The Acoustic Correlates of $[\pm\text{ATR}]$ Vowels: An Analysis by Reference Levels of Anyi Vowels

Ettien Koffi

*St. Cloud State University, enkoffi@stcloudstate.edu*

Follow this and additional works at: [http://repository.stcloudstate.edu/stcloud_ling](http://repository.stcloudstate.edu/stcloud_ling)

Part of the [Applied Linguistics Commons](http://repository.stcloudstate.edu/stcloud_ling)

**Recommended Citation**


Available at: [http://repository.stcloudstate.edu/stcloud_ling/vol5/iss1/9](http://repository.stcloudstate.edu/stcloud_ling/vol5/iss1/9)
THE ACOUSTIC CORRELATES OF [±ATR] VOWELS: AN ANALYSIS BY REFERENCE LEVELS OF ANYI VOWELS

ETTIEN KOFFI

Abstract
This paper pursues two goals simultaneously. First, it seeks to describe the [±ATR] vowels of Anyi Morofu acoustically by investigating six correlates: F0, F1, F2, F3, intensity, and duration. Anyi has a nine-vowel system that is divided evenly among four [+ATR] and four [-ATR] vowels. The vowel /a/ is ambidextrous in that it straddles both groups. The second goal is to demonstrate that reference levels can be used reliably to determine which one(s) of the six acoustic correlates is/are the most acoustically robust. The data that serves as the basis for this analysis comes from 10 participants who produced nine carrier sentences in which a monosyllabic verb of /hV/ structure occurred. Each sentence was repeated three times. The participants produced a total of 270 vowels (10 x 9 x 3). All in all, 1,602 acoustic tokens (270 x 6) are examined in this paper.

1 Introduction
Stewart (1967:197) accidentally “discovered” the phonetic feature [±ATR] by reading Hockett’s (1958:78-79) explanation of the articulation of tense and lax vowels in American English. Since then almost two-dozen instrumental studies have been devoted to these vowels. Kang and Ko (2012) write that cine-radiographic, MRI, endoscopic, and ultrasound studies have been used to investigate [±ATR] vowels in African languages. They also provide an overview of the acoustic studies of [±ATR] vowels that go as far back as Lindau (1979). Except for Retord’s (1977) massive cine-radiographic dissertation on Anyi Sanvi vowels, no other instrumental description exists for Anyi vowels. This paper seeks to provide acoustic phonetic information about [±ATR] vowels in Anyi. It also seeks to determine which of the six acoustic correlates – F0, F1, F2, F3, intensity, and duration – is/are perceptually robust.

1.1. Background Information on the Language and Participants
Anyi belongs to the Akan language family that stretches from eastern Côte d’Ivoire to western Ghana. In Côte d’Ivoire, more than a million people (1,072,618) speak Anyi, which is subdivided into seven main dialects, as shown in the map below:
Seventy-five percent of all Anyi people speak the Morofu dialect (804,000). Ten participants from the Bongouanou prefecture, the main administrative center of the Moronou region, took part in this study. They are bilingual in Anyi and French and work as literacy teachers for CATA (the Anyi Literacy and Translation Center), a non-governmental organization dedicated to adult literacy in the Anyiland. The participants, who are all males, range in age from 30 to 50. At the time of the data collection there were no female literacy teachers, but this situation has changed. Three female teachers have now joined the CATA. The elicitation task consisted in producing the following sentences:

1. <ɔ’a hi> (he/she has refused to eat it)
2. <ɔ’a hɪ> (he/she has caught it)
3. <ɔ’a he> (he/she has shared)
4. <ɔ’a he> (he/she is late)
5. <ɔ’a hʌ> (it has boiled)
6. <ɔ’a hʊ> (a nonsense word)
7. <ɔ’a ho> (he/she has dug a whole)
8. <ɔ’a hɔ> (he/she has left)
9. <ɔ’a ha> (he/she has bitten).

The elicitation word in each sentence begins with /h/. These words were chosen intentionally in order to replicate Peterson and Barney’s (1952) methodology as much as possible. Countless studies of vowels have followed this methodology. Ladefoged (1996:112) explains the benefits of choosing /h/ in these kinds of acoustic phonetic studies as follows:

1 These figures are projections based on the official data from the 2000 census. It takes into account an overall 3% annual population grown in Côte d’Ivoire.
As the positions of the articulators during the sound [h] are similar to those of the surrounding sounds, such as the adjacent vowels, the frequency components in [h] sounds have relative amplitudes similar to those in vowels; but the complex wave has a smaller amplitude and no fundamental frequency, as it is not generated by regular pulses from the vocal cords.

Since [h] exists in Anyi as an allophone of /k/, Peterson and Barney’s methodology can be replicated without any problem. The entire duration of the vowel, from the onset to the offset, was measured. It was not deemed necessary for this study to take measurements at various points in the vowel because the environment in which the vowel occurred did not foster co-articulation. Furthermore, the methodology used by Peterson and Barney that is being replicated in this study did not sample vowels at multiple intervals. The onset of each vowel was easily identified because of the frication noise contained in [h]. However, it was more challenging to determine the offset of vowels. In annotating the offset, Thomas (2011:142) proposes three options:

… The same problem crops up frequently with vowels before a pause. In these cases, you have another choice to make. One option is to look for a spot where the vocal fold vibrations become more or less unrecognizable or start looking more like staticky patterns of aspiration than the sharper pattern usually evident with vocal fold vibrations. Often, the best way to determine this spot is by moving the cursor to different spots and listening; after a certain point, all you hear is aspiration, and that point is where you mark the offset. The other option is to mark the offset at the end of the recognizable aspiration, though this point may be quite difficult to define.

For this study, the offset of the vowel was determined by following the second option in Thomas’ recommendation, that is, demarcating the offset right before the point at which aspiration is heard. The measurements for one speaker were done manually to ensure that the offsets of vowels are identified accurately. Once the pattern was well established, Ryan’s (2005) Grid-maker script for Praat was used to annotate all the vowels produced by the rest of the speakers. Subsequently, Yoon’s (2008) Stress-analysis script for Praat was employed to collect all the relevant information displayed in all the tables. Six acoustic correlates, F0, F1, F2, F3, intensity, duration, are investigated so as to assess their relative robustness in the perception of [±ATR] vowels. The recordings were done on an Olympus Digital Voice Recorder WS-710. The participants wore a Panasonic head-mounted, noise cancellation fixed microphone in order to minimize environmental noise. The sentences were recorded in a quiet room on the premises of the CATA. The Institutional Review Board (IRB) of Saint Cloud State University in MN, USA, approved data collection prior to the fieldwork trips that took place in the summers of 2012 and 2013.² WavePad Sound Editor, Version 5.17 by NCH Software was used to sample of all the files at 44100 Hz at the rate of 16 bits per sample. File segmentation and editing were done through WavePad, but the measurements of the data were collected via Praat.

² The methodology and the equipment are the same as described in Koffi (2015).
Previous studies, Retord (1977:96), Quaireau (1987:27), Koffi (2009:11) have determined that Anyi has a nine-vowel system consisting of four [+ATR] vowels, four [-ATR] vowels, and of an ambidextrous vowel /a/ that straddles both groups. Table 1 classifies them according to height, backness, and [±ATR].

<table>
<thead>
<tr>
<th></th>
<th>Front</th>
<th>Central</th>
<th>Back</th>
</tr>
</thead>
<tbody>
<tr>
<td>[+ATR]</td>
<td>i, e</td>
<td></td>
<td>u, o</td>
</tr>
<tr>
<td>[-ATR]</td>
<td>i, e</td>
<td>a</td>
<td>a, o</td>
</tr>
</tbody>
</table>

Table 1: Classification of ATR Vowels

Burmeister (1983:161) contends that the Sanvi dialect of Anyi has a contrast between [ʌ] and [a]. However, Retord’s thorough instrumental analysis did not uncover any [ʌ] vowel. As for the Morofou dialect, Koffi’s (2015) acoustic vowel space study indicates conclusively that it has a nine-vowel system, as shown in Figure 4 in section 1.6. Phonologically, all the dialects of Anyi exhibit a harmonic system which causes all the vowels in a lexical root to be either [+ATR] or [-ATR]. The central vowel [a] is ambivalent in that it can partner with vowels of either set. The Morofu dialect distinguishes itself from other Anyi dialects for having three additional harmonic requirements. Koffi (2009:51-65) provides numerous examples to show that the vowels in lexical roots (excluding loanwords and compounds) must agree in fronting, in rounding, and in height.

1.2 Succinct Literature Review

It has been almost 50 years since Stewart (1967) brought [±ATR] vowels to the attention of the worldwide linguistic community. Over these five decades, numerous phonological descriptions have shed some light on the various harmonic systems that operate in many of these languages. Though phonological accounts abound, to my knowledge, there have been only about two-dozen studies that have examined [±ATR] vowels instrumentally. Five studies employed cine-radiography, MRI, endoscopic, and ultrasound technologies. The rest of the studies used an acoustic phonetic methodology. The literature review section of this paper is purposefully short because it focuses only on the last three acoustic phonetic studies available at the time of writing. There is no reason for an extensive literature review that will repeat the same pieces of information found in the literature review sections of earlier studies. The three studies reviewed in this section and discussed throughout the paper are Starwalt (2008), Kang and Ko’s (2012), and Quinn-Wriedt (2013).

Starwalt (2008) investigated F1, F2, F3, bandwidth, spectral flatness, and center of gravity (p.76) of [±ATR] vowels in 11 African languages for her Ph.D. dissertation. Kang and Ko (2012) described [±ATR] vowels in three languages in the Altaic family. Their study is important because it shows that the [±ATR] feature is not confined to Africa. The authors examined eight acoustic correlates (p. 191): F0, F1, F3, amplitude (A1, A2, A3), bandwidth (B1), Harmonics (H1, H2), and center of gravity. Quinn-Wriedt (2013) devoted her Ph.D. dissertation to Maasai vowel harmony. She investigated F1, F2, and duration. The consensus among these recent acoustic phonetic studies, as well as the previous ones, is that F1 is by far the most robust cue in the intelligibility of [±ATR] vowels. The other cues are not as salient. In these three studies, as in previous studies, the authors relied heavily on statistical analyses to determine the robustness of the acoustic correlates of [±ATR] vowels. By contrast, this paper
uses reference levels to assess the salience of F0, F1, F2, F3, intensity, and duration of [±ATR] vowels in Anyi Morofu.

1.3 Reference Levels and their Uses

Everest and Pohlmann (2015:23) note that reference levels are “widely used to establish a baseline for measurements.” Professionals and laypeople alike routinely rely on reference levels to interpret data. An example of the thermometer in a doctor’s office or in the family medicine cabinet will suffice to illustrate the usefulness of reference levels. When one does not feel well and suspects having a fever, one takes one’s temperature. The readings on the thermometer help to assess one’s condition in order to determine the next course of action. A reading of 97.0 to 99.5 °F (35.0 to 37.5 °C) indicates normal body temperature, that of 99.5 to 100.9 °F (37.5 to 38.3 °C) is interpreted as a low grade fever, that of 101 to 103 °F (38.3 to 40 °C) means a high fever, and that of 104-106 °F (41.5 to 42 °C) is interpreted as a very high fever that requires immediate medical attention.

In the hearing sciences for examples, audiologists use audiometers to assess the extent and depth of hearing loss (Ruggero and Santos-Sacchi 1998:1125-1137). Noise dosimeters are used to measure residential or occupational noise levels (Hoover and Keith 1998:1021-1043). Architectural acousticians determine the acoustical qualities of auditoriums, concert halls, large rooms, classrooms, etc. on the basis of baseline measurements (Tocci 1998:985-1003, Wetherill 1998:1005-1019). Audio engineers also use baseline measurements to assess the musical quality of sound equipment in cars, buses, airplanes, etc. (Beauchamp and Maher 1998:1427-1438). In other words, reference levels are widely used in industries and various occupations to gauge the significance of what is being measured. Even though huge amounts of experimental and empirical speech data exist, and even though such data have accumulated for more than a century, phoneticians rarely use reference levels in their analyses. Instead, they rely almost exclusively on various statistical analyses to interpret the significance of their findings. To be sure, there is nothing wrong with using statistics. However, limiting the quest of segmental intelligibility of speech sounds only to statistical measurements prevents researchers from exploring other analytical possibilities that are equally as valid. For example, the use of reference levels obviates the need to refer to complicated statistical jargon that is not easily decipherable by average readers. Furthermore, reference levels have the advantage of being applicable to all languages in the same way that reference levels in the biomedical field, in chemistry, in physics, etc. apply to all humans. Last but not least, reference levels help to put into practice the Occam Razor Principle to which linguistics wholehearted subscribes: “Widely accepted as a principle of science ever since the 12th century, it means that if two theories account for a phenomenon equally well, that theory which is simpler is better” (Radford 1986:25-6).

When using reference levels to interpret speech data, it must be kept firmly in mind that the human ear analyzes frequency and intensity logarithmically, but it analyzes duration arithmetically (Repp 1987:6). It is also important to highlight the fact that all human beings perceive speech the same way, irrespective of their native languages. All human beings use the same auditory area that Everest and Pohlmann (2015:49) represents graphically as follows:3

3 Unless otherwise specified, all the graphics are Everest and Pohlmann (2009), Fifth edition.
These uncontroversial facts about human audition make analyses based on reference levels attractive because the attendant findings are language independent. There is much to be said about Figure 2, but suffice it to note that human beings can perceive speech sounds ranging from 20 to 20,000 Hz. Their ears are not equipped to hear sound frequencies below 20 Hz (infrasonic sounds), nor can they hear frequencies above 20,000 Hz (ultrasonic sounds). The frequency bands that humans use to decipher vowels and consonants are located within the 250 to 5,000 Hz range. However, human ears are at their best in perceiving speech sounds in the 1,000 to 4,000 Hz range (Roseberry-McKibbin and Hegde 2000:550-70).

1.4 The Ear as a Transducer and a Critical Band Filter

The thermometer mentioned in 1.3 is a fitting example to illustrate how transducers and filters work in general. The College Standard Dictionary defines a transducer as follows, “Any device by means of which the energy of one power system may be transmitted to another system, whether of the same or different type.” The thermometer takes body temperature and translates it into biomedical terms that are easily understood by most people. It also has preset filters that inform its users as to whether or not they are running a fever; and the type of fever that they are running. Similarly, the human auditory system takes speech signals and transduces them into frequency, intensity, duration, nerve impulses, etc. The process is extremely complicated and not fully understood. Yet, empirical and experimental studies have shed some light on it and great strides have been made as a result. Now, computerized devices that imitate how humans perceive sounds have been invented for all kinds of purposes. The core principle underlying the
ears and all human-made transducers can be simplified (or oversimplified\textsuperscript{4}) as follows. The incoming speech signals go through a series of filters that discriminate among speech sounds according to their aerodynamic properties, most importantly, frequencies and intensity. Some filters sort out low frequency sounds, others filter out mid range frequency sounds, while other do the same for high frequency sounds. Theoretically, humans are capable of perceiving frequencies that range from 20 to 20,000 Hz, as noted earlier. Most manufactured transducers can also “hear” sounds like humans. The human capacity to perceive intensity runs from about 0 to 120 dB, but some devices have far greater dynamic ranges than humans. Human capabilities are summarized pictorially in Figure 2. These are, however, only theoretical possibilities. In actuality, the critical frequency bands that humans use to produce and perceive speech run from 20 to 5,000 Hz, as indicated in Table 2.

<table>
<thead>
<tr>
<th>No</th>
<th>Reference Levels</th>
<th>Formant Frequencies</th>
<th>Frequency Bands</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>1 Hz</td>
<td>F0</td>
<td>20–500 Hz</td>
</tr>
<tr>
<td>2.</td>
<td>60 Hz</td>
<td>F1</td>
<td>1,000 Hz</td>
</tr>
<tr>
<td>3.</td>
<td>200 Hz</td>
<td>F2</td>
<td>2,000 Hz</td>
</tr>
<tr>
<td>4.</td>
<td>400 Hz</td>
<td>F3</td>
<td>3,000 Hz</td>
</tr>
<tr>
<td>5.</td>
<td>600 Hz</td>
<td>F4</td>
<td>4,000 Hz</td>
</tr>
<tr>
<td>6.</td>
<td>800 Hz</td>
<td>F5</td>
<td>5,000 Hz</td>
</tr>
</tbody>
</table>

Table 2: Reference Levels and Formant Frequencies

Speech sounds such as vowels contain areas of concentrated acoustic energy known as formant frequencies. There are many such formants, but experts contend that the first five are worth investigating. Furthermore, for vowel intelligibility, F1 and F2 are deemed the most relevant. However, for the purposes of investigating the acoustic correlates that are most salient for discriminating between [±ATR] vowels in Anyi Morofu, the present analysis will include F0, F3, intensity, and duration.

In order to use reference levels to assess the salience of various acoustic correlates of speech, we must also familiarize ourselves with the concept of “octave.” Everest and Pohlmann (2015:602) define an octave simply as “the interval between two frequencies having a ratio of 2:1.” An important distinction is made between reference levels based on full octave bands (also known as eight octave bands) and fractional octave bands such as the 1/3-octave bands (Everest and Pohlmann 2015:12-4, 86, 529). The latter is more useful in analyses involving suprasegmentals. However, since this study deals only with segments, the eight-octave band system with center frequencies at 63, 125, 250, 500, 1,000, 2,000, 4,000, and 8,000 Hz will suffice. It is also worth remembering that the frequency doubles from one octave band to the nearest one, as shown in the previous list and below:

\textsuperscript{4} Fant (1998:1253) argues that oversimplification is unavoidable in the acoustical analysis of speech production and perception.
The first octave at which pitch is audible is 63 Hz. However, it is rounded down to 60 Hz to make calculations simpler. Octave-based reference levels are used primarily in general acoustics for audio-frequency measurements. Only a handful of acoustic phonetic analyses make use of them. Labov et al. (2006) and Labov et al. (2013) refer to them but implicitly. Baart (2010: 67) does the same in the following quote:

People do not perceive frequency differences in a linear fashion. A frequency difference of, say, 200 Hz is much more noticeable for people (as in the difference between 200 and 400 Hz) than if higher frequencies are involved (as in the difference between 2000 and 2200 Hz).

Gandour (1978:57), Fry (1979:68), and Kent and Read (2002) refer explicitly to octave band frequencies in their phonetic analyses. Manufactured transducers that imitate the human auditory spectrum have filter settings that are divided into “octaves,” which are internationally agreed upon reference levels (Pope 1998:1346). Let’s now see how the information presented so far can be used to determine which of the six acoustic correlates mentioned above is/are most prominent in perceiving [±ATR] vowels in Anyi.

1.5 The Relevance of F0 in the Intelligibility of Anyi ATR Vowels

F0 is a technical acoustic term. Its perceptual equivalent is pitch. Both terms are used interchangeably. Some phoneticians are of the opinion that F0 does not encode any segmental information. Miller (1989:2126) echoes this view by making the following statement: “It is well known, that under most conditions, the identity of a perceived vowel depends strongly on the formant values of the spectrum and is independent of voice pitch.” The fact that F0 is not deemed salient for the identification of vowels did not prevent Peterson and Barney (1952) and Hillenbrand et al. (1995) from providing F0 measurements for English vowels. In like manner, F0 measurements are provided for Anyi vowels in Table 3A:

---

5 The low frequency band is based adult males’ voice. Miller (1989:2122) provides the following reference levels for men, women, and children. The F0s for men, women, and children between the ages of 7 to 10 are respectively 133, 225, and 263 Hz. Fry (1979:68) has the following F0 for the three groups: 120 Hz for men, 225 Hz for women, and 265 Hz for children. The 13 Hz difference between the F0 of men in Miller and Fry is perceptually insignificant.
In the F0 frequency band, an acoustic distance of ≥ 1 Hz is needed to distinguish between two contiguous segments on the same octave band.

The human ear is very sensitive and can perceive a pitch difference of 0.3% Hz between two sounds within the same octave band (Young 2011:609). For ease of calculation, it has been recommended that the reverence level for F0 be rounded up to 1 Hz. Gandour (1978:57) cites a number of empirical and experimental studies in which the participants had no problem detecting pitch differences of 1 Hz on an octave band of 80 to 160 Hz. Fry (1958:144) also indicates that the magnitude of pitch variation within the same octave band does not make any difference in perception after a change of 1 Hz has been observed. In other words, once the ear has detected a pitch change of 1 Hz in the same octave band, it stops computing F0 information.

Starwalt (2008:76-80) reviewed the literature on the connection between F0 and [+ATR] vowels in African languages. For her own study, she concluded that, because spectral tilt findings were unreliable, she would not concentrate on F0. Kang and Ko (2012:191) mentioned that they would measure F0 but failed to discuss it in the rest of their paper. Quinn-Wriedt (2013) skipped F0 altogether. As for Anyi Morofu, the analysis by reference levels indicates clearly that F0 is not a salient cue for discriminating between [+ATR] and [-ATR] vowels. We see that two [+ATR] vowels, [e] (146 Hz) and [o] (152 Hz) have higher F0 values than their respective [-ATR] counterparts [ɛ] (140 Hz) and [ɔ] (139 Hz). However, we also see that two [-ATR] vowels, [ɪ] (142 Hz) and [ʊ] (147 Hz), have higher F0 values than their respective [+ATR] equivalents [i] (141 Hz) and [u] (144 Hz). In other words, the perception of [+ATR] and [-ATR] vowels in Anyi does not rest on F0.

Now, let’s contrast the information in the previous section about [+ATR] and [-ATR] vowels with those in Table 3B regarding tense and lax vowels in American English:

<table>
<thead>
<tr>
<th></th>
<th>[i]</th>
<th>[ɪ]</th>
<th>[e]</th>
<th>[ɛ]</th>
<th>[æ]</th>
<th>[a]</th>
<th>[ɔ]</th>
<th>[o]</th>
<th>[ʊ]</th>
<th>[u]</th>
<th>[ʌ]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tense (+) vs. Lax (-)</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+/−</td>
<td>-</td>
<td>+</td>
<td>−</td>
<td>+</td>
<td>−</td>
<td>−</td>
</tr>
<tr>
<td>P&amp;B</td>
<td>F0</td>
<td>136</td>
<td>135</td>
<td>NA</td>
<td>130</td>
<td>127</td>
<td>124</td>
<td>129</td>
<td>NA</td>
<td>137</td>
<td>141</td>
</tr>
<tr>
<td>HILL</td>
<td>F0</td>
<td>138</td>
<td>135</td>
<td>129</td>
<td>127</td>
<td>123</td>
<td>123</td>
<td>121</td>
<td>129</td>
<td>133</td>
<td>143</td>
</tr>
<tr>
<td>Difference</td>
<td>+1 Hz/+3Hz</td>
<td>+3 Hz</td>
<td>+3 Hz/+0Hz</td>
<td>+8 Hz</td>
<td>+4 Hz/+10 Hz</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 3B: F0 and Tense/Lax in American English

In all the examples in Table 3B, tense vowels have higher F0s than their lax counterparts. The situation of [æ] and [ɑ] is unclear because experts do not agree as to whether [ɑ] is a tense or lax vowel in American English, hence the discrepancy between the measurements found in
Peterson and Barney (P&B) and those in Hillenbrand (Hill). The fact that F0 clearly distinguishes tense vowels from their lax counterparts, but fails to discriminate between [±ATR] vowels is an indication that these features are not acoustically synonymous. Ladefoged and Maddieson (1996:300) are therefore right in cautioning linguists not to use them interchangeably.

1.6 The Relevance of F1 in the Intelligibility of Anyi ATR Vowels

F1 correlates with height and goes by the logic of inverse proportionality. This means that lower vowels have higher F1 values, while higher vowels have lower F1 values. In other words, we expect [-ATR] vowels to have higher F1 formant values than their [+ATR] counterparts. The acoustic vowel space in Figure 4 highlights the relationship between [+ATR] vowels and their [-ATR] counterparts. [+ATR] are circled for ease of identification.

![Acoustic Vowel Space of Anyi](image)

Figure 4: Acoustic Vowel Space of Anyi

The reference level formulation for determining the robustness of the F1 correlate is as follows:

---

6 Double vowels stand for [-ATR] vowels.
In the F1 frequency band, pairs of contiguous vowels whose acoustic distance is \( \geq 60 \) Hz are clearly perceived, but those whose acoustic distance falls below 60 Hz, may be subject to confusion.

As noted in 1.3, 63 Hz is the lowest octave on the F1 formant frequency band. However, to make calculations simpler, it is rounded down to 60 Hz. A cursory look at Table 4 shows that F1 is a robust cue for distinguishing between [+ATR] and [-ATR] vowels in Anyi Morofu vowels:

<table>
<thead>
<tr>
<th></th>
<th>[i]</th>
<th>[ɪ]</th>
<th>[e]</th>
<th>[ɛ]</th>
<th>[u]</th>
<th>[ʊ]</th>
<th>[o]</th>
<th>[ɔ]</th>
<th>[a]</th>
</tr>
</thead>
<tbody>
<tr>
<td>ATR</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>±</td>
</tr>
<tr>
<td>F1</td>
<td>348</td>
<td>399</td>
<td>392</td>
<td>589</td>
<td>388</td>
<td>523</td>
<td>477</td>
<td>635</td>
<td>925</td>
</tr>
<tr>
<td>Difference</td>
<td>51 Hz</td>
<td>197 Hz</td>
<td>135 Hz</td>
<td>158 Hz</td>
<td>NA</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 4: F1 and [±ATR] Vowels in Anyi

F1 is useful for assessing the salience of [±ATR] vowels for two important reasons. First, Ladefoged and Maddieson (1996:286) observe that “all languages have some variations in vowel quality that indicate contrast in the vowel height dimension.” Second, Ladefoged and Johnson (2015:207) state that F1 contains “on average 80% percent of the energy in the vowel.” This means that it is by far the most important acoustic correlates of vowels in all languages. Kent and Read (2002:33) provide the following explanation for the acoustic robustness of F1:

The first formant is typically the most intense formant, largely because of the interaction with the amplitudes of the other formants. One way of thinking about this is to say that F1 rides on the low-frequency tails of the other formant curves, so that F1 is boosted in amplitude relative to the other formants. Loudness judgments of speech tend to be highly correlated with the amplitude of F1, which is not surprising given that this formant tends to be the strongest.

An analysis by reference level shows that F1 is a robust cue for distinguishing between [+ATR] and [-ATR] vowels in Anyi. In every case, the acoustic distance between [+ATR] vowels and their [-ATR] counterparts is higher or equal to 60 Hz. In other words, one can tell a [+ATR] vowel from an [-ATR] because of the perceptual differences associated with height. Starwalt (2008), Kang and Ko’s (2012), and Quinn-Wriedt (2013) used statistical measurements such as ANOVA, t-test, and the like, to arrive at the same conclusion at which I have arrived by using reference levels. Employing reference levels does not negate the use of statistics, but as Rabiner (1998:1267) puts it, this approach “only provides first-order statistics (mean values) of the spectral parameters of the reference.”

Two quick observations need to be made about [i] and [ɪ] on the one hand, and [i] and [e] on the other. For the vowels [i] and [ɪ], even though the mean acoustic distance between them is less than 60 Hz, confusion is not expected to be widespread in everyday encounters because the acoustic distance between [i] and [ɪ] is \( \geq 60 \) Hz for four of the ten participants. This means that most people still produce them in ways that do not interfere with intelligibility.
Table 5: F1 Formant Frequencies with a focus on [ɪ] and [e]

<table>
<thead>
<tr>
<th>Speaker</th>
<th>[ɪ]</th>
<th>[e]</th>
<th>[ɛ]</th>
<th>[ɛ]</th>
<th>[u]</th>
<th>[ʊ]</th>
<th>[o]</th>
<th>[ɔ]</th>
<th>[a]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Speaker 1</td>
<td>325</td>
<td>368</td>
<td>355</td>
<td>556</td>
<td>423</td>
<td>539</td>
<td>481</td>
<td>624</td>
<td>942</td>
</tr>
<tr>
<td>Speaker 2</td>
<td>280</td>
<td>407</td>
<td>408</td>
<td>576</td>
<td>329</td>
<td>595</td>
<td>429</td>
<td>639</td>
<td>885</td>
</tr>
<tr>
<td>Speaker 3</td>
<td>307</td>
<td>435</td>
<td>408</td>
<td>623</td>
<td>345</td>
<td>431</td>
<td>414</td>
<td>653</td>
<td>983</td>
</tr>
<tr>
<td>Speaker 4</td>
<td>291</td>
<td>344</td>
<td>368</td>
<td>573</td>
<td>384</td>
<td>583</td>
<td>429</td>
<td>677</td>
<td>882</td>
</tr>
<tr>
<td>Speaker 5</td>
<td>291</td>
<td>415</td>
<td>361</td>
<td>601</td>
<td>396</td>
<td>510</td>
<td>469</td>
<td>639</td>
<td>980</td>
</tr>
<tr>
<td>Speaker 6</td>
<td>304</td>
<td>402</td>
<td>493</td>
<td>662</td>
<td>342</td>
<td>546</td>
<td>556</td>
<td>662</td>
<td>823</td>
</tr>
<tr>
<td>Speaker 7</td>
<td>338</td>
<td>378</td>
<td>381</td>
<td>536</td>
<td>473</td>
<td>506</td>
<td>544</td>
<td>634</td>
<td>981</td>
</tr>
<tr>
<td>Speaker 8</td>
<td>444</td>
<td>469</td>
<td>449</td>
<td>597</td>
<td>420</td>
<td>491</td>
<td>445</td>
<td>654</td>
<td>940</td>
</tr>
<tr>
<td>Speaker 9</td>
<td>255</td>
<td>360</td>
<td>350</td>
<td>584</td>
<td>365</td>
<td>545</td>
<td>544</td>
<td>628</td>
<td>815</td>
</tr>
<tr>
<td>Speaker 10</td>
<td>654</td>
<td>421</td>
<td>356</td>
<td>583</td>
<td>405</td>
<td>490</td>
<td>468</td>
<td>635</td>
<td>1028</td>
</tr>
<tr>
<td>Mean</td>
<td>348</td>
<td>399</td>
<td>392</td>
<td>589</td>
<td>388</td>
<td>523</td>
<td>477</td>
<td>635</td>
<td>925</td>
</tr>
</tbody>
</table>

The situation is, however, different for [ɪ] and [e] where confusion is commonplace because, for eight of the ten participants, the acoustic distance between the two vowels is below 60 Hz. Furthermore, for six of them [ɪ] and [e] are perceptually indistinguishable from each other because the acoustic distance between them is ≤ 20 Hz. Koffi (2015) describes the results of dictation test that shows that Morofu speakers confuse these two sounds aurally. A full-blown phoneme identification test using the same methodology as Miller and Nicely (1952) is needed to gauge the depth and breadth of the confusion between these two vowels. In the meantime, Koffi (2015) contends that Anyi Morofu is inexorably moving from a nine-vowel system to an eight-vowel system, and possibly to a seven-vowel system at a later date. This is not unheard of in languages of the Akan family. Mensah (1983:430) shows that [ɪ] has disappeared in Krobou and has been replaced by [e]. Similarly, Baule, a language closely related to Anyi Morofu, has moved to a seven-vowel system. The vowel [ʊ] has been swallowed up by [o]. Anyi Morofu is following suit.

1.7 The Relevance of F2 in the Intelligibility of Anyi ATR Vowels

The F2 formant is found in the 2,000 Hz frequency range. Articulatorily, it correlates with the phonetic feature [±back]. Vowels with the feature [-back] have higher F2 formant values, while those with the feature [+back] have lower F2 formant values. F2 values of central vowels range from 1,400 to 1,600 Hz because they are neither front nor back. The reference level formulation for assessing the robustness of the [±ATR] cue is as follows:

\[
\text{In the F2 frequency band, pairs of contiguous vowels whose acoustic distance is } \geq 200 \text{ Hz are clearly perceived, but those whose acoustic distance falls below 200 Hz may be subject to confusion.}
\]

---

7 This situation is reminiscent of the merger between [a] and [ɔ] in many dialects of American English. Koffi (2013:12-4) has shown that the confusion between [a] and [ɔ] in Central Minnesota English is due to the fact that the acoustic distance between these two vowels in only 4 Hz.
Steven (2004:80) explains that confusion is likely because “F2 exhibits a discontinuity or abrupt jump in frequency. Thus there tends to be a range of values of F2 within 100 Hz or so where the frequency of the spectrum is unstable.” The measurements in Table 6 show that F2 does not discriminate between [+ATR] and [-ATR] vowels in Anyi because the acoustic distance between pairs of like vowels is below the 200 Hz threshold, except for [o] (1392 Hz) and [ɔ] (1056 Hz).

<table>
<thead>
<tr>
<th></th>
<th>[i]</th>
<th>[ɪ]</th>
<th>[e]</th>
<th>[ɛ]</th>
<th>[u]</th>
<th>[ʊ]</th>
<th>[o]</th>
<th>[ɔ]</th>
<th>[a]</th>
</tr>
</thead>
<tbody>
<tr>
<td>ATR</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>±</td>
</tr>
<tr>
<td>F2</td>
<td>2206</td>
<td>2174</td>
<td>2141</td>
<td>2038</td>
<td>1249</td>
<td>1182</td>
<td>1392</td>
<td>1056</td>
<td>1486</td>
</tr>
<tr>
<td>Difference</td>
<td>32 Hz</td>
<td>103 Hz</td>
<td>67 Hz</td>
<td>336 Hz</td>
<td>NA</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 6: F2 and [±ATR] Vowels in Anyi

It is not surprising that F2 is not a robust cue for discriminating between [±ATR] vowels in Anyi because, as Ladefoged and Maddieson (1996:286, 290) have noted, the languages of the world make much more limited use of the front-back dimensions to encode phonemic contrasts. Statistically based analyses have also found that F2 is not a robust cue in several languages that have [±ATR] vowels. Quinn-Wriedt (2013:41,43) performed various tests on F2 and concluded that it is not a robust correlate in Maasai. Kang and Ko’s (2012:192-3) ANOVA tests did not find F2 to be perceptually salient in the three Altaic languages that they studied. Starwalt (2008) also ran several statistical measurements and came to the conclusion that F2 was not perceptually salient in 10 of the 11 African languages that she analyzed acoustically. Ife seemed to be the only exception. In this case, as in the previous others under consideration in this paper, the analyses by reference levels have yielded the same results as the statistically based analyses that other linguists have employed in their attempt to determine the salience of the acoustic correlate of [±ATR] vowels.

1.8 The Relevance of F3 in the Intelligibility of Anyi ATR Vowels

The F3 formant is found in the 3,000 Hz frequency range. It provides information about the position of the lips in the production of sounds. When the lips are rounded or protruded, the F3 formant values are lower than if the lips are spread or retracted. The reference level formulation for F3 is as follows:

*In the F3 frequency band, pairs of contiguous vowels whose acoustic distance is ≥ 400 Hz are clearly perceived, but those whose acoustic distance falls below 400 Hz, may be subject to confusion.*

Narayan (2008:210) writes that, perceptually, F3 is susceptible to noise and other distortions. This may explain why greater energy is needed to perceive the robustness of acoustic correlates that depend mostly on lip positions.

<table>
<thead>
<tr>
<th></th>
<th>[i]</th>
<th>[ɪ]</th>
<th>[e]</th>
<th>[ɛ]</th>
<th>[u]</th>
<th>[ʊ]</th>
<th>[o]</th>
<th>[ɔ]</th>
<th>[a]</th>
</tr>
</thead>
<tbody>
<tr>
<td>ATR</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>±</td>
</tr>
<tr>
<td>F3</td>
<td>3174</td>
<td>3055</td>
<td>2877</td>
<td>2738</td>
<td>2832</td>
<td>2716</td>
<td>2897</td>
<td>2545</td>
<td>2506</td>
</tr>
<tr>
<td>Difference</td>
<td>119 Hz</td>
<td>139 Hz</td>
<td>116 Hz</td>
<td>352 Hz</td>
<td>NA</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 7: F3 and [±ATR] Vowels in Anyi
The measurements in Table 7 show that F3 is not a robust cue for distinguishing between [±ATR] vowels in Anyi. Kang and Ko (2012:193) have come to the same conclusion and have stated categorically that F3 is irrelevant for discriminating between [±ATR] vowels in the three Altaic languages that they analyzed. Quinn-Wriedt (2013:41,43) did not investigate F3 in Maasai. Starwalt (2008) displayed F3 measurements for some of the languages she studied, but did not use them to assess the robustness of [±ATR] vowels.

1.9 The Relevance of Intensity in the Intelligibility of Anyi ATR Vowels

Styler (2013:21) states that “absolute intensity measures as given by Praat are largely meaningless.” However, this statement needs to be qualified with “unless they are interpreted logarithmically.” The smallest intensity difference that the human ear can perceive on a logarithmic scale is 1 dB (Fry 1979:92-93). A difference of 1 dB on a logarithmic scale corresponds to a difference of 3 dB on a perceptual scale. Furthermore, a logarithmic difference of 2 dB translates into 5 dB on a perceptual scale (Everest and Pohlmann 2015: 52, 86).

When calculating the robustness of the intensity correlate, we concern ourselves with the 3 dB threshold. Hasen (2001:41) notes that when the intensity difference between two sounds is ≤ 3 dB, the ear can barely discriminate between them. However, intensity differences of ≥ 5 dB are clearly heard. The reference level formulation for gauging the robustness of the intensity correlate is as follows:

Segments whose intensity difference is ≥ 5 dB are clearly perceptible, but those whose intensity difference is ≤ 3 dB are barely aurally distinguishable.

The measurements in Table 8A indicate that intensity does not discriminate between [±ATR] vowels in Anyi. In fact, Kang and Ko (2012) and Quinn-Wriedt (2013) did not investigate intensity. Starwalt (2008:291-318) did, but she noted that her results were inconclusive.

<table>
<thead>
<tr>
<th></th>
<th>[i]</th>
<th>[ɪ]</th>
<th>[e]</th>
<th>[ɛ]</th>
<th>[u]</th>
<th>[ʊ]</th>
<th>[o]</th>
<th>[ɔ]</th>
<th>[a]</th>
</tr>
</thead>
<tbody>
<tr>
<td>ATR</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>±</td>
</tr>
<tr>
<td>Intensity</td>
<td>75</td>
<td>75</td>
<td>76</td>
<td>75</td>
<td>76</td>
<td>76</td>
<td>77</td>
<td>76</td>
<td>76</td>
</tr>
<tr>
<td>Difference</td>
<td>0 dB</td>
<td>1 dB</td>
<td>0 dB</td>
<td>1 dB</td>
<td>NA</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 8A: Intensity Values of [±ATR] Vowels in Anyi

A conclusive finding would have been surprising since loudness has not been found to be phonemic in any language. It has been suggested that a bandwidth analysis might clarify the confusion between of [i] and [ɛ]. There are three main reasons for being skeptical as to whether or not bandwidth measurements can offer more insights than the F1-based explanation sketched in 1.7. The first reason is that Starwalt (2008:223-290) did not find any significant differences between the bandwidth of [+ATR] vowels and their [-ATR] counterparts for nearly all the languages she studied. Secondly, if intensity is not a robust correlate, it is extremely doubtful that bandwidth would because, more often than not, intensity and bandwidth go hand in hand. Thirdly, Kent and Read (2002:131) write that “experiments have shown that changing bandwidth of formants has very little effect on vowel perception. Apparently, the ear is not very sensitive to such changes.”
Anyi Morofu is not the only language in which intensity is not a robust acoustic cue. Lehiste and Peterson’s (1959:432) study of American English vowels shows clearly that intensity does not discriminate between tense and lax vowels:

<table>
<thead>
<tr>
<th></th>
<th>[i]</th>
<th>[ɪ]</th>
<th>[e]</th>
<th>[ɛ]</th>
<th>[æ]</th>
<th>[ɑ]</th>
<th>[ɔ]</th>
<th>[ʊ]</th>
<th>[u]</th>
<th>[ʌ]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tense/Lax</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+/−</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>Intensity</td>
<td>75.1</td>
<td>78.1</td>
<td>78.6</td>
<td>79.3</td>
<td>79.4</td>
<td>80.2</td>
<td>80.6</td>
<td>79.7</td>
<td>78.4</td>
<td>78.2</td>
</tr>
<tr>
<td>Difference</td>
<td>3 dB</td>
<td>1.3 dB</td>
<td>.8 dB</td>
<td>.9 dB</td>
<td>.2 dB</td>
<td>.2 dB</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 8B: Intensity Values of Tense and Lax Vowels in American English

Lehiste and Peterson measured the intensities of vowels in CV and VC environments. Tables 3 and 4 on page 432 of their article indicate that irrespective of the environment, the intensities of vowels were unchanged. The intensity differences of 0.3 dB between [ʌ] and [æ], of 0.5 dB between [ʌ] and [ɑ], and of 0.9 dB between [ʌ] and [ɔ] are perceptually insignificant. Even though there is a 3 dB difference between [i] and [ɪ], this intensity level is barely perceptible. It can, therefore, be concluded that intensity does not discriminate between tense and lax vowels in American English. The examples from Anyi and American English are in agreement with Ladefoged’s (2003:93) view that intensity “is seldom one of the distinguishing characteristics of a language.”

1.10 The Relevance of Duration in the Intelligibility of Anyi ATR Vowels

The use of duration as an acoustic correlate between [±ATR] vowels calls for a brief little detour into the acoustics of speech perception. Blumstein and Stevens (1980:660-661) conducted four experiments in which they found that it took the participants 25 ms to reliably identify stops consonants. These 25 ms cover the interval between the onset and the offset of the sound being produced. Repp (1987:10) summarizes these findings as follows, “The human auditory system integrates over about 25 ms and thus extracts the acoustic property relevant to place of articulation.” Everest and Pohlmann (2015:60-61) cite other experiments that have shown that integration occurs during the first 35 ms. Notice that there is a 10 ms discrepancy between these experimental findings. However, there is no discrepancy at all. Some studies such as Blumstein and Steven measure only the interval between the onset and the offset of the sound itself, whereas others include the time lag between when the sound is produced and when the ears perceive it. Miller (1989:2122) explains that it takes 10 ms for the ear to amplify incoming acoustic signals. Kent and Read (2002:11) state that “the minimal time resolution for general analysis purposes is about 10 ms.” The reference level formulation for the acoustic correlate of duration can be stated as follows:

A durational distance of ≥ 10 ms is needed to perceive clearly between two speech sounds.

Now, let’s contrast the duration characteristics of tense and lax vowels in Midwest English (MWE) vowels with those of [±ATR] vowels of Anyi Morofu. The measurements of...
MWE are from Hillanbrand et al. (1995) because Peterson and Barney (1952) did not measure duration:

<table>
<thead>
<tr>
<th>Lexical Set</th>
<th>fleece</th>
<th>kit</th>
<th>face</th>
<th>dress</th>
<th>trap</th>
<th>lot</th>
<th>cloth</th>
<th>goat</th>
<th>foot</th>
<th>goose</th>
<th>strut</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vowels</td>
<td>[i]</td>
<td>[ɛ]</td>
<td>[ɛ]</td>
<td>[æ]</td>
<td>[ə]</td>
<td>[ʊ]</td>
<td>[ʊ]</td>
<td>[ɔ]</td>
<td>[ɝ]</td>
<td>[ʊ]</td>
<td>[ɻ]</td>
</tr>
<tr>
<td>MW DUR</td>
<td>243</td>
<td>192</td>
<td>267</td>
<td>189</td>
<td>278</td>
<td>267</td>
<td>283</td>
<td>265</td>
<td>192</td>
<td>237</td>
<td>188</td>
</tr>
<tr>
<td>Difference</td>
<td>78 ms</td>
<td>11 ms</td>
<td>18 ms</td>
<td>45 ms</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 9A: Duration of Tense and Lax Vowels in American English

We see that duration is a robust acoustic correlate that distinguishes between tense and lax vowels in MWE. For each pair, the durational distance is ≥ 10 ms. Now let’s contrast this with the durational distances between [+ATR] and [-ATR] vowels in Anyi Morofu.

<table>
<thead>
<tr>
<th>Vowels</th>
<th>[i]</th>
<th>[ɪ]</th>
<th>[ɛ]</th>
<th>[ɛ]</th>
<th>[ʊ]</th>
<th>[ʊ]</th>
<th>[ɔ]</th>
<th>[ʊ]</th>
<th>[a]</th>
</tr>
</thead>
<tbody>
<tr>
<td>ATR</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>±</td>
</tr>
<tr>
<td>Duration</td>
<td>118</td>
<td>113</td>
<td>121</td>
<td>114</td>
<td>117</td>
<td>116</td>
<td>121</td>
<td>109</td>
<td>102</td>
</tr>
<tr>
<td>Difference</td>
<td>5 ms</td>
<td>7 ms</td>
<td>1 ms</td>
<td>12 ms</td>
<td>NA</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 9B: Duration of [+ATR] Vowels in Anyi

Table 9B shows that the only pair of [+ATR] vowels in Anyi where duration matters is between [o] and [ɔ]. In the remaining three pairs, it is not a robust cue. In fact, Stewart (1967:210) used duration as the main reason for not ascribing the labels “tense” and “lax” to Akan vowels:

As it was repeatedly stated by different observers, the tense vowels are necessarily lengthened in comparison with the corresponding lax phonemes. Tense vowels have the duration needed for the production of the most clear-cut, optimal vowels and in relation to them the lax vowels appear as quantitatively and qualitatively reduced, obscured, and deflected from their tense counterpart toward the neutral formant pattern. Now African languages do not, to my ears, have any difference in length between the unadvanced and advanced vowels; in Akan, in fact, verb stems with a final vowel form the past tense by lengthening that vowel irrespectively of whether it is unadvanced or advanced.

Starwalt (2008:77-78) reviewed the literature on duration and [+ATR] vowels in African languages and arrived at the following conclusion, “In sum, the results across the languages for which duration has been measured are mixed. Differences tend not to be significant and no particular pattern emerges. As no particular pattern is linked to [ATR] emerged from duration measurements in this study, the results are not reported.” Quinn-Wriedt (2013:21-44) also reviewed the literature on [+ATR] vowels and duration and found only one language where it is reportedly salient. The language is question is Akure, spoken in the Democratic Republic of Congo. As for her own study of Maasai, she made the following observation:

Because duration did not appear to be a useful cue for distinguishing the vowel pairs overall, no subsequent tests were done to look for possible confounds. Duration was also ignored in further explorations of the alternations with harmony system of Maasai (p. 44).
Kang and Ko (2012) did not investigate the relationship between duration and $\pm$ATR vowels in the three Altaic languages that they studied.

### 1.11. Summary

The paper set out to kill two birds with one stone. First, it sought to describe $\pm$ATR vowels in Anyi Morofu acoustically. Secondly, it wanted to demonstrate that using reference levels achieves the same results as using complicated statistical calculations. Reference levels have the advantage of having cross-linguistic validity because thresholds used in determining robustness of acoustic cues are language and speaker independent. Reference level-based analyses have an additional benefit in that they ensure comparability data because they follow ANSI (American National Standards Institute) and ISO (International Organization for Standardization) standards (Pope 1998:1346). They also give phoneticians the opportunity to put into practice the Occam Razor Principle to which linguistics subscribes. Both goals have been achieved. $\pm$ATR vowels of Anyi Morofu have been described acoustically. In the process, F0, F1, F2, F3, intensity, and duration have been ranked as to which of these correlates is most salient. The analyses based on reference levels have yielded the same results as those that rely heavily on more advanced statistical measurements. Irrespective of the methodology used, F1 is the most robust acoustic correlate in the intelligibility of $\pm$ATR vowels. This is not surprising for three reasons. First, all languages use height as a distinguishing feature for vowels. Second, F1 alone contains 80% of the acoustic energy found in vowels. Third, reference levels are valid analytical constructs used in many scientific fields.

### ABOUT THE AUTHOR

**Ettien Koffi** is a professor of Linguistics. He teaches the linguistics courses in the TESOL/Applied Linguistics MA program in the English Department at Saint Cloud State University, MN. He has written three linguistic books: *Language Society in Biblical Times* (1996), *Paradigm Shift in Language Planning and Policy: Game Theoretic Solutions* (2012), and *Applied English Syntax* (2015). He is the author of many peer-reviewed articles on various topics in linguistics. His primary area of specialization is at the interface between acoustic phonetics and phonology. He has extensive experience in emergent orthographies and in the acoustic phonetic and phonological description of dialect variation. He can be reached via email at: enkoffi@stcloudstate.edu.

### Reference


