Chapter 2

From co-occurrence patterns to rhythmic alignment: Ongoing investigations of Twi consonants and vowels

Kenneth de Jong

Indiana University

This chapter describes the results of several phonological and phonetic research projects investigating the system of secondary articulations found in Twi, in particular the extensive combinations of secondary articulations found with different fricatives. These patterns give evidence of a historical source for the secondary articulations as lying in the interaction between vowel articulations superimposed on various obstruent articulations, and how these affect acoustic output patterns. The chapter goes on to describe ongoing research into the relative timing of consonant and vowel articulations in speakers of Twi, by means of rhythmic metronome repetition tasks.

1 Introduction

As a linguistics professor at the University College London and then at the School of Oriental and African Studies (SOAS) from the 1920's to the 1950's, J. R. Firth's take on phonological analysis was and remains somewhat unusual in the context of modern linguistics. The crux of Firth's approach to phonological systems, vis-à-vis the more common approaches found especially in the American academy, is an insistence that phonological systems are multisystemic, and on the phonological side, that one cannot assume that particular phono-segmental structures exist before one begins to work with a language. The approach asserts that once

one begins looking at larger spans of time, one finds that language users express many different facets of meaning beyond just the lexical contrasts that comprise the morpho-syntactic structures linguists typically examine.

As noted above, the Firthian approach to phonological systems questions some of the basic assumptions of many traditional phonological models. The most obvious divergence from what we typically expect of such systems is a skepticism about the segment. In many ways, the Firthian approach prefigured the move toward multi-tiered representations which began to be popularized in Goldsmith (1976). While structuralist analyses of the Bloomfieldian tradition focused on isolating categories of contrasts from which languages build a lexicon, and tended to represent the signal as a sequence of such categories, these alternative representations opened up the possibility of segments arising from the bundling of various properties which may have different temporal spans. Such research has highlighted languages with harmonic systems, like Turkic and Uralic vowel harmony systems, and the Tupi nasal harmony systems found in South America, with the point being that one does not get segments for free as some necessary building block of language.

Perhaps the most read phonological work from the 1960s, and certainly the most read work on English phonology, *The sound pattern of English* (Chomsky & Halle 1968), was particularly notable for its persistent attempts to remove any reference to non-segmentally organized information from the lexicon. In this work, prosodic stress was argued to be predictable, and hence, unnecessary. Tone was only referenced in passing, and syllabic organization was also done away with. Ideas laid out in Davis (1988), however, document many reasons why we need to understand these prosodic phenomena, especially stress and syllabification, in order to make sense of phonological systems, including the English phonological system.

In this Firthian tradition, de Jong & Obeng (2000) illustrate via Twi, a variety of Akan, aspects of the language's segmental phonology that bear upon our understanding of the fundamental units of speech and the cross-linguistic relationship between segments and prosodic structure. The current paper revisits this work and expands upon its findings through novel research on the relative timing of consonant and vowel articulation as explored through rhythmic metronome repetition tasks.

1.1 Twi fricatives and affricates

Twi boasts 11 different fricative qualities in our first survey of the system: [f, s, ɕ, h, fʲ, sʲ, fʷ, sʷ, hʷ, sᵋ, ɕᵋ]. A brief summary of examples from the Asante variety is laid out in Table 1, cross-classifying the examples by primary articulation in

the rows, and secondary articulations shaping those primary articulations in the columns. These examples are from de Jong & Obeng (2000: 686–687).

Table 1: Examples of contrasting fricatives in Asante Twi

Plain	Palatalized	Labialized	Labio-palatalized
fa 'to take'	*fʲa* 'to embellish'	*fʷa* 'to endorse'	–
sa 'to dance'	*æsʲa* 'tree sp.'	*sʷa* 'to carry on one's head'	*sᶣa* 'to carry one one's head'
çi 'to burn'	–	*çʷa* 'scrotum'	*çᶣa* 'scrotum'
ha 'here'	–	*hʷa* 'whack!'	–

In analyzing the system further, there are complicating irregularities concerning the idea that the Twi sound system includes eleven different fricative segments, and these irregularities come in three categories.

First, many of these fricatives are associated with secondary articulations, as explicated in the layout of Table 1. While there are four basic articulations, one labial, two coronal, and one glottal (or possibly dorsal), the addition of palatalization, lip rounding, and even a combination of palatalization with lip rounding – labio-palatalization – are what expands the inventory of qualities from four to the impressive eleven.

Second, these secondary articulations are distributed in a way that is not independent of the primary articulation. Co-occurrence restrictions are fairly common in phonological systems, but the particular patterning in Twi is not typical. For more on the complexities of secondary articulations and their phonotactic basis, see Ofori's chapter 2 in this volume. As Ofori discusses, there are some constraints in Akan on segmental combinations that one can broadly ascribe to OCP constraints (see Leben 1973, and an extensive thread of following research) which insist that two collocated items differ from one another. One example of this in Akan is the following: a palatalized articulation cannot combine with any dorsal articulation. However, most other constraints do not fit this pattern, for example those involving lip rounding. Typologically, the most typical restriction on rounding combining with consonantal articulation is a proscription of labialization from combining with labials. Thus, [fʷ] is typologically rare. However, it is perfectly fine in Twi systems. Similarly, labio-palatalization only occurs with primary coronal articulations [çᶣ sᶣ], those most similar to the palatal element in the secondary articulation. This labio-palatalization, however, is not just a secondary property of the palatal articulation on the alveolo-palatal fricatives, because it also appears with the alveolar sibilant.

Third, these fricative articulations also exhibit a complex pattern of co-occurrence with the following vowel; again, see an impressive array of examples speaking to this fact in Ofori's chapter in this volume. This pattern of co-occurrence suggests strongly that, at least historically, the large array of fricatives in Twi have arisen partly due to consonants interacting with the vowel system. That is, the current consonant system reflects patterns of organization at roughly the syllabic level, which is where consonants and following vowels are organized into a production complex. De Jong & Obeng's (2000) investigation into these irregularities in segmental make-up and contextual constraints led them back into the question of syllabic organization as a root cause for their occurrence. Specifically, the integrated system of consonants in Twi appears to be the outcome of contrasts in the vowel system migrating into the consonant system. In this process of migrating, the system takes on a very different shape, since the acoustic space in which obstruents are articulated is very different from the acoustic space that the vowel contrasts came from. I next turn to the evidence for this sort of migration.

1.2 Phonotactic restriction and evidence for vowel migration

Akan vowel systems are all quite similar, with each language typically exhibiting ten vowels, composed of a five-vowel system, each vowel having a [± ATR] pair. With respect to the consonant system, the ATR contrasts are essentially irrelevant, so the patterns can be more simply grasped by treating the vowel system as having five vowels. In this five-vowel system, the high and mid vowels pattern similarly, so there are three categories of vowels to be considered: front vowels ([i/ɪ] & [e/ɛ]), back, rounded vowels ([u/ʊ] & [o/ɔ]), and low vowels ([a/ɑ]). The patterns of co-occurrence found with the fricatives also extend, with some minor variations, to the plosives in the language, so the effects described here affect the entire obstruent system.

Looking at obstruents occurring before front vowels, one notes a systematic gap, that posterior consonants do not appear before front vowels. Similarly, posterior consonants ([k g h]) exhibit only a two-way contrast in secondary articulations; they systematically exhibit minimal contrasts in rounding, but never contrasts in palatalization. This gap strongly suggests that any historical contrastive front vowel elements that appeared after the posterior consonants have been incorporated into the consonant system in the form of primary distinctions between alveolo-palatal consonants and velar/glottal consonants. Or, more directly:

[k] + [i/ɪ] > [tɕ]
[g] + [i/ɪ] > [dʑ]
[h] + [i/ɪ] > [ɕ]

Obeng (p.c.) has recently mentioned that there are older speakers in his memory with productions reflecting the earlier form of the language before this posited change, so the time-depth of this process would be in the early 20th century. The resulting co-occurrence pattern is not completely consistent, as there are a small number of current forms such as [kita] 'to hold,' but by-and-large the velars and glottals do not appear before front vowels or with palatal secondary articulations.[1]

This hypothetical historical pathway to the alveolo-palatal series, then, further suggests a historical source for the labio-palatalization articulations as coming from rounded dorsal consonants appearing before front vowels, or more directly:

[kʷa] + [i/ɪ] > [tɕᵩ]
[gʷ] + [i/ɪ] > [dʑᵩ]
[hʷ] +[i/ɪ] > [ɕᵩ]

That is, the labio-palatal articulation is actually, at least historically, a combination of previous consonant rounding and palatal articulation from the following vowel, both becoming incorporated into a complex secondary articulation of the consonants.

This overall conjecture about the dorsal series and front vowels is further borne out by the complicated system appearing before low vowels. The low vowel context exhibits the greatest number of contrasting secondary articulations on the obstruent. For labial and alveolar consonants before [a/ɑ], there is a systematic, three-way contrast in plain, labial, and palatal. Minimal pairs exhibiting these secondary articulation contrasts are readily available; with [b], [d], and [t], there's a systematic four-way lexical contrast, as exemplified in (1) with [t].

(1) a. *ta* 'to plaster'
     b. *tʲa* 'to step on'
     c. *tʷa* 'bottle'
     d. *tᵩa* 'to pay'

---

[1]It is possible that these exceptions are the result of subsequent changes or borrowings, which obscure the overall pattern.

However, with posterior consonants, alveolo-palatal, and velar/glottal, there is largely only a two-way contrast. Alveolo-palatal consonants contrast plain vs. labio-palatal, while velars and glottals contrast plain with rounded, as in (2). Palatalization does not co-occur with posterior consonants before [ɑ], as in (3).

(2) a. *dʑa* 'to leave behind'
     b. *dʑᵁa* 'to butcher'
     c. * *dʑʲa*
     d. * *dʑʷa*

(3) a. *ha* 'here'
     b. * *hᵁa*
     c. * *hʲa*
     d. *hʷa* 'whack!'

Putting all these pieces together, it appears that the systematic, two-way rounding contrast in consonants before [ɑ] can be split into a four-way contrast by incorporating an intervening [i] vowel: [tɑ] contrasting with ([t] + [i/ɪ] + [ɑ]>) [tʲɑ], and [tʷɑ] contrasting with ([tʷ] + [i/ɪ] + [ɑ]>) [tᵁɑ]. Posterior consonants also contrast in rounding, alveolo-palatals contrasting plain with labio-palatals, and velars/glottals contrasting plain with rounded. However, since the palatal element of the front vowel has already been incorporated into the primary articulation system, it is not available anymore to divide the secondary articulation contrasts. So, [tɑ] ⟺ [tʲɑ], and [tʷɑ] ⟺ [tᵁɑ], but [gɑ] ⟺ [dʑɑ], and [gʷɑ] ⟺ [dʑᵁɑ].

The final piece to this story concerns the rounding contrast itself. This rounding systematically does not contrast before back vowels, except in the case of labio-palatal contrasts in the alveolo-palatal consonants in some varieties, such as Akwapem: [dʑo] ' to cool' vs. [dʑᵁow] 'to harvest palm nuts.' A sense of linguistic neatness would suggest a similar source for the rounding secondary articulation and the palatal secondary articulation. If the palatalization arises as incorporation of a front vowel articulation into the consonant system, then perhaps consonant rounding could also be the result of the incorporation of a back vowel into the consonant system. The suggested historical patterns with example anterior and posterior obstruents, then, are summarized in Table 2.

To sum up, all this patterning might suggest that the whole secondary articulation system is actually better thought of as a vowel system with the peculiarity of being articulated in conjunction with the consonants. Or, to put it another way,

Table 2: Posited sources for contrasting fricatives in Asante Twi

Posited historic source	Current form
s + ɑ	sɑ
h + ɑ	hɑ
s + i + ɑ	sʲɑ
k + i + ɑ	çɑ
t + u + ɑ	sʷɑ
k + u + ɑ	hʷɑ
t + u + i + ɑ	sᶣɑ
k + u + i + ɑ	çᶣɑ

maybe the co-production pattern of vowels and consonants is systematically different from what one expects from other languages, and so, the main difference between Kwa languages like the Akan varieties and, for example, Indo-European languages, is a different kind of syllabic organization of the contrasting elements. Here, we call the organization syllabic, because it is the syllabic level that coordinates the production of consonant and vowel articulations.

1.3 Phonetic factors in syllabic organization in Twi

In this line of inquiry, a first question is why the non-low vowels might appear in such close temporal proximity with the onset consonants in the first place. What is it about these languages that has encouraged the development of what look like complex consonants out of consonant + vowel sequences?

A second, related question is whether these consonantal complexes are actually just consonants and vowels, and not complex consonants at all. That is, could the current production patterns actually better be thought of as an accidental acoustic byproduct of syllabic organization, without reference at all to the segmental structure of the consonant system? Leading in this direction is the presence of a lot of apparent variation in the production of secondary articulations. Rounding sometimes varies with labio-palatalization, as in [pʷie] ~[pᶣie] 'to exit,' and palatalization varies with an actual vowel, as in [sie] ~[sʲe] 'to bury.'

This second line of questioning, however, does not seem entirely explanatory of the situation, since there are examples of lexical items which violate the co-occurrence patterns expected, if the whole secondary articulation pattern is just the systematic fall-out of a general syllabic organization creating the patterns.

Some forms lexically require the non-palatal rounding with the velars, in contrast to the palatal rounding. For example, [kʷa] 'to paint with clay' contrasts with [ækᶣaŋ] 'to cheat.' Similarly, some forms require the palatal secondary articulation without the palatal primary articulation, and one finds forms such as [hᶣaŋ] 'to sponge off someone.' If the different secondary articulations are just unspecified variants of following vowels produced by a general syllabic articulation frame, then we would not expect lexically specific items that require a particular primary or secondary articulation.

De Jong & Obeng (2000) indicated a large amount of temporal overlap in the production of vowels and consonants, suggesting something like a syllabic frame which might account for the patterning. Palatography of consonants with different following vowels in various languages generally indicates that a following vowel will shift consonant closures in the direction of the vowel articulation. However, the extent of this effect in individual speakers is notable and, thus, it was deemed worthy of further investigation.

To quantify the degree of this temporal overlap, this earlier study explored a statistical framework which attempted to determine how far from the edges of a consonant one might be able to detect the identity of a neighboring vowel in the acoustic patterns on the other side of the consonant. The corpus included consonants intervening between two vowels, and then formant measurements were taken from the center of both vowels, and on the preceding and trailing edges of the consonant. Directional asymmetries were quite obvious, with the effect of a following vowel showing up in the middle of the initial vowel in one of the speakers, and in the preceding edge of the consonant in both speakers. Effects of the initial vowel on the following vowel, however, were not detectable at all. All this work suggested strongly that the Twi speakers were timing vowel articulations such that they strongly overlapped with a previous consonant.

Further detailed analyses, however, showed that the phonetic behavior of the speakers was not well accounted for just by stating that vowel and consonant articulations overlap. The main difficulty documented in de Jong & Obeng (2000) is in the lingual articulation of the labio-palatal consonants. If labio-palatal consonants are actually just the result of simultaneous production of three different articulations (primary articulation, front vowel/palatal articulation, and rounding articulation), then one would expect the three articulations to behave the same as in other words which are missing one of the articulations. Thus, alveolo-palatal articulations should be the same with or without the secondary articulations added to them. However, this turns out not to be the case; static palatography of the labio-palatal consonants shows that the lingual obstruent closures are shifted in a posterior direction from their plain counterparts.

The effect, then, indicates that, even though the secondary articulations of consonants in Twi may have historically arisen from vowel contrasts, the vowel contrasts have not remained separate from the consonants they are articulated with. The fact that the point of articulation of the alveolo-palatals is shifted in a posterior direction when articulated with labio-palatalization strongly suggests that the secondary articulations and the primary articulations have been fused together.

It is also not inexplicable that the primary closure would be shifted specifically in a posterior direction. This posterior shifting and lip rounding couple together to help lower the frequency of the dominant formant imposed on the noise corresponding to the obstruent. Both lip protrusion and lingual backing have complementary effects on the length of an anterior resonator between the two constrictions; both make that tube longer, and hence the resonating noise in the tube takes on a lower frequency, and results in a prominent formant (corresponding to F3 in the vowel) in the frication noise. The analysis of the noise spectra of these consonants (reflecting observations about fricative contrasts recently reviewed in Shadle 2023) indicates the presence of just such a very low anterior-cavity formant frequency in these labio-palatalized alveolo-palatal fricatives and affricates.

The conclusion of this earlier phonetic work, then, is that, yes, the consonant system of Twi has been heavily augmented by incorporating former vowel articulations into the former consonant system. This process of migration of vowel contrasts into the consonant system appears to be at least partly driven by how these vowel articulations generate distinct noise categories in the consonant system, so the function of the vowel articulations does not work exactly like it did in the vowel system. The consonant system has a different acoustic dimensionality than does the vowel system, and so some of the complexity of the migration process is driven by this mapping from the old, "sending" vocalic system to the "receiving" consonant system. As suggested by one of the reviewers of this chapter, such migrations from vowel to consonant systems are not uncommon in linguistic systems. For example, Faytak (2022) has explored the diachronic context of "fricative vowels" including in Sūzhōu Chinese, as well as some Grassfields Bantu languages, where vowels display more consonant-like articulatory/acoustic targets. Similarly, Voorhoeve (1976) also briefly describes a situation in Medʉmba, another Grassfields Bantu language, in which vowels /o/, /u/ and /e/, /i/ (the latter having been classified as a "super-high vowel" in Bantu reconstructions) are perceptually quite similar except for the fricativizing effect of /u/ and /i/ on a preceding consonant. Recent work on Lutuv (Bonhert et al. 2022), a language of the Chin State in Myanmar, has found a vowel system including six high vowels,

with the high central rounded vowel [ʉ] being "fricativized." What the observations from Twi bring to the fore is the question of how such consonant-like vowel articulations are acoustically related to other vowels and to the idea that vowels and consonants are separated contrastive systems.

The other outstanding puzzle that this earlier work on Twi points out is the overall question of why the temporal coordination of consonants and vowels in Twi is so different from previously heavily examined languages in the first place. Or, to state the question more generally, how do consonant and vowel articulations get coordinated in languages, and how do languages differ in managing such coordination?

1.4 Studying temporal coordination in Twi

Languages such as Twi re-open some very long-standing questions in speech production, concerning how it is that the detailed articulations of speech become coordinated with one another in highly skilled, fluent speech. This topic has been the center of many different threads of research, but the thread that forms the core of the rest of this chapter is that pursued by speech and linguistic researchers at Haskins laboratories, especially during the 1980s and 1990s.

In the late 1980s, several elements of a synthetic speech model came together in an effort to understand how speech unfolds in time. One particularly detailed element of this larger model was Task Dynamics (Salzman & Munhall 1989), which sought to elucidate how a linguistic representation of speech composed of contrasting gestures might be interpreted as action regimes associated with these different contrastive elements – action regimes that then become orchestrated together in larger speech units corresponding variously to segments, syllables, or sub-syllabic units such as onsets and codas, all as part of a single orchestration scheme. Task Dynamic mechanisms interpret this orchestration as a mapping onto control-regimes for actual anatomical articulators, which then play out as schematized movement patterns. These are then interpreted by an articulatory synthesizer into actual acoustic signals.

Particularly relevant to the current discussion, the prosodic organization of these various gestures was pursued in heavily cited papers by Browman & Goldstein (1986, 1990). Their approach, along with similar work by others (e.g., de Jong 2003) envisioned syllabic organization as consisting of specified timing relations between the gestures associated with consonants and those associated with vowels. Subsequent work by Goldstein (e.g., with particularly interesting cross-linguistic analyses in Goldstein et al. 2007) drew on earlier coordination

models by Tuller, Kelso, and colleagues (e.g., Tuller & Kelso 1991) to hypothesize that syllabic structures correspond to stable phasing relations between the consonant and vowel systems.

These models, based as they are in the motor control literature, indicate that there should be a stereotypical set of timing relations found in all human languages, as long as the individual segments are executed by gestures of similar type to the ones seen in commonly examined western languages. However, the work discussed here on Twi suggests the language exhibits a tendency towards a different timing relation between consonant articulation and high vowel articulation, one in which a following vowel is synchronized with the release of consonant, such that the peak articulation of the vowel occurs simultaneous to the consonant release. The patterns found here are strikingly different from what is found in similar research with paleography on French, where there is very little overall difference in coronal contact for consonants before front and back vowels (Dart 1991).

While there is much direct measurement of production patterns that this set of speculations would suggest, I would like to close this section proposing a somewhat different approach to analyzing timing relations motivated by these very interesting questions that Twi raises for speech science. This research thread concerns the field of rhythmic organization.

1.5 Segmental make-up and P-centers

Recent work has revived interest in an experimental paradigm which became popular in the 1970s, that of the P-center (Allen 1972, Morton et al. 1976, Rapp 1971). In this work, two different threads of research, one in speech perception and one in speech production, came together in observations by Morton et al. (1976). They were examining memory patterns in speech perception, and encountered the difficult problem of figuring out how people perceive speech to have a regular timing pattern. It turned out that it was quite difficult to identify anything in the acoustic signal that precisely corresponded to the perception of regular speech timing, since the balance of acoustic information about consonants and vowels contributed to the timing pattern. The general effect was that listeners, given the task of adjusting recordings of speech syllables to sound regular, would align the beginning of vowels to occur at regular intervals. Onset consonants, preceding the vowel, would tend to pull the alignment point to an earlier time. Different consonants also would affect alignment to different degrees.

This complex pattern matches up quite well with what researchers such as Rapp and Allen found with production tasks, such as having people produce

syllables in time with a metronome or having them tap a finger in time with speech. From this literature, the concept of "P-center" was formulated, as a way of avoiding deciding whether the center of the syllable is based on perception or production, since both tasks give evidence for the effect.

Recent work on the P-center has suggested that the location of the P-center is language specific, and sensitive to the syllabic inventory of the language. Specifically, Chow et al. (2015) examined production in time to a metronome by Cantonese speakers, finding that the speakers were timing the beginning of the consonant with the metronome beats, rather than the vowel. The authors speculated that this alternative alignment pattern is due to the relatively simple onset structure in Cantonese, wherein (in many analyses) there are no consonant clusters. More recent work with speakers of Mandarin (Lin & de Jong 2023), however, has found a pattern similar to that found in languages such as English, with general alignment of a point near the beginning of the vowel with the metronome beats.

Of particular interest to the current chapter are two points: First, that the P-center might be sensitive to the structure of the language of the speakers, rather than just to the auditory or production substance of the speech. Besides the unusual results of Chow et al. (2015), a similar sensitivity of the P-center alignment to linguistic structure, specifically morphological constituency, has been found in Medʉmba (Franich 2018). Second, that the crucial point of alignment resides in the nexus between the consonant onset and the vowel. These two points together raise the more specific question of what counts as a consonant, and what counts as a vowel, especially in a case such as Twi where there is evidence that vowels can, at least historically, migrate into the consonant position. Are these vowels which are heavily co-produced with the onset consonant, or are they actually part of the consonants themselves?

1.6 P-centers in Twi: A first look

To begin exploring rhythmic organization in Twi, I employed the protocol used in a previous study of Mandarin (Lin & de Jong 2023). The logic of this approach is that, if secondary articulations in Twi are actually vowels which are produced in time with preceding consonants, then we would expect syllables with secondary articulations, such as [tʲɑ] 'to step on,' to exhibit an earlier timing of the metronome with the respect to the initial consonant. The metronome would line up with the secondary articulation, rather than with the following vowel. This logic might not extend to labialization, however, as our reconstruction places rounding as possibly a part of the earlier consonantal system. So we might not

expect a difference in timing between syllables with labialization, such as [tʷɑ] 'bottle,' and ones without it, such as [tɑ] 'to plaster.'

A preliminary look at the recordings for this project suggests a difference in syllable timing, however, not in the expected direction.[2] Figure 1 presents spectrograms of parts of three production trials, comparing [tɑ] (top), [tʲɑ] (middle), and [tʷɑ] (bottom). The general pattern one finds is that the metronome signal, which shows on the spectrogram as dark squares in the lower frequency ranges, appears near the beginning of the vocalic portion of each syllable production, with quite a bit of variation in the exact timing. Looking in more detail at the middle syllable in the images in Figure 1, the metronome pulse in the middle train ([tʲɑ]) appears after the consonant release, while in the other trials ([tɑ] and [tʷɑ]), the metronome tends to overlay the release of the consonant. This is actually the reverse of what was expected, since the palatal articulation should reflect the onset of the vowel, while the labial articulation might be part of the consonant that would precede the metronome. However, there is also extensive variation in the productions. For example, the last syllable in the top trial has the consonant release happening well before the metronome, while the first syllable has the metronome centered on the release. So, the next step in this journey is to determine the distribution of the different types and to generalize the protocol for use with different speakers.

The variation noted above also led in another direction, that of querying the fact that this sort of repetitive production is strongly rhythmic in nature. Given what we know about rhythmic productions, based on a well-developed research program exemplified in works like Cummins & Port (1998), Cummins (2009), Taijma (1998), and Anderson (2018), we are amply aware of the existence of different modes of phasing in rhythmic production. Given a repetitive cycle, either explicit in a metronome, or just internally generated by a speaker (or musician), the speech content can entrain to different parts of the cycle in any harmonic relation. So, our preliminary discussions have led us into the question of whether the variation in P-center studies might be due to different harmonic entrainment.

Previous work on the perceptual side, using tasks where participants are to adjust recordings to make them sound regular have not found strong evidence for different rhythmic modes (Whalen et al. 1991). Some probing of the possibilities did produce various production modes, as illustrated in Figure 2.

The three types illustrated in Figure 2 seemed perfectly stable ways of repeating syllables to a metronome, as expected from previous work on the subject.

---

[2]These recordings were made with the help of my previous co-author and collaborator Samuel Obeng, whose voice is featured in them.

(a) Productions of /tá/ 'to plaster' [tɑ]

(b) Productions of /tìà/ 'to step on' [tʲɑ]
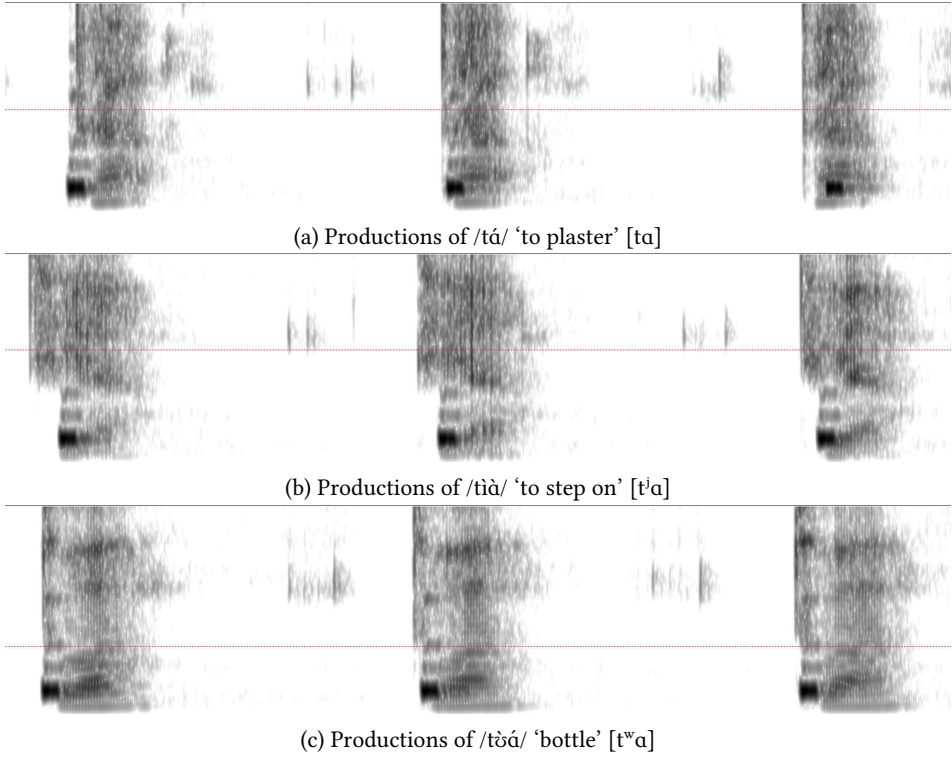

(c) Productions of /tòá/ 'bottle' [tʷɑ]

Figure 1: Three-syllable extracts from metronome speech trains. The metronome is visible as the square dark spot which appears overlaid near the beginning of each syllable.

The most obvious mode is the "on-beat" mode in the top panel, and this was clearly the mode that the productions in the P-center protocol generated. The other modes subdivided the metronome cycle, making harmonic repetitions at higher rates. So, for example, the bottom, "off-beat" mode has the metronome and the syllable alternating regularly so an event happens at twice the rate of the metronome cycle. The "pick-up" mode in the middle is a bit more complicated as the syllable timing suggests a sub-division of this double-time mode, yielding a repetition pattern at four times the metronome frequency. Here, there are two other events (at halfway and three-quarters of the way between metronome pulses) that need to be supplied by the speaker (or musician).

It is unclear at this juncture the extent to which speakers' tendencies toward different modes may affect their patterning in P-center production experiments. It seems clear that the dominant alignment patterns with the metronome and syllabic beats being synchronized is the most obvious way of performing the
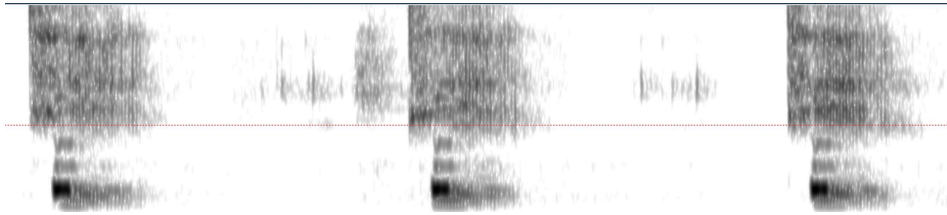
(a) Metronome embedded in the speech image, from the original task which created the productions in Figure 1
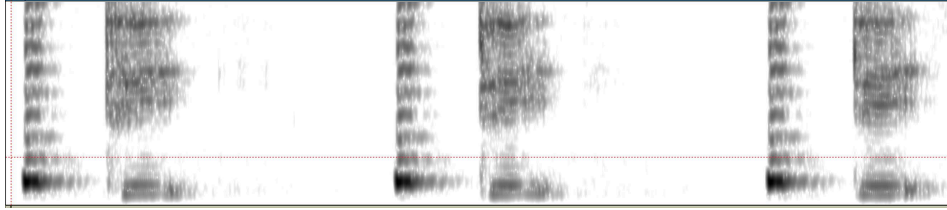
(b) "Pick-up" mode with the metronome beating before the onset of each syllable.
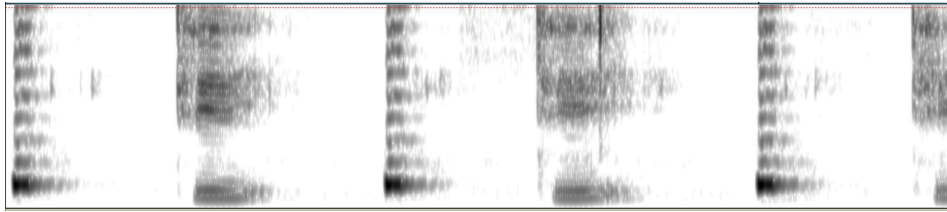
(c) "Off beat" mode with the metronome appearing half-way between each syllable onset.

Figure 2: Three-syllable extracts of three trials probing different rhythmic modes. The metronome appears as a dark square with harmonically related stripes above it. All syllables are productions of /tí/ 'head' [ti].

task, and that most speakers fall into it naturally. Also, the different modes are different enough that we should be able to distinguish them from the on-beat pattern. However, we have encountered speakers who, for some reason, avoid the on-beat pattern early in the trials, but then will gradually migrate into an on-beat pattern later on, and it is not clear whether certain types of syllables are more likely to cause this migration than others. This remains a topic for current investigation.

Of course, these observations are just preliminary, but they suggest the possible fruitfulness of extending this analysis to see how general these patterns are with other speakers of Twi, and other consonants. But a general conclusion to this work is clear: The Akan languages are a rich foundation for moving the discipline of linguistics forward, beyond the comfortable patterns found in more typically studied languages.

Kenneth de Jong

Acknowledgments

References

Allen, George D. 1972. The location of rhythmic stress beats in English: An experimental study. *Language and Speech* 15. 72–100. DOI: 10.1177/002383097201500110.

Anderson, Jonathan C. 2018. *The phonetics and phonology of speech rhythm in Akan.* Indiana University, Bloomington. (Doctoral dissertation).

Bonhert, Amanda, Kelly Berkson & Sui-Hnem Par. 2022. Vowel sounds in Hnaring Lutuv. *Indiana University Working Papers in South Asian Languages and Cultures* 3. DOI: 10.14434/iwpsalc2022.v3i1.35447.

Browman, Catherine P. & Louis Goldstein. 1986. Towards an articulatory phonology. *Phonology* 3. 219–252. DOI: 10.1017/S0952675700000658.

Browman, Catherine P. & Louis Goldstein. 1990. Gestural specification using dynamically-defined articulatory structures. *Journal of Phonetics* 18. 299–320. DOI: 10.1016/S0095447019303766.

Chomsky, Noam & Morris Halle. 1968. *The sound pattern of English.* Cambridge, MA: MIT Press.

Chow, Ivan, Michel Belyk, Vance Tran & Steven Brown. 2015. Syllable synchronization and the P-center in Cantonese. *Journal of Phonetics* 49. 55–66. DOI: 10.1016/j.wocn.2014.10.006.

Cummins, Fred. 2009. Rhythm as entrainment: The case of synchronous speech. *Journal of Phonetics* 37. 16–28. DOI: 10.1016/j.wocn.2008.08.003.

Cummins, Fred & Robert Port. 1998. Rhythmic constraints on stress timing in English. *Journal of Phonetics* 26. 145–171. DOI: 10.1006/jpho.1998.0070.

Dart, Sarah N. 1991. Articulatory and acoustic properties of apical and laminal articulations. *UCLA Working Papers in Phonetics* 79.

Davis, Stuart. 1988. *Topics in syllable geometry*. New York: Garland Press.

de Jong, Kenneth. 2003. Temporal constraints and characterizing syllable structuring. *Papers in Laboratory Phonology* 6. 253–268. DOI: 10.1017/CBO9780511486425.015.

de Jong, Kenneth & Samuel G. Obeng. 2000. Labio-palatalization in Twi: Prosodic, contrastive, and quantal effects in linguistic micro-evolution. *Language* 76. 682–703. DOI: 10.2307/417139.

Faytak, Matthew. 2022. Place uniformity and drift in the Suzhounese fricative and apical vowels. *Linguistics Vanguard* 8. 569–581. DOI: 10.1515/lingvan-2021-0071.

Franich, Kathryn. 2018. Tonal and morphophonological effects on the location of perceptual centers (P-centers): Evidence from a Bantu language. *Journal of Phonetics* 67. 21–33. DOI: 10.1016/j.wocn.2017.11.001.

Goldsmith, John. 1976. *Autosegmental phonology*. MIT. (Doctoral dissertation).

Goldstein, Louis, Ioana Chitoran & Elisabeth Selkirk. 2007. Syllable structure as coupled oscillator modes: Evidence from Georgian and Tashlhiyt Berber. In Jürgen Trouvain & William Barry (eds.), *Proceedings of the XVI International Congress of Phonetic Sciences*, 241–244. Saarbrücken, Germany: Universität des Saarlandes.

Leben, William. 1973. *Suprasegmental phonology*. MIT. (Doctoral dissertation).

Lin, Yu-jung & Kenneth de Jong. 2023. The perceptual center in Mandarin Chinese syllables. *Journal of Phonetics* 99. 101–245. DOI: 10.1016/j.wocn.2023.101245.

Morton, John, Steve Marcus & Clive Frankish. 1976. Perceptual centers (P-centers). *Psychological Review* 1. 405. DOI: 10.1037/0033-295X.83.5.405.

Rapp, K. 1971. A study of syllable-timing. *University of Stockholm, Speech Tranmission Laboratory - Quarterly Progress and Status Report* 12. 14–19.

Salzman, Elliot L. & Kevin G. Munhall. 1989. A dynamical approach to gestural patterning in speech production. *Ecological Psychology* 1. 333–382. DOI: 10.1207/s15326969eco0104_2.

Shadle, Christine. 2023. Alternatives to moments for characterizing fricatives: Reconsidering Forrest et al. *Journal of the Acoustical Society of America* 153. 1412–1426. DOI: 10.1121/10.0017231.

Taijma, Keiichi. 1998. *Phonetics and phonology of speech rhythm: A cross-linguistic exploration in rhythmic speech production*. Indiana University. (Doctoral dissertation).

Tuller, Betty & J. A. Scott Kelso. 1991. The production and perception of syllable structure. *Journal of Speech and Hearing Research* 34. 501–508. DOI: 10.1044/jshr.3403.501.

Voorhoeve, Jan. 1976. *Contes bamileke*. Tervuren: Musée Royal de l'Afrique Centrale.

Whalen, D. H., Andre M. Cooper & Carol A. Fowler. 1991. P-center judgments are generally insensitive to the instructions given. *Haskins Laboratories Status Report on Speech Research* 108. 141–146. DOI: 10.1159/000261843.