we might predict a phonological pattern in which the F1 of /i/ raises before a /k/ compared with before a /b/, as in (16).

(16) /i/ → $[i]_{HighF1}$ / ______ k V

/i/ → $[i]_{LowF1}$ / ______ b V

Because this type of pattern is not found in Yorùbá, one might propose that the reasoning taken up here is not valid. However, I suggest a possible explanation as to why we might not expect the consonants to have the best split in a more representative data set. In a more extensive study many more consonants would be present, resulting in a continuum of F1 averages for V1 /i/ depending on the intervocalic consonant, as can be seen for Mọ̀bà speaker Mb10 in section 4.3.2.1. With more consonants—and the same number of vowels—the influence of any one consonant would be reduced. This hypothesis could easily be tested in an acoustic study.

5.4.1. Exception driven strategy

An unstated assumption in the decision tree analysis above is that the model learner notes all of the input data points with an F1 value (in addition to context), so that each point enters into the learning algorithm with a quantitative value. This strategy can be referred to as an **average driven strategy**. The model determines the average F1 value for the vowel for different contexts, and determines the context with the tightest fit. This is what the algorithm above did. I suggest an alternative strategy a learner might use for positing the sets. In an **exception driven strategy**, the learner observes that certain [i]s, for example, had a particularly high F1—the exceptions—and subsequently determines the generalization that best captures the contexts in which these aberrant [i]s are uttered.[7]

An analogy may help distinguish the two strategies, namely how we determine what characteristics typify bad drivers. An observer of traffic might classify drivers

[7] This strategy resembles the **warping of perceptual space** concept proposed by Boersma, Escudero and Hayes (2003).

into normal drivers and bad drivers. When a bad driver is identified, the observer notes the characteristics about the situation such as the color of the car, the make of the car, the state on the license plate, the age and sex of the driver, whether the driver is wearing a hat, what bumper stickers are present, even the phase of the moon. After observing many bad drivers, if a characteristic has been repeated often enough, then the observer makes the generalization that, for example, people with Massachusetts plates are bad drivers. The fact that the observer has seen some drivers from Massachusetts who are not bad drivers does not change the generalization formed about Massachusetts drivers. This would be an exception driven strategy. Alternatively, the observer could use an average driven strategy by determining characteristics regarding every driver, including whether or not they are a bad driver. She could then determine an average ratio of bad to good drivers for each characteristic; the characteristics with the highest ratios would then be considered the characteristics of bad drivers. The exception driven strategy has intuitive appeal both for this driving scenario and for vowels—it is perhaps easier to notice aberrant characteristics than normal ones. I do not model an exception driven strategy here; a future study could test this strategy using z-scores calculated for each speaker. My intent here has not been to determine the precise way such learning would take place, as this too is beyond the scope of this dissertation, but rather to schematize possible scenarios for how a language might change due to phonetic influence.

5.5. Grammar change

We can consider that for any given phonological grammar, A, there is a certain probability of a change to another grammar, B, over some arbitrary period of time. The probability of change can not be the same for each grammar; to go from Grammar A to grammar C might be less probable, as illustrated in Figure 5.10.

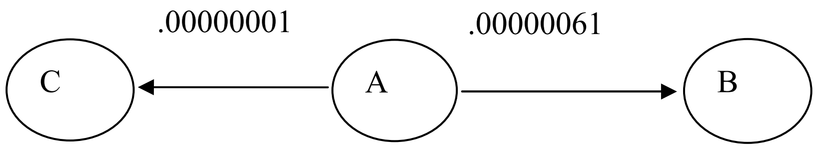


Figure 5.10. Grammars A, B, C with probabilities of change from one grammar to another for some arbitrary time.

Each ellipse represents a possible grammar. The arrows and the corresponding ratios represent the probability of a change from one grammar to another within an arbitrary period of time, say one generation, with arbitrary values for probabilities. This figure shows that if a language currently has grammar A, there is a higher probability of changing to grammar B than to grammar C.

Now turn to Figure 5.11.

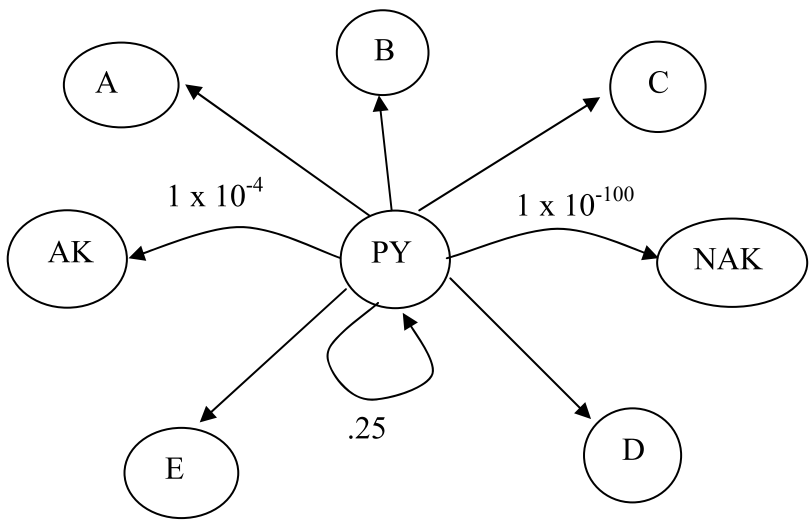


Figure 5.11. A language with grammar PY (Proto-Yorùbá like) is more likely to change to grammar AK (one like Àkùrẹ́) than to grammar NAK (grammar with anti-harmony system); it is even more likely to stay the same.

PY represents a grammar of Proto-Yorùbá, assuming it to be similar to the grammar of Mọ̀bà in vowel harmony as I have claimed in Chapter 2. Grammar AK is the grammar like Àkùrẹ́'s, that is, one with high vowel harmony. Grammar NAK is a grammar with an anti-harmony system, as in (8) above. The numbers represent probabilities of going from one grammar to the other, and as above, the numbers are contrived for this illustration. If all the possible grammars were included, the probabilities would add up to one. The probability of going from a PY grammar to an AK grammar is relatively low—indeed this has possibly only happened once, in the Central Èdè dialects. The probability of change in the direction of an NAK grammar is virtually zero, for the reasons mentioned above—that is, because the data do not support it. In this way, the space of conceivable grammars is greatly reduced simply because the probability of getting to most grammars in the space is virtually zero. Note that the grammar space is reduced, not solely because of limitations in the cognitive system,[8] but because language change, as mediated by phonetics, will only lead to a small subset of the grammar space.

5.6. Historical account in Proto-Yorùbá

I have shown that there are patterns in coarticulation that are similar in many respects to the patterns of vowel harmony, and that elements of the vowel harmony pattern can be extracted from the patterns of coarticulation using a simple decision tree learning algorithm. These findings provide strong evidence that [ATR] high vowel harmony has its origin in coarticulatory effects.

While this study is not primarily an historical study of the Yorùbá language group, the findings lend support to a particular view of the historical development of the Yoruboid languages. In chapter two (§2.8), I argued for a classification of the historical development of the Yorùbá languages, shown again in (17), using evidence from distribution of vowels in existing languages.

---

[8] Although cognitive limitations must certainly also impact on the grammar space.

(17) Proposed classification of Mọ̀bà with NW and Central Èdè dialects (from chapter 2, (82).

Proto-Yorùbá (7 surface oral vowel vowels)

NW Èdè (7)
SY
Ẹ̀gbá
Ọ̀yọ́
Ọ̀ṣùn

Proto-Central Èdè (7)

Mọ̀bà (7)

Central Èdè (9)
Àkùrẹ́
Ìfẹ̀
Ìjẹ̀ṣà
Èkìtì

This view is strengthened as a result of the findings presented in this chapter. Of particular interest is the innovation from a system of seven surface oral vowels in Proto-Yorùbá and Proto-Central Èdè to a system of nine surface oral vowels with high vowel harmony in some Central Èdè dialects, while the remainder of the dialects maintain seven surface oral vowels. The results of the phonetic study and the analysis presented here provide an explanation of how a transition from seven to nine surface vowels might have occurred.[9] The high vowel harmony found in Àkùrẹ́, and other similar systems, looks like a phonologization of vowel-to-vowel coarticulation in an earlier language. Moreover, in the specific case of Àkùrẹ́ and closely related Central Èdè dialects, the results suggest that high vowel harmony emerged from coarticulation

[9] If it were found, contrary to the evidence I present, that Proto-Yorùbá already had a nine surface oral vowel system with high vowel harmony, the central claim of my study—that vowel harmony emerges from coarticulation—would not be invalidated. Such a finding would however invalidate the classification in (17).

in Proto-Central Èdè, or at the very least the results suggest that such a transition is plausible.

5.7. Coarticulation rooted in the physical domain

If the origin of phonological vowel harmony lies in the phonetic domain, one may wonder if the problem has been shifted from one domain to another, without additional insight. This might be the case if there were no explanation for coarticulation. However, coarticulation is rooted in articulation and planning, at least in part due to the task of moving a mass, the tongue and other articulators, through space and time to make more or less discrete sounds (Kühnert and Nolan 1999, Whalen 1990). If the vocal tract were turned off between phones, and there was adequate time to position the vocal tract for each sound, then we might expect that there would be no coarticulation. However, speakers do not necessarily speak slowly, and importantly they continue to emit sounds during the transition between phones. Because of this, the transitions between phones are necessarily part of the speech stream. In addition, the articulatory realization of consonants varies depending on context, in this case the adjacent vowels. When the articulation of the consonant changes, this bleeds into adjacent segments, in this case vowels. This is nicely demonstrated by Figure 5.12, based on x-ray drawings from Öhman (1966), which shows different tongue positions for the consonant [g]; in one case the /g/ is before a high front rounded vowel [y] (dashed line), in the other case the /g/ is before a low vowel [ɑ] (solid line).

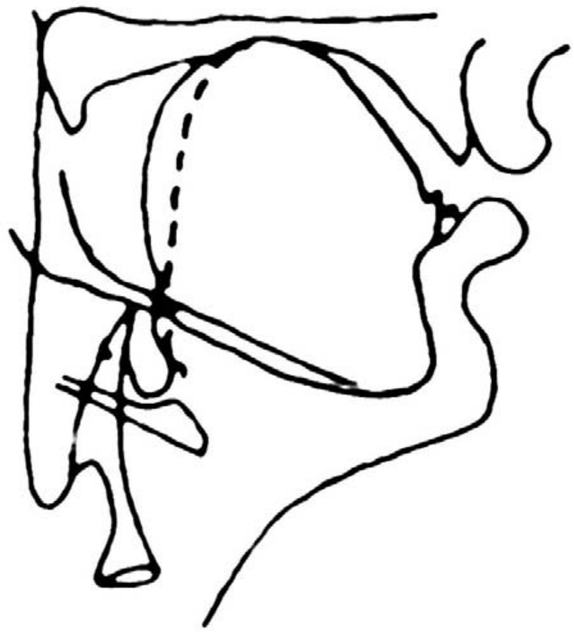


Figure 5.12. The /g/ of /ygɑ/ (solid line) and the /g/ of /ygy/ (dashed line) a few milliseconds after closure from a tracing from an x-ray motion picture, following Öhman (1966).

In both cases the preceding vowel is the high front rounded /y/. The differing position of the tongue in /g/ affects the realization of the preceding vowel /y/. This can be seen in the spectrogram tracings from averages of the formants for both VCV words, shown in Figure 5.13.

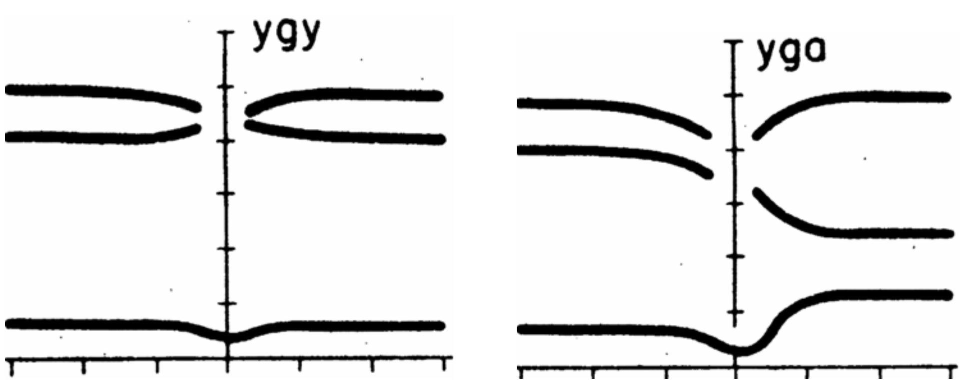


Figure 5.13. Averages of tracings from spectrograms /ygy/, left, and /yga/ in one Swedish speaker, following Öhman (1966).

Note in particular that the decrease of F2 as the /y/ of /yga/ goes into the /g/ corresponding with the backness of the tongue shown in the /g/ of /yga/ in Figure 5.12, above.

While some languages exhibit more vowel to vowel coarticulation than others, it is implausible for a language to exhibit no coarticulation, at least in the edges of adjacent segments, due to the constraints imposed by moving physical entities from one point to another. Ohman (1966) found no vowel-to-vowel coarticulation in Russian, but Choi and Keating (1991), using a more sensitive methodology, did find vowel-to-vowel coarticulation, albeit weak. This study is not a study of articulation, and I have not directly measured the vocal tract. However to some degree one can infer the vocal tract's shape from the acoustics. In the coarticulation examined in Mọ̀bà and SY, the F1 is higher in front of a set of vowels with a higher F1 than in front of the rest of the vowels, consistent with the idea that the vocal tract is preparing for the consonant and the following vowel during the initial vowel. Although much remains to be understood in the study of coarticulation, the coarticulation found in Mọ̀bà and SY is firmly rooted in the physical domain.

5.8. Final words

I have provided evidence that suggests ATR high vowel harmony in Àkùrẹ́ Yorùbá has its roots in coarticulation. The coarticulatory patterns contain the information necessary for a learner to infer a vowel harmony pattern given fairly general assumptions, without an *a priori* notion of features. Recall from chapter 1, that universal grammar is an argument of last resort. Anderson (1976) acknowledges that UG can only be posited after the effects of non linguistics domains have been ruled out. If the analysis presented here is correct, then certain elements traditionally considered part of UG, can be reconsidered as originating in phonetics.

It is true that the case for /u/ is not as strong as the case for /i/ with regard to the phonetics coming from phonology. This is unfortunate for my study, and weakens the case. One possible explanation is that the phonologization process is a hybrid of phonetic and *a priori* abstract features. However, unless these same inconclusive results are replicated for other potential phonetics-phonology doublets, the more likely alternative is that the methodology was not sufficient to attain more general results.

Assuming /u/ is found to pattern sufficiently similarly to /i/, the phonological rule that emerges may be formulated as in (18).

(18) [+high] → [−ATR] / ____ C [−ATR]

(As before, the choice of rule over constraint is a matter of convenience.) However, the specific choice of the vowels in the rule does not actually require an independent grouping called [−ATR]; rather the set /a, ẹ, ọ/ is extracted from the phonetics as the group that triggers the change, as was shown in the analysis. Recall that those vowels were chosen regardless of whether the choice of questions was limited to natural classes. Therefore, it may be sufficient to state the harmony rules as in (19).

(19) /i/ → [ị] / ____ C [a | ẹ | ọ]
/i/ → [i] / elsewhere

/u/ → [ụ] / ____ C [a | ẹ | ọ]
/u/ → [u] / elsewhere

It is outside the scope of this dissertation to determine the degree of generality that the learner imposes on phonological patterns. I am in no position to make the claim that (19) is more correct than (18). While I have shown evidence to suggest that *a priori* feature [ATR] is not required in a learner's inference of a phonological pattern, I make no claim about the ultimate level of generalization. Indeed, in the same way that I have generalized the class of consonants in both (18) and (19), it seems intuitively likely that the classes of trigger vowels and target vowels are generalized by the learner. However, I do claim that these classes or features are formed as a result of experience. In the case at hand, the [−ATR] set of /a, ẹ, ọ/ is induced from the acoustic data to which the learner is exposed, both for current learners of Àkùrẹ́, as well as for learners of proto-Central Yorùbá at the time of the formation of high vowel harmony. That is, if phonology originates from the phonetics, we do not have to posit a pre-existing independent structure—the feature—to account for the patterning of the phonology; the patterning can be induced initially from the phonetics, and at a later stage from the phonology of the speakers.

Finding a phonetic explanation for a phonological phenomenon does not reduce phonology to phonetics. Once phonologized, the pattern becomes a pattern in the mind, diachronically—but not synchronically—rooted in the phonetics. (Although it should be noted that coarticulation must have a mental component as well, since coarticulation differs from language to language.) Of course, the forces of phonetics continue to act on all speech sounds, since human speech sounds are uttered in the physical world. For example, coarticulatory effects continue to be seen in Àkùrẹ́—note the higher F1 of /i/ before /a/ compared with before /ẹ/ and /ọ/ in Figure 5.9. Beddor and Yavuz (1995) find that in Turkish, a language with vowel harmony, coarticulation is found in disharmonic words. Because the forces of change continue to act on all languages, phonological patterns continue to change in ways that may ultimately result in patterns that may not be considered phonetically naturally, such as /r/ insertion in English (Hale and Reiss 2000).

As discussed in Chapter 4 (§4.3.2.1), we expect the degree of vowel-to-vowel coarticulation to differ from one language to the next due in part to the differences in characteristics of the phoneme inventory. In languages with fewer vowels, with fewer or no consonant clusters, or with fewer consonants, for example, we expect, all else being equal, to find more vowel to vowel coarticulation than in languages with a crowded vowel system, many consonant clusters, or many consonants. For the same reason, vowel harmony of various types is arguably more likely to develop in some systems over others.

In chapter 1, I argued that SPE and some other phonological theories use the existence of universal patterns as an argument for an articulated (and innate) Universal Grammar. The UG, consisting of rules, conventions, and/or constraints, is said to reduce the types of grammars available to all language learners, and by the same token, all speakers, resulting in the universal patterns which are evident in languages. I have shown in this study, following work of Ohala (1990, 1993, 1994, 1995, and elsewhere), Myers, (1997, 2000b), Hale (1999), Hale and Reiss (2000), Silverman (2000), and Buckley (2000), that at least some of the universal patterns instead emerge as a result of sound change, the direction of which is influenced by phonetic phenomena. Thus, the naturalness of phonology is not inherent to phonology. Myers (2000b) states: “to express phonetic naturalness in phonology, it isn’t necessary to use phonetically natural constraints, since the correspondence between the two domains is guaranteed by the diachronic process of phonologization.”

In order to keep this study manageable, I restricted the experiment in several ways. Now that my hypothesis has been confirmed on a restricted set, further studies might expand the study to include a larger set of conditions, for example with more intervocalic consonants, to more closely approximate the input to which a learner might be exposed. Further studies could examine the vowel /a/, to determine how coarticulation could lead to the low vowel participating in vowel harmony. It is expected that the coarticulatory effects on /a/ will be weaker than they were for the high vowels, since /a/ allophony in ATR harmony is not as common as high vowel allophony. In addition, while I have proposed how learners could extract high vowel allophonic patterns from coarticulatory patterns, I have not suggested how these allophones then become triggers for further harmony. Finally, I expect that other dialects of Yorùbá similar to SY and Mọ̀bà would exhibit similar coarticulatory patterns. Other languages with similar surface vowels, such as Defaka (Ijoid, Niger-Congo) (Shryock et al. 1996), should also exhibit coarticulatory patterns similar to those found for Mọ̀bà and SY. A language like Italian is not considered an ATR language, but like SY and Mọ̀bà it has seven vowels. (See Disner (1983) for a comparison of vowels in Standard Yorùbá and Italian.) An acoustic study of coarticulation in Italian vowels would shed light on whether the patterns found in this study are found in languages that do not exhibit ATR type harmony.

APPENDIX

Subject Questionnaire

Date:

Name:

Age:

Where were you brought up? What language(s) was/were spoken in your home?

How long did you live there?

What language was spoken in the town where you lived?

What language(s) is/are spoken in your home currently?

Where else did you live in Nigeria and outside of Nigeria?

What other languages do you speak?

Word Lists

Àkùrẹ́ word list:[1]

1. igi
2. epo
3. ake
4. agbo
5. ata
6. abo
7. ọkọ
8. ibọ
9. ẹkẹ
10. abọ
11. ike
12. aku
13. ẹba
14. abu
15. ọba
16. akọ
17. ukọ
18. oke
19. uku
20. eki
21. abi
22. uko
23. ọka
24. ubu
25. ako
26. ẹbọ
27. aba
28. oko
29. aki
30. eke
31. uba
32. ubi
33. ibe
34. uki
35. ọkẹ
36. ube
37. ẹka
38. iki
39. ọbẹ
40. ibu
41. obe
42. oki
43. ọbọ
44. ẹbẹ
45. ibo
46. eku
47. obu
48. ẹkọ
49. uka
50. ebe
51. ibi
52. iku
53. ika
54. obi
55. eko
56. ubọ
57. ibẹ
58. iko
59. ebo
60. uke
61. akẹ
62. ubẹ
63. obo
64. ake
65. ikọ
66. abe
67. ebu
68. oku
69. ubo
70. aka
71. abẹ
72. ikẹ
73. ukẹ
74. ebi
75. iba
76. ọkẹ
77. ubo
78. obe
79. ubọ
80. ako
81. ube
82. akẹ
83. ibi
84. ubu
85. oke
86. abẹ
87. ọba
88. oki
89. abi
90. eko
91. ukẹ
92. ika
93. aki
94. ibu
95. ike
96. aba
97. eke
98. ake
99. ikẹ
100. iko
101. aku
102. iki
103. ẹkẹ
104. ẹkọ
105. obu
106. ukọ
107. ẹbẹ
108. ọkọ
109. ẹbọ
110. uka
111. ebu
112. uku
113. ibẹ
114. oko
115. ibọ
116. abu
117. ubi
118. ẹba
119. uki
120. ebi
121. ọbẹ
122. aka
123. uba
124. ubẹ
125. ọbọ
126. obi
127. akọ
128. abo
129. ibo
130. ẹka
131. iku
132. ibe
133. obo
134. uke
135. ebe
136. uko
137. oku
138. eku
139. eki
140. abe

[1] Words 1-5, 286-288 are filler words and are not measured.

141. ebo
142. ikọ
143. ọka
144. abọ
145. iba
146. ọkọ
147. obo
148. obe
149. aba
150. abọ
151. ọkẹ
152. ọbọ
153. uku
154. ubo
155. ake
156. obi
157. eku
158. uki
159. ọka
160. abẹ
161. oki
162. ibu
163. ẹka
164. ẹkọ
165. ubu
166. ukẹ
167. ubọ
168. iba
169. ukọ
170. uba
171. obu
172. ikọ
173. ẹba
174. oko
175. iko
176. uka
177. ẹbọ
178. aki
179. ube
180. ebo
181. ẹkẹ
182. ọba
183. ibẹ
184. ẹbẹ
185. ako
186. akẹ
187. ọbẹ
188. abu
189. ikẹ
190. ibọ
191. ubi
192. ika
193. abo
194. uke
195. aka
196. ebu
197. aku
198. eko
199. ebe
200. oku
201. abe
202. ebi
203. ibo
204. eke
205. oke
206. abi
207. akọ
208. iki
209. ubẹ
210. ike
211. ibi
212. uko
213. eki
214. iku
215. ibe
216. ibu
217. iko
218. ibọ
219. ubẹ
220. ubu
221. eko
222. ubo
223. uku
224. ubi
225. ubọ
226. ẹka
227. eku
228. ibi
229. aki
230. obi
231. oku
232. ebo
233. iki
234. ube
235. obe
236. ikọ
237. abe
238. oki
239. ake
240. eki
241. ebi
242. aba
243. ọka
244. ẹbẹ
245. ẹkẹ
246. abo
247. iku
248. akẹ
249. abọ
250. ibẹ
251. abu
252. eke
253. uki
254. ọbọ
255. uka
256. ọbẹ
257. abẹ
258. oko
259. ẹkọ
260. ebu
261. aku
262. ẹba
263. ọkọ
264. uba
265. ibo
266. ikẹ
267. aka
268. ẹbọ
269. ike
270. ọkẹ
271. iba
272. oke
273. ukọ
274. uke
275. obu
276. ebe
277. akọ
278. ukẹ
279. abi
280. ibe
281. obo
282. uko
283. ika
284. ọba
285. ako
286. ata
287. igi
288. ẹja

Mọ̀bà word list:[2]

1. igi
2. epo
3. ake
4. agbo
5. ata
6. abo
7. ọkọ
8. ibọ
9. ẹki
10. ẹkẹ
11. abọ
12. ike
13. aku
14. ẹba
15. abu
16. ọki
17. ọba
18. akọ
19. ukọ
20. oke
21. uku
22. eki
23. abi
24. uko
25. ọka
26. ubu
27. ako
28. ẹbọ
29. aba
30. oko
31. aki
32. eke
33. uba
34. ọbu
35. ubi
36. ibe
37. uki
38. ọkẹ
39. ọku
40. ube
41. ẹka
42. iki
43. ẹbu
44. ọbẹ
45. ibu
46. obe
47. oki
48. ọbọ
49. ẹbẹ
50. ibo
51. eku
52. obu
53. ẹkọ
54. uka
55. ebe
56. ibi
57. iku
58. ika
59. obi
60. eko
61. ubọ
62. ibẹ
63. iko
64. ọbi
65. ebo
66. uke
67. akẹ
68. ubẹ
69. obo
70. ake
71. ikọ
72. abe
73. ebu
74. oku
75. ubo
76. aka
77. abẹ
78. ẹku
79. ikẹ
80. ẹbi
81. ukẹ
82. ebi
83. iba
84. oku
85. abe
86. oko
87. ake
88. ọbẹ
89. iki
90. ibe
91. eke
92. ako
93. ika
94. aki
95. obu
96. eki
97. ẹkẹ
98. obo
99. ẹbọ
100. ukọ
101. iko
102. aka
103. ọbu
104. ebi
105. ẹki
106. aba
107. obi
108. uku
109. ikẹ
110. ọkẹ
111. ibẹ
112. iku
113. ubu
114. ebu
115. ubi
116. ibu
117. abu
118. uke
119. ẹka
120. ubọ
121. uko
122. abọ
123. ọbọ
124. iba
125. ọkọ
126. ikọ
127. ọba
128. ube
129. akọ
130. ọka
131. akẹ
132. ọku
133. eku
134. oki
135. ubẹ
136. ike
137. ẹku
138. uba
139. aku
140. uki
141. abẹ
142. uka
143. ibi
144. abi
145. ẹbẹ
146. obe
147. abo
148. ẹkọ
149. ẹbi
150. ọki

[2] Words 1-5, 318-320 are filler words and are not measured.

151. eko
152. ebe
153. ibo
154. ẹbu
155. ẹba
156. ukẹ
157. ebo
158. oke
159. ibọ
160. ọbi
161. ubo
162. oke
163. ọkọ
164. ebu
165. abu
166. ibo
167. abi
168. obo
169. ẹkẹ
170. iko
171. uke
172. ike
173. ẹki
174. ukẹ
175. iki
176. ibi
177. ibẹ
178. ọbọ
179. uku
180. ọku
181. ibọ
182. ebo
183. aku
184. eku
185. ọka
186. ẹbu
187. ọkẹ
188. eke
189. ube
190. ubẹ
191. akẹ
192. abọ
193. obe
194. ubọ
195. aba
196. eko
197. oko
198. uko
199. uka
200. ebe
201. ẹka
202. ikọ
203. ibe
204. akọ
205. ẹba
206. obi
207. ake
208. aki
209. ọba
210. ọki
211. eki
212. abo
213. iku
214. abẹ
215. uki
216. ako
217. oki
218. ẹbi
219. ubu
220. oku
221. ubo
222. ikẹ
223. ẹkọ
224. ọbi
225. obu
226. ọbẹ
227. ika
228. abe
229. uba
230. ubi
231. ọbu
232. ẹbọ
233. ukọ
234. ibu
235. ẹku
236. ẹbẹ
237. iba
238. ebi
239. aka
240. uka
241. ọbẹ
242. ọba
243. ẹkọ
244. ika
245. eke
246. oki
247. ebu
248. eko
249. akọ
250. ọki
251. iki
252. uku
253. aki
254. ọbi
255. ebi
256. ẹkẹ
257. ẹbi
258. ẹki
259. iku
260. ibe
261. aka
262. iba
263. ọka
264. ikẹ
265. eku
266. ẹbọ
267. ibẹ
268. ọbu
269. abẹ
270. ako
271. ube
272. uke
273. aku
274. abe
275. ọkẹ
276. ọkọ
277. eki
278. ubi
279. ẹbu
280. ibo
281. uko
282. ẹba
283. ibi
284. oku
285. ake
286. ebo
287. ebe
288. abọ
289. obe
290. oko
291. ẹku
292. abi
293. oke
294. obo
295. ẹbẹ
296. ukọ
297. uki
298. ibọ
299. ukẹ
300. ikọ
301. ubu
302. ọku
303. ọbọ
304. ike
305. obi
306. ubọ
307. obu
308. uba
309. abu
310. ibu
311. abo
312. ubẹ
313. akẹ
314. aba
315. ubo
316. iko
317. ẹka
318. ata
319. igi
320. ẹja

SY Word List:[3]

1. ata
2. igi
3. ẹja
4. epo
5. ibe
6. ake
7. igi
8. agbo
9. akọ
10. ẹkẹ
11. ikẹ
12. ako
13. ike
14. ikọ
15. ọbi
16. aku
17. ẹku
18. ẹbẹ
19. oke
20. oko
21. ọkẹ
22. abe
23. ọkọ
24. ọki
25. oku
26. ibi
27. obe
28. obu
29. eke
30. ọbẹ
31. abo
32. akẹ
33. ọbọ
34. iku
35. ibọ
36. ẹki
37. eku
38. ebi
39. ake
40. aka
41. obo
42. ibu
43. ọka
44. ẹba
45. ẹbọ
46. abu
47. ibẹ
48. oki
49. iko
50. ẹka
51. iba
52. abẹ
53. iki
54. ika
55. eki
56. aki
57. ẹkọ
58. abi
59. ọku
60. ẹbi
61. ọbu
62. ibe
63. ebo
64. ebe
65. ibo
66. ebu
67. abọ
68. ọba
69. ẹbu
70. eko
71. aba
72. obi
73. ika
74. ọbi
75. ebi
76. aki
77. ọkọ
78. ake
79. ike
80. eki
81. ẹbẹ
82. oku
83. abo
84. oko
85. obo
86. ibu
87. ibi
88. ẹku
89. akẹ
90. abu
91. ibẹ
92. oki
93. ikẹ
94. obu
95. ọbọ
96. ẹbi
97. aka
98. aku
99. ọka
100. ọku
101. ẹkẹ
102. abe
103. iku
104. akọ
105. oke
106. iba
107. ọki
108. ako
109. ọbẹ
110. eke
111. ẹka
112. ebo
113. ibe
114. aba
115. ibọ
116. ebu
117. obi
118. abọ
119. obe
120. ibo
121. ẹbọ
122. ọba
123. ikọ
124. ẹki
125. iko
126. ẹkọ
127. abẹ
128. ẹba
129. ebe
130. eku
131. ọbu
132. ẹbu
133. iki
134. ọkẹ
135. abi
136. eko
137. ọki
138. ẹbi
139. ẹbẹ
140. akọ
141. ẹkọ
142. iki
143. aba
144. ako
145. ẹbọ
146. ike
147. abi
148. aki
149. obo
150. aka

[3] Words 1-5, 265-268 are filler words and are not measured.

151. abọ
152. ọbẹ
153. iba
154. ikẹ
155. ẹbu
156. obe
157. ebi
158. ibọ
159. abẹ
160. akẹ
161. ibẹ
162. aku
163. oku
164. oke
165. ẹkẹ
166. eke
167. ebu
168. ẹku
169. iko
170. ika
171. abu
172. ikọ
173. ake
174. ebo
175. obu
176. ẹka
177. ibi
178. ibe
179. ọba
180. obi
181. ọbi
182. ẹki
183. iku
184. abe
185. ọbu
186. oki
187. eku
188. ọbọ
189. ọka
190. ibo
191. oko
192. ọku
193. ebe
194. abo
195. ẹba
196. eki
197. eko
198. ibu
199. ọkọ
200. ọkẹ
201. ẹkọ
202. ibọ
203. oki
204. ibe
205. oke
206. ẹki
207. ake
208. abẹ
209. aku
210. ebe
211. ẹbẹ
212. abi
213. ẹbi
214. eku
215. ẹka
216. iku
217. abu
218. ọbọ
219. abo
220. ọkọ
221. ẹku
222. aba
223. obe
224. ọbi
225. aki
226. abe
227. ẹbọ
228. ebo
229. ọku
230. ọbu
231. ibẹ
232. abọ
233. oko
234. eke
235. akẹ
236. eko
237. ọkẹ
238. eki
239. akọ
240. ike
241. ọki
242. ebi
243. iko
244. obu
245. iba
246. ikọ
247. ẹkẹ
248. iki
249. ẹbu
250. ọba
251. obo
252. ako
253. ebu
254. aka
255. ẹba
256. ibu
257. obi
258. ibi
259. oku
260. ika
261. ikẹ
262. ọka
263. ọbẹ
264. ibo
265. ata
266. igi
267. ẹja
268. epo

REFERENCES

Abraham, R. C. (1958) *Dictionary of Modern Yoruba*, University of London Press, London.

Adetugbo, Abiodun (1967) *The Yorùbá Language in Western Nigeria: Its Major Dialect Areas*, Doctoral Dissertation, Columbia University, New York.

Ajiboye, Oladiipo (1991) *Àtúnyẹ̀wò Fonọlọjì Mọ̀bà*, MA thesis, University of Ilorin, Nigeria.

Akinkugbe, Olufẹmi Odutayọ (1978) *A Comparative Phonology of Yorùbá Dialects, Iṣẹkiri, and Igala*, Doctoral Dissertation, University of Ibadan, Nigeria.

Akinlabi, Akinbiyi (1992) "Two-tone analysis of Yoruba," *Research in Yoruba Language and Literature* 2, 59-76.

Akinlabi, Akinbiyi (in preparation) *A Phonological Grammar of Yorùbá*.

Akinlabi, Akinbiyi and Mark Liberman (1999) "The tonal phonology of Yorùbá clitics," presented at the Linguistic Society of America meeting, Los Angeles, California.

Andersen, Torben (1987) "An outline of Lulubo phonology," *Studies in African Linguistics* 18, 39-65.

Anderson, Stephen (1981) "Why phonology isn't 'natural'," *Linguistic Inquiry* 12, 493-539.

Archangeli, Diana and Douglas Pulleyblank (1989) "Yoruba vowel harmony," *Linguistic Inquiry* 20, 173-217.

Archangeli, Diana and Douglas Pulleyblank (1993) "Two rules or one… or none? (ATR) in Yoruba," in Joshua S. Guenter, Barbara A. Kaiser, and Cheryl C. Zoll, editors, *Proceedings of the Nineteenth Annual Meeting of the Berkeley Linguistics Society*, Berkeley Linguistics Society, Berkeley.

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

Armstrong, Robert (1965) "Comparative wordlists of two dialects of Yorùbá with Igala," *Journal of West African Languages* 2.2, 51-78.

Awobuluyi, A. Oladele (1967) "Vowel and consonant harmony in Yoruba," *Journal of African Languages* 6, 1- 8.

Awobuluyi, Oladele (1985) "On the reality of vowel coalescence in Yoruba," *Studies in African Linguistics* supplement 9, 11-14.

Bakovic, Eric (2000) *Harmony, Dominance and Control*, Doctoral Dissertation, Rutgers University, New Brunswick, New Jersey.

Bamgboṣe, A. (1965) "Assimilation and contraction in Yoruba," *Journal of African Languages* 2, 21-27.

Bamgboṣe, A. (1967) "Vowel harmony in Yoruba," *Journal of African Languages* 6, 268-273.

Bamiṣilẹ (1987) *Fọnọlọji Yorùbá-Mọ̀bà*, M.A. Thesis, University of Ilọrin, Nigeria [unavailable].

Bamiṣilẹ, Rẹmi (1994) "Justification for the survival of vowel coalescence as a phonological process in Yorùbá," *African Languages and Cultures* 7.2, 133-142.

Beckman, Mary, Kenneth De Jong, Sun-Ah Jun, and Sook-Hyang Lee (1992) "The interaction of coarticulation and prosody in sound change," *Language and Speech* 35, 45-48.

Beddor, Patrice Speeter and Handan Kopkalli Yavuz (1995) "The relation between vowel-to-vowel coarticulation and vowel harmony in Turkish," *Proceedings of the 16th International Congress of Phonetic Sciences*, vol. 2, 44-51, Stockholm.

Bendor-Samuel, J. (1989) *The Niger-Congo Languages*, University Press, London.

Bennett, Patrick R. and Jan P. Sterk (1977) "South central Niger-Congo: A reclassification," *Studies in African Linguistics* 8, 241-273

Bickley, Corine (1982) "Acoustic analysis and perception of breathy vowels," *MIT Working Papers in Speech Communication* 1, 71-81.

Boersma, Paul, Paolo Escudero and Rachel Hayes (2003) "Learning abstract phonological from auditory phonetic categories: An integrated model for the acquisition of language-specific sound categories," *Proceedings of the 15th International Congress of Phonetic Sciences*, 1013-1016.

Bradlow, Ann R. (1993) *Language-Specific and Universal Aspects of Vowel Production and Perception: A Cross-Linguistic Study of Vowel Inventories*, Doctoral Dissertation, Cornell University, Ithaca, NY.

Bradlow, Ann R. (1995) "A comparative acoustic study of English and Spanish vowels," *Journal of the Acoustical Society of America* 97:3, 1916-1924.

Browman, Catherine P. and Louis Goldstein (1992) "Articulatory phonology: An overview," *Phonetica* 78, 78-86.

Buckley, Eugene (2000) "What should phonology explain?," handout from a talk at SUNY Buffalo, NY.

Cahill, Mike (1996) "ATR harmony and feature geometry in Kɔnni," paper presented at the 27th Annual Conference of African Linguistics, University of Florida, Gainesville.

Capo, Hounkpati B. C. (1985) “On the high non-expanded vowels in Yoruboid,” *Studies in African Linguistics* 16, 103-121.

Capo, Hounkpati B. C. (1989) “Defoid,” in J. Bendor-Samuel, editor, *The Niger-Congo Languages*, University Press, London.

Casali, Rod (2000) “An introduction to ATR vowel harmony in African languages”, unpublished manuscript.

Choi, John D. and Patricia Keating (1991) “Vowel-to-vowel coarticulation in three Slavic languages,” *UCLA Working Papers in Phonetics* 78, 78-85.

Chomsky, Noam and Morris Halle (1968) *The Sound Pattern of English*, Harper and Row, New York.

Christaller, Johann Gottlieb (1875) *A Grammar of the Asante and Fante Language called Tshi (Chwee, Twi)*, Basel Evangelical Missionary Society, Basel (Reprinted in 1964 by Gregg Press, Ridgemont, New Jersey).

Clements, G. N. (1981) “Akan vowel harmony: A nonlinear analysis,” *Harvard Studies in Phonology* II, 108-177.

Clements, G. N. (1985) “The geometry of phonological features,”*Phonological Yearbook* 2, 225-252.

Clements, G. N. (1991) “Vowel height assimilation in Bantu languages,” *Working Papers of the Cornell Phonetics Laboratory* 5, 37-76, Ithaca, N.Y., Cornell University.

Clements, G.N. and Elizabeth Hume (1995) “Internal organization of speech sounds,” in John A. Goldsmith, editor, *The Handbook of Phonological Theory*, Blackwell.

Cohn, Abigail C. (1990) *Phonetic and Phonological Rules of Nasalization*, Doctoral dissertation, UCLA.

Cohn, Abigail C. (1993) “Nasalization in English: phonology or phonetics,” *Phonology* 10, 43-81.

Cohn, Abigail C. (1998) “The phonetics-phonology interface revisited: Where’s phonetics?,” *Proceedings of the 1998 Texas Linguistics Society Conference* 41, 25-40.

Connell, Bruce (1990) “Aspects of pitch realisation in Yoruba,” *Phonology* 7, 1-29.

Disner, Sandra Ferrari (1983) *Vowel Quality: The Relation Between Universal and Language Specific Factors*, Doctoral Dissertation, UCLA, Los Angeles, California.

Donegan, Patricia Jane and David Stampe (1979) “The study of natural phonology,” in Daniel A. Dinnsen, editor, *Current Approaches to Phonological Theory*, Indiana University Press, Bloomington, Indiana.

Fischer-Jørgensen, E. (1970) “Phonetic analyses of breathy (murmured) vowels in Gujarati,” *Indian Linguistics* 28, 71-140.

Flemming, Edward (2001) “Scalar and categorical phenomena in a unified model of phonetics and phonology,” *Phonology* 18, 7-44.

Fougeron, C. and P. A. Keating (1997) “Articulatory strengthening at edges of prosodic domains,” *Journal of the Acoustical Society of America* 101:6, 3728-3740.

Fresco, Edward Max (1970) “Topics in Yoruba dialect phonology,” *Studies in African Linguistics* Suppl. 1, 1-219.

Fulop, Sean, Ethelbert Kari, and Peter Ladefoged (1998) “An acoustic study of the tongue root contrast in Degema vowels,” *Phonetica* 55, 80-98.

Gaitenby, Jane (1965) “The elastic word,” *Haskins Laboratories Status Report on Speech Research* SR-2, 3.1-3.12.

Gordon, Matthew and Peter Ladefoged (2001) “Phonation types: A cross-linguistic overview,” *Journal of Phonetics* 29, 383-406.

Green, M. M. and G. E. Igwe (1964) *A Descriptive Grammar of Igbo,* Oxford University Press, London.

Hale, Mark (1999) “Historical phonology and phonological theory in the 21st century,” paper presented at the Phonology 2000 Symposium, Cambridge, Massachusetts.

Hale, Mark and Charles Reiss (2000) “Phonology as cognition,” in Noel Burton-Roberts, Philip Carr, and Gerard Docherty, editors, *Phonological Knowledge: Conceptual and Empirical Issues,* Oxford University Press, Oxford, UK.

Halle, Morris (1995) “Feature geometry and feature spreading,” *Linguistic Inquiry* 26, 1-46.

Halle, Morris and Ken Stevens (1969) “On the feature [ADVANCED TONGUE ROOT],” *Quarterly Progress Report* 94, 209-215.

Harris, J. (1987) “Non-structure-preserving rules in lexical phonology,” *Lingua* 72, 255-292.

Haudricourt, A.-G. (1961) “Bipartition et tripartition des systèmes de tons dans quelques langues d’Extrême-Orient,” *Bulletin de la Société de Linquistique de Paris* 56, 163-180.

Hayes, Bruce (1999) “Phonetically-driven phonology: The role of optimality theory and inductive grounding,” in Michael Darnell, Edith Moravscik, Michael Noonan, Frederick Newmeyer, and Kathleen Wheatly, editors, *Functionalism and Formalism in Linguistics, Volume I: General Papers,* John Benjamins, Amsterdam.

Hess, Susan (1992) “Assimilatory effects in a vowel harmony system: An acoustic of advanced tongue root in Akan,” *Journal of Phonetics* 20, 475-492.

Hombert, Jean-Marie (1977) “Consonant types, vowel height, and tone in Yorùbá,” *Studies in African Linguistics* 8:2, 173-190.

Hombert, Jean-Marie (1978) “Consonant types, vowel quality, and tone,” in Victoria Fromkin, editor, *Tone: A Linguistic Survey*, Academic Press, San Diego, California.

Huffman, M. (1987) “Measures of phonation type in Hmong,” *Journal of the Acoustical Society of America* 81, 495-504.

Hyman, Larry (1976) “Phonologization,” in A. Juilland, editor, *Linguistic Studies Offered to Joseph Greenberg*, Vol. 2, 407-418, Anma Libri, Saratoga, California.

Hyman, Larry (1988) “Underspecification and vowel height transfer in Esimbi,” *Phonology* 5, 255-73.

Keating, Patricia (1984) “Aerodynamic modeling at UCLA,” *UCLA Working Papers in Phonetics* 54, 18-28.

Keating, Patricia (1990) “The window model of coarticulation: Articulatory evidence,” in John Kingston and Mary Beckman, editors, *Papers in Laboratory Phonology I: Between the Grammar and Physics of Speech*, 451-470.

Keating, P., B. Lindblom, J. Lubker, and J. Kreiman (1994) “Variability in jaw height for segments in English and Swedish VCVs,” *Journal of Phonetics* 22, 407-422.

Keating, Patricia, Taehong Cho, Cécile Fougeron, and Chai-Shune Hsu (1999) “Domain-initial strengthening in four languages”, *UCLA Working Papers in Phonetics* 97, 139-151.

Kenstowicz, Michael (1994) *Phonology in Generative Grammar*, Blackwell, Cambridge, Massachusetts.

Kirchner, Robert (1997) “Contrastiveness and faithfulness,” *Phonology* 14, 83-111.

Kirchner, Robert (1998) *An Effort-Based Approach to Consonant Lenition*, Doctoral Dissertation, UCLA, Los Angeles, California.

Klatt, Dennis (1975) “Vowel lengthening is syntactically determined in a connected discourse,” *Journal of Phonetics* 3, 129-140.

Klatt, Dennis H. and Laura C. Klatt (1990) “Analysis, synthesis, and pereception of voice quality variations among female and male talkers,” *Journal of the Acoustical Society of America* 87:2, 820-857.

Kühnert, Barbara and Francis Nolan (1999) "The origin of coarticulation," in William J. Hardcastle and Nigel Hewlett, editors, *Coarticulation: Theory, Data and Techniques,* Cambridge University Press.

Labov, William, Mark Karan and Corey Miller (1991) "Near mergers and the suspension of phonemic contrast," *Language Variation and Change* 3, 33-74.

Ladd, D. and Silverman, K. (1984) "Vowel intrinsic pitch in connected speech," *Phonetica* 41, 31-40.

Ladefoged, Peter (1964) *A Phonetic Study of West African Languages: An Auditory-Instrumental Survey*, Cambridge University Press.

Ladefoged, Peter (1983) "The linguistic use of different phonation types," in D. Bless and J. Abbs, editors, *Vocal Fold Physiology: Contemporary Research and Clinical Issues*, 351-360, College Hill Press, San Diego.

Ladefoged, P. and N. Antoñanzas-Barroso (1984) "Computer measures of breathy voice quality," *UCLA Working Papers in Phonetics* 61, 79-86.

Ladefoged, P. and I. Maddieson (1990) "Vowels of the world's languages," *Journal of Phonetics* 18, 93-122.

Ladefoged, P., I. Maddieson, and M. Jackson (1988) "Investigating phonation types in different languages," in O. Fujimura, editor, *Vocal Physiology: Voice Production, Mechanisms and Functions,* Raven Press, New York, 297-317.

Laniran, Yetunde (1993) *Intonation in Tone Languages: The Phonetic Implementation of Tones in Yoruba*, Doctoral Dissertation, Cornell University, Ithaca, NY.

Lare, Kpapo (1990) *Morphologie et Classification des Noms en Ntcham (Bassar),* Master's Dissertation, University of Benin, Lomé, Togo.

Lavoie, Lisa (1999) *Phonological Patterns and Phonetic Manifestations of Consonant Weakening*, Doctoral Dissertation, Cornell University, Ithaca, NY.

Lehiste, I. and Peterson, G. E. (1961) "Some basic considerations in the analysis of intonation," *Journal of the Acoustical Society of America* 33, 419-425.

Lindau, M. (1975) *Features for Vowels*, Doctoral dissertation, UCLA.

Lindau, M. (1978) "Vowel Features," *Language* 54:3. 541-563.

Lindau, Mona (1979) "The feature EXPANDED," *Journal of Phonetics* 7, 163-176.

Lindau, Mona and Patricia Wood (1977a) "Vowel spaces in some West African languages," ms., University of Ibadan, Nigeria.

Lindau, Mona and Patricia Wood (1977b) "Acoustic vowel spaces," *University of California Working Papers in Phonetics* 38, 41-47.

Lindblom (1963) "Spectrographic study of vowel reduction," *Journal of the Acoustical Society of America* 35, 1773-1781.

Maddieson, Ian (1985) “Phonetic cues to syllabification,” in Victoria Fromkin, editor, *Phonetic Linguistics*, 203-222, Academic Press, Orlando.

Majors, Tivoli Jane (1998) *Stress Dependent Harmony: Phonetic Origins and Phonological Analysis*, Doctoral Dissertation, University of Texas at Austin.

Malécot, A. (1956) “Acoustic cues for nasal consonants: An experimental study involving a tape-splicing technique,” *Language* 32, 274-284.

Manuel, Sharon Y. (1990) “The role of contrast in limiting vowel-to-vowel coarticulation in different languages,” *Journal of the Acoustical Society of America* 88, 1286-1298.

Matisoff, J. A. (1973) “Tonogenesis in Southeast Asia,” in L. Hyman, editor, *Consonant Types and Tone*, 71-95, University of Southern California.

McCarthy, John (1988) “Feature geometry and dependency: A review,” *Phonetica* 43, 84-108.

McCarthy, John and Alan Prince (1993) “Generalized alignment,” *Yearbook of Morphology* 1993 79–154.

Myers, Scott (1997) “Expressing phonetic naturalness in phonology,” in I. Roca, editor, *Constraints and Derivations in Phonology*, Oxford University Press, Oxford.

Myers, Scott (2000a) “Boundary disputes: The distinction between phonetic and phonological sound patterns,” in N. Burton-Roberts, P. Carr, and G. Docherty, editors, *Phonological Knowledge: Conceptual and Empirical Issues*, Oxford University Press, Oxford.

Myers, Scott (2000b) “The typology of voicing patterns and the limitations of factorial typology,” handout from talk at Cornell University.

Nesper M. and I. Vogel (1986) *Prosodic Phonology*, Foris, Dordrecht.

Newport, Elissa and Richard Aslin (2004) “Learning at a distance I. Statistical learning of non-adjacent dependencies,” *Cognitive Psychology* 48, 127–162.

Odden, David (1991) “Vowel geometry,” *Phonology* 8, 261-290.

Ohala, John (1990) “The phonetics and phonology of aspects of assimilation”, in John Kingston and Mary E. Beckman, editors, *Papers in Laboratory Phonology I: Between the Grammar and Physics of Speech*, 258-275, Cambridge University Press.

Ohala, John (1993) “Coarticulation and phonology,” *Language and Speech* 36, 155-170.

Ohala, John (1994) “Towards a universal, phonetically-based, theory of vowel harmony,” *Proceedings of the International Conference on Spoken Language Processing (ICSLP)*, 491-494, Yokohama, Japan.

Ohala, John (1995) “Phonetic explanations for sound patterns: Implications for grammars of competence,” Proceedings of the 16th International Congress of Phonetic Sciences, vol. 2, 52-59, Stockholm.

Ohala, John and Manjari Ohala (1993) “The phonetics of nasal phonology: Theorems and Data,” *Phonetics and Phonology* 5, 225-249, Academic Press.

Öhman, S. E. G. (1966) “Coarticulation in VCV utterances: Spectrographic measurements,” *Journal of the Acoustical Society of America* 39, 151-168.

Ọla Orie, Ọlanikẹ (2001) “Two harmony theories and high vowels in comparative Yorùbá,” presented at the Linguistic Society of America meeting, Washington, D.C.

Oller, D. Kimbrough (1973) “The effect of position in utterance on speech segment duration in English,” *Journal of the Acoustical Society of America* 54, 1235-1247.

Omisore, Folasade Omolola (1989) *A Comparative Study of Ife, Ijesa, and Ekiti Dialects*, BA Essay, University of Ilọrin, Nigeria.

Oyelaran, Olasope O. (1973) “Yoruba vowel co-occurrence restrictions,” *Studies in African Linguistics* 4, 155-82.

Parkinson, Frederick (1996) *The Phonological Representation of Vowel Height, its Phonetic Correlates and Role in Constraint-Based Phonology*, Doctoral dissertation, Ohio State University.

Peperkamp, S. and E. Dupoux (2004) “The acquisition of abstract phoneme categories,” talk presented at the Ninth Conference on Laboratory Phonology, University of Illinois at Urbana-Champaign.

Perkell, J. (1971) “Physiology of speech production: A preliminary study of two suggested revisions of the features specifying vowels,” *Quarterly Progress Report of the MIT Research Laboratory of Electronics* 102, 123-139.

Peterson, G. E. and Barney, H. L. (1952) “Control methods used in a study of the vowels,” *Journal of the Acoustical Society of America* 24, 175-184.

Pierrehumbert, Janet B. (1990) “Phonological and phonetic representation,” *Journal of Phonetics* 18, 375-394.

Pierrehumbert, Janet B. (2001a) “Why phonological constraints are so coarse-grained,” in J. McQueen and A. Cutler, editors, *SWAP special issue*, *Language and Cognitive Processes* 16 5/6, 691-698.

Pierrehumbert, Janet B. (2001b) “Exemplar dynamics: Word frequency, lenition, and contrast,” in Joan Bybee and Paul Hopper, editors, *Frequency and the Emergence of Linguistic Structure*, John Benjamins, Amsterdam, 137-157.

Pierrehumbert, Janet B. (2003) “Probabilistic phonology: Discrimination and robustness,” in R. Bod, J. Hay, and S. Jannedy, editors, *Probability Theory in Linguistics*, MIT Press, 177-228.

Pike, Kenneth (1947) "Tongue root position in practical phonetics," *Phonetica* 17, 129-140.

Prince, A. and P. Smolensky (1993) *Optimality Theory: Constraint Interaction in Generative Grammar*, Ms., University of Colorado and Rutgers University.

Pulleyblank, Douglas (1988a) "Vocalic underspecification in Yoruba," *Linguistic Inquiry* 19, 233-270.

Pulleyblank, Douglas (1988b) "Vowel delection in Yoruba," *Journal of African Languages and Linguistics* 10, 117-136.

Pycha, Anne, Pawel Nowak, Eurie Shin, and Ryan Shosted (2003) "Phonological rule-learning and its implications for a theory of vowel harmony," in Gina Garding and Mimu Tsujimura, editors, *Proceedings of the 22nd West Coast Conference on Formal Linguistics* 22, 101-114, Cascadilla Press, Somerville, Massachusetts.

Recasens, Daniel (1984) "V-to-C coarticulation in Catalan VCV sequences: An articulatory and acoustic study," *Journal of Phonetics* 12, 61-73.

Recasens, Daniel (1985) "Coarticulatory patterns and degrees of coarticulatory resistance in Catalan CV sequences," *Language and Speech* 28, 97-114.

Recasens, Daniel (1987) "An acoustic analysis of V-to-C and V-to-V coarticulatory effects in Catalan and Spanish sequences," *Journal of Phonetics* 15, 299-312.

Sagey, Elizabeth (1986) *The Representation of Features and Relations in Nonlinear Phonology*, Doctoral Dissertation, MIT, Cambridge, Massachusetts.

Shryock, Aaron, Peter Ladefoged and Kay Williamson (1996) "Phonetic structures of Defaka," *The Journal of West African Languages* 26, 3-27.

Silverman, Daniel (1997) "Laryngeal complexity in Otomanguean vowels," *Phonology* 14, 235-261.

Silverman, Daniel (2000) "Hypotheses concerning the phonetic and functional origins of tone displacement in Zulu," *Studies in African Linguistics* 29:2, 3-32.

Sproat, Richard and Osamu Fujimura (1993) "Allophonic variation in English /l/ and its implications for phonetic implementation," *Journal of Phonetics* 21, 291-311.

Stahlke, Herbert (1972) "The development of the three-way tonal contrast in Yoruba," in Erhard Voeltz, editors, *3rd Annual Conference on African Linguistics*, 139-145.

Stahlke, Herbert (1976) "The noun prefix in Yoruba," *Studies in African Linguistics,* Supplement 6, 243-253.

Steriade, Donca (2000) "The phonology of perceptibility effects: The P-map and its consequences for constraint organization," Ms., UCLA.

Stevens, S. S., J. Volkmann, and E. B. Newman (1937) “A scale for the measurement of the psychological magnitude pitch,” *Journal of the Acoustical Society of America* 8, 185-190.

Stewart, J. M. (1967) “Tongue root position in Akan vowel harmony,” *Phonetica* 16, 185-204.

Tiede, Mark K. (1996) “An MRI-based study of pharyngeal volume contrasts in Akan and English,” *Journal of Phonetics* 24, 399-421.

Trubetzkoy, N. S. (1939) *Grundzüge der Phonologie*, Vandenhoeck & Ruprecht, Göttingen (Translated by Christiane Baltaxe, 1969, *Principles of Phonology*, Berkeley, University of California Press).

Watkins, Justin W. (1997) “Can phonation types be reliably measured from sound spectra? Some data from Wa and Burmese,” in Stefan Ploch, editor, *SOAS Working Papers in Linguistics and Phonetics* 7, 321-339.

Westbury, J. R., and Keating, P. A. (1985) “On the naturalness of stop consonant voicing,” *Working Papers in Phonetics (UCLA)* 60, 1-19.

Westermann, D., and Ida C. Ward (1933/1964) *Practical Phonetics for students of African Languages*, Oxford University Press.

Whalen, D. H. (1990) “Coarticulation is largely planned,” *Journal of Phonetics* 18, 3-35.

Whalen, D. H. and Andrea G. Levitt (1995) “The universality of intrinsic F_0 of vowels,” *Journal of Phonetics* 23, 349-366.

Wightman, Colin, Stefanie Shattuck-Hufnagel, Mari Ostendorf, and Patti Price (1992) “Segmental duration in the vicinity of prosodic phrase boundaries,” *Journal of the Acoustical Society of America* 91, 1707-1717.

Williamson, Kay (1982) “The classification of East Niger-Congo,” *JOLAN* 1, 101-105.

Williamson, Kay (1989) “Benue-Congo overview,” in J. Bendor-Samuel, editor, *The Niger-Congo Languages*, University Press, London.

Zsiga, Elizabeth (1992) “A mismatch between morphological and prosodic domains: Evidence from two rules,” *Phonology* 9, 101-135.

Zsiga, Elizabeth (1993) *Features, Gestures, and the Temporal Aspects of Phonological Organization*, Doctoral Dissertation, Yale University.

Zsiga, Elizabeth (1995) “An acoustic and electropalatographic study of lexical and postlexical palatalization in American English,” in Bruce Connell and Amalia Arvaniti, editors, *Phonology and Phonetic Evidence: Papers in Laboratory Phonology* IV, Cambridge University Press, Cambridge.

Zsiga, Elizabeth (1997) “Features, gestures, and Igbo vowels: An approach to the phonology-phonetics interface,” *Language* 73, 227-274.