Linguistic Portfolios

Volume 13 Article 7

2024

THE ACOUSTIC PHONETIC PROPERTIES OF VOICELESS FRICATIVES IN ANYI

Ettien Koffi
St. Cloud State University

THE ACOUSTIC PHONETIC PROPERTIES OF VOICELESS FRICATIVES IN ANYI

ETTIEN KOFFI

ABSTRCT

African languages are well endowed with fricatives. Common ones /f, v, s, z, ʒ, ʃ, h /, less common ones /β, ɸ, s', ç, ɣ, χ, ɦ/, and rare ones /ʃ^{w}, ɕ, ɕ^{w} , zɥ/ are all found in West African languages (Ladefoged 1968:45-66). Yet, surprisingly, there is a severe paucity of data on the acoustic phonetic properties of fricatives in these languages. This paper seeks to remedy this situation by providing a comprehensive overview of voiceless fricatives in Anyi, an Akan language spoken in Côte d'Ivoire. Twelve correlates, i.e., F0, F1, F2, F3, F4, Center of Gravity, intensity, duration, and bandwidths (B1, B2, B3, B4) are extracted from /f/ and /s/ when they occur before /i, ɪ, e, ɛ, a, ɔ, o, ʊ, u/. The findings to be discussed are based on 5,436 tokens, that is, three repetitions of /f/ and /s/ produced by nine native speakers of Anyi, across 12 correlates. The paper helps to answer questions about the putative role of these correlates in speech intelligibility in Anyi and possibly other Akan languages.

Keywords: Voiceless Fricatives, Intensity of Fricatives, Duration of Fricatives, Center of Gravity of Fricatives, Formants of Fricatives, Bandwidths of Fricatives

1.0 Introduction

The segments /f/ and /s/ are respectively the third and first most common fricatives in world languages (Maddieson 1984:42). They are also the only two phonemic fricatives in Anyi. They are studied in this paper for three reasons; the first is to satisfy acoustic phonetic curiosity since Anyi fricatives have not been measured before; the second is to gauge which correlates are robust for intelligibility; and the third is to provide data that can be used in the future for speech synthesis in Anyi. These interconnected reasons call for a comprehensive overview. Comprehensiveness is achieved by extracting 12 correlates, namely, F0, F1, F2, F3, F4, B1, B2, B3, B4, Center of Gravity (CoG), intensity, and duration. The two voiceless fricatives occur in syllable onsets and are followed by nine oral vowels, that is, /i, ɪ, e, ɛ, a, ɔ, o, ʊ, u/. By casting the analytical net very wide, I contribute an enormous amount of data that can help address the severe paucity of acoustic phonetic data on fricatives in Anyi and other African languages. However, collecting such a large amount of data also presents great organizational challenges. After several attempts, I have settled on subdividing the paper into 11 chunks. The first provides a bird's eye view on fricatives. The second describes the articulatory and phonological characteristics of fricatives in Anyi. The third deals with the methodology and the participants. Thereafter, sections and subsections are devoted to individual or groupings of correlates. The paper concludes with a summary section that highlights the most significant takeaways.

1.1 Lamenting the Paucity of Data

Many authorities have lamented the paucity of data on the acoustic phonetic properties of fricatives in world languages. For instance, Ladefoged and Maddieson (1996:173) note that "There have been surprisingly few studies of the acoustics of fricatives." Kent and Read (2002:168) add that "The acoustic description of fricatives has considerable room for improvement." Despite diligent searches in various databases for a long time, I have not come across any acoustic phonetic measurements of fricatives in African languages, except for a handful of spectrographic tracings in Ladefoged (1968:45-66). His book is replete with a language-by-language list of fricatives. However, listing fricatives and accounting for them acoustically are two different things. The lists of segments on pages 51 and 52 include four fricatives for Fante, /f, s, ɕʷ, h/ and five for Twi /f, s, ɕ, ʃʷ, h/. These two languages are relevant to this paper because they are closely related to Anyi. Yet, Anyi has only three fricatives, /f, s, h/. In fact, /h/ is not a bona fide fricative because it occurs only as an allomorph of /k/ and /t͡ʃ/. The fricatives /v/ and /z/ also occur in Anyi, not as full-fledged phonemes, but as allomorphs of /f/ and /s/ when they are preceded by non-count prefix /n/.

Anyi is not the only language where there is an asymmetry between voiceless and voiced fricatives. In fact, it is in good company with other world languages in not having phonemic voiced fricatives. Of the 266 languages that have /s/, only 96 have a corresponding /z/. Similarly, of the 135 languages that have /f/, only 67 have /v/ (Maddieson 1984:45, Table 3.2). Johnson (2012:156) explains the relative rarity of voiced fricatives in world languages as follows:

> Voiced fricatives are relatively unusual in the languages of the world, undergo a variety of phonetically motivated alternations, and are surprisingly difficult to produce. This difficulty which may underlie the cross-linguistic and phonological patterns, arises because of high volume velocity is needed to produce the turbulent noise characteristic of fricatives, and the vibrating vocal cords impede the flow of air through the vocal tract. … Because of a certain degree of airflow is necessary in order to produce turbulence, voiced fricatives may lose frication, and become glides.

Since the fricatives /v, z, h/ are allomorphs, not bona fide phonemes, they are not addressed in this paper. The exclusion of /h/ deserves an additional comment. This segment, though a fricative, is often omitted in the study of fricatives. For example, Jongman et al.'s (2000) acoustic phonetic study of fricatives in American English excludes it. Maddieson (1984:57) explains why:

> The classification of segments such as /h/ and /ɦ/ has been the subject of considerable disagreement. Although they have often been considered as members of the class of fricatives, some linguists have preferred to put them into a special class of "laryngeals" together with /ʔ/, and others have emphasized their similarity to vowels and approximants.

The exclusion of /h/ from consideration in this paper is justified on these grounds and also by the fact that it is not a full-blown phoneme. It is an allomorph of /k/ and /tʃ/ when they occur in a grammatical construction similar to the present perfect in English.

1.2 A Quick Review of the Literature on Fricatives

Jongman et al.'s (2000) acoustic phonetic study is considered the most authoritative source on fricatives to date. They provided a wide variety of measurements based on data obtained from 20 speakers of American English, 10 females and 10 males. Another important source of useful insights and measurements is Kochetov's (2017) acoustic phonetic account of sibilant fricatives in Russian. He extracted data from eight participants: four females and four males. Jassem (1962) studied fricatives in Swedish, American English, and Polish. Unfortunately, his measurements are based on only three speakers, one speaker per language. Outside of European languages, an important source on fricatives is Gordon et al. (2002). They provide various measurements, including Center of Gravity (CoG). These are the main acoustic phonetic sources on fricatives. These sources do not include any African languages. This lack of data makes it hard to determine if the measurements found in Anyi are in the mainstream or not. In other words, this paper is most likely the very first attempt to provide acoustic phonetic measurements of fricatives in an African language. This does not mean that the fricatives of African languages have not been studied phonologically. Yet, such studies are of little value for the current investigation because they are by and large impressionistic. It means that analysists relied on their own auditory acuity or that of a handful of native speakers to arrive at conclusions about fricatives. Studies that provide measurements on fricatives on African languages were non-existent at the time of the writing of this paper. For this reason, every effort is made to describe fricatives comprehensively by measuring 12 correlates. Casting such a large net is recommended because, for auditory intelligibility and for speech synthesis, we must single out which correlate(s) is/are robust. Ladefoged and Maddieson (1996:139) said as much in the quote below:

> The acoustic structure of fricatives seems to vary widely from individual to individual, but this reality reflects only the unfortunate fact that we do not yet know what it is that we ought to be describing. We do not know how to sum up what is constant, and what is linguistically and perceptually most relevant in acoustic terms. As we do not yet have an adequate model for the acoustics of fricatives, we are in a position comparable to having to describe vowels without having a notion of formants, or at least peaks in the spectrum. Our best guess is that what matters for fricatives (more especially for sibilant fricatives) is the overall intensity, the frequency of the lower cut-off point in the spectrum, and something corresponding to the center of gravity and dispersion of the spectral components above a certain threshold.

Even though this statement was made almost 30 years ago, there have not been many studies on fricatives since. Anybody who thinks otherwise has not investigated fricatives as thoroughly as I have. Except for the studies mentioned in this section, the most recent studies of fricatives in American English have been published by me and/or my students.

1.3 The Fricatives of Anyi

Three main articulatory features are used to classify fricatives according to place of articulation (POA), manner of articulation (MOA), and voicing. This scheme helps describe fricatives in Anyi as in Table 1:

		POA						
		Bilabial	Labiodental	Alveolar	Palatal	Velar	Labiovelar	Glottal
MOA	- voice		f	s				(h)
	+ voice		(v)	(z)				

Table 1: List of Anyi Fricatives

The segments /v, z, h/ are inside of parentheses because, as noted earlier, they are not full-blown phonemes. Only /f/ and /s/ are true phonemes in Anyi. Their status as phonemes is confirmed by the existence of lexical minimal pairs such as */fá/* (to take, to hold) and */sá/* (to draw water, to hide) or */fú/* (to dig a hole) vs */sú/* (to produce, to yield). The segment /f/ is labiodental, and /s/ is alveolar according to their POA. However, traditional impressionistic classifications of fricatives such as the one in Table 1 have been challenged for lacking accuracy. Shadle (1985:26, 136-7), for example, contends that fricatives can be divided articulatorily into three groups: front, mid, and back in regard to the area of greatest constriction. Jongman (1998:1721) concurs and places [f] among the fricatives that are produced in the front, while [s] belong to the "mid" group. From Shadle's (1985:26, 136-7) description of the aerodynamics of /f/ and /s/, it appears that in English, for /s/, air molecules pass over the tongue and meet an obstacle in the alveolar area where the air jet is subsequently directed "towards the lower teeth." However, "the fricative /f/ is produced by forcing air between the upper and lower lip so that it strikes the upper lip before exiting the mouth." Kochetov (2017:322, Table 1) paints a picture of a very complex articulation system in Russian fricatives. So, in the absence of acoustic phonetic data, the fricatives of Anyi displayed in Table 1 should be taken with a grain of salt. It is only after Center of Gravity (CoG) and F2 measurements are provided that we can be sure about the articulatory status of /f/ and /s/.

2.0 Participants, Methodology, and Interpretive Framework

The data that serves as the basis of this paper was recorded in the summer of 2014. The recordings were approved by the Institutional Review Board (IRB) of Saint Cloud State University where I teach. The participants are bilingual in Anyi and French, with Anyi being their dominant language. The recordings took place in a quiet room with cement walls and a cement floor on the premises of the Anyi Literacy and Translation Center, otherwise known by its French acronym of CATA. An Olympus WS-710 Digital Recorder was used for the recording and the participants wore a head mounted Krome microphone with noise cancelation capabilities. The recordings were later on exported as .wav files and sampled at 44100 Hz, with 16-bit quantization. Praat, Version 6.3.09 of March 2, 2023 was used to annotate and extract all the relevant correlates.

The participants read out loud the monosyllabic words in Table 2. They all occur in a CV context, where C stands either for the fricatives /f/ or /s/ and V stands for any one of the nine oral vowels in the language, which are /i, ɪ, e, ɛ, a, ɔ, o, ʊ, u/.

	Words with [f]	**English Translation**	**Words with [s]**	**English Translation**
1.	[fi]	to originate	[si]	to build
2.	[fɪ]	to vomit	[sɪ]	to know
3.	[fe]	nonce word	[se]	to speak
4.	[fɛ]	to suffer	[sɛ]	to be dumbfounded
5.	[fa]	to take	[sa]	to draw (water)
6.	[fɔ]	to lose weight	[sɔ]	to light (fire)
7.	[fo]	nonce word	[so]	to thicken
8.	[fʊ]	to climb	[sʊ]	to carry on the head
9.	[fu]	to dig	[su]	to bear fruit

Table 2: Wordlist Corpus

Each participant produced 27 monosyllabic words that begin with /f/.[1] Twelve acoustic correlates were extracted from each /f/. So, the nine participants produced 2,196 tokens of /f/ (9 participants x 9/f/s x 3 repetitions x 12 correlates). Speaker 9M's audio file for /s/ was intact. So, the 10 participants produced 3,240 tokens of /s/ (10 participants x 9 /s/s x 3 repetitions x 12 correlates.). All in all, the participants produced **5,436** tokens. Speaker 10F is the only female participant. The data on her /f/ and /s/ are placed in the appendix and not included in the calculations of formant measurements because women's formants are ordinarily 20% higher than their male counterparts.

2.1 Methodology and Correlate Extraction

Two different methodologies were adopted to extract the measurements. The first extracted measurements only from the beginning to the end of the frication noise of /f/ and /s/, as shown in Figure 1. Measurements that were collected via this approach are F0, intensity, duration, and CoG.

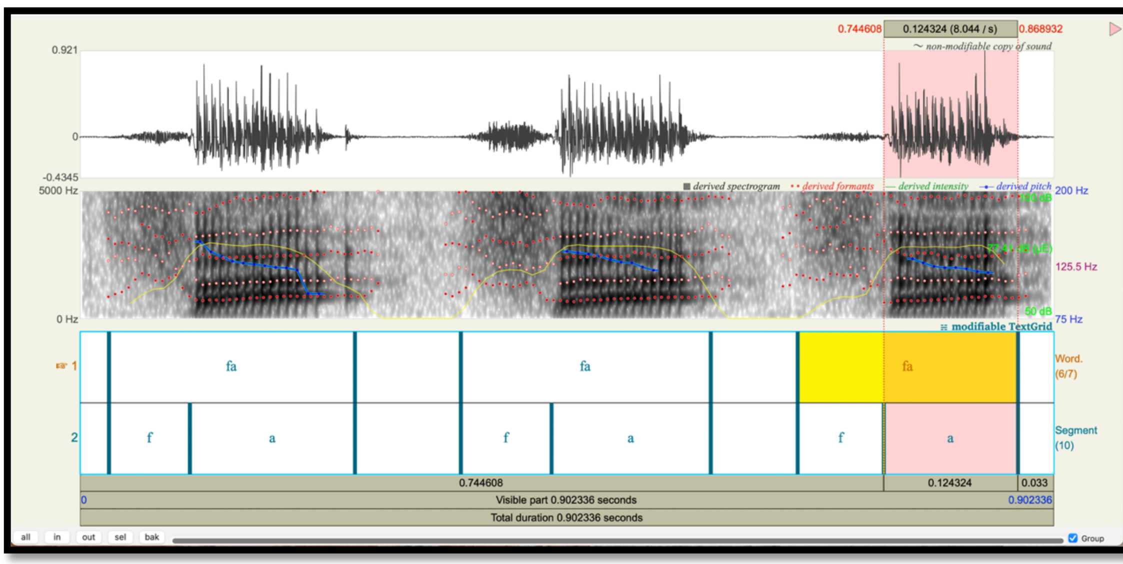

Figure 1: Sample TextGrid Annotation

The second methodology consisted of extracting formant measurements from 20 msec after the end of frication noise, as shown in Figure 2. This duration is used because it is believed that spectral characteristics of segments are most stable within this timeframe.

[1] This is so because the file containing the fricative /f/ was missing from Speaker 9 M's pronunciation. So, his /f/s were excluded from the total count. In other words, the words containing /f/ were produced by nine participants instead of 10.

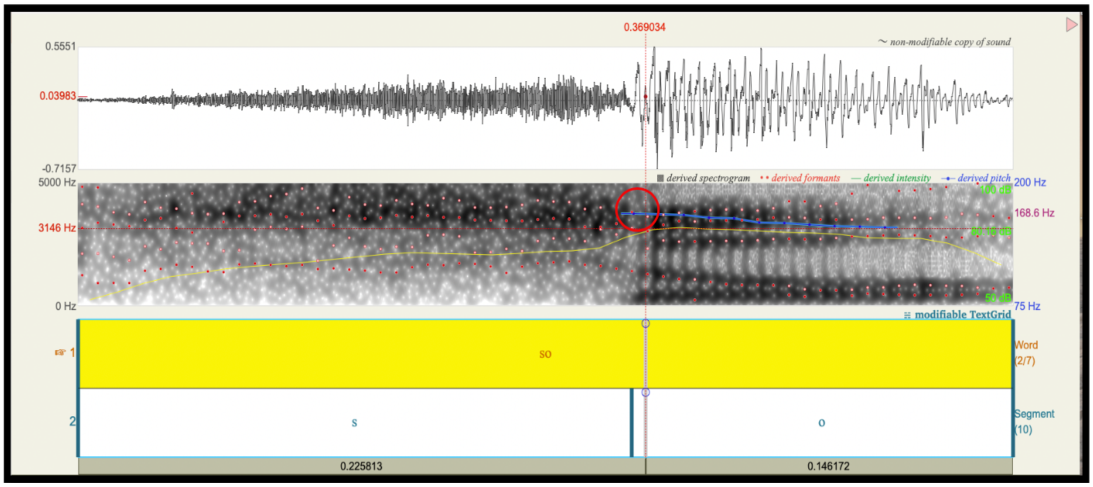

Figure 2: Illustration of Prevoicing

Once all the TextGrid annotations were complete, the Extract Sound Correlates Script (Lokousa 2023) was used to compile all the measurements and extract all correlates automatically except for CoG. The latter was extracted manually because the results of automatic extraction deviated markedly from human supervised extraction.

2.2 Data Tabulation and Presentation

I experienced considerable difficulties with data tabulation and presentation. I went back and forth several times about the best way to present the data. I eventually settled on presenting the data by correlates. Table 3 gives an overview of how key correlates are tabulated and discussed:

	F0	F1	F2	F3	F4	Dur	Ints	CoG
/s/1 before /a/	74	1259	2192	3227	4279	204	68	10068
/s/2 before /a/	74	1139	2248	3370	4464	171	70	9711
/s/3 before /a/	74	1254	2279	3296	4432	200	68	10928
AVG before /a/	**74**	**1217**	**2239**	**3297**	**4391**	**191**	**68**	**10235**

Table 3: Sample of Extracted Data

Table 3 shows that a speaker produced /s/ three times and each time it occurred before the vowel /a/. The measurements obtained are listed under each correlate. Yet, in reporting the measurements in master tables, only the arithmetic means are listed. This is so because the focus of this paper is not on intraspeaker variability but on the averaged value of fricatives. The measurements are only those of /f/ and /s/ as they occur before different vowels.

2.3 A Remark about Interspeaker Variability

Fricatives carry very important identity vectors, meaning that there is a great deal of interspeaker variability among the speakers of the same language (Koffi 2023a). This is not a new insight; it has been known for millennia. For example, in ancient Israel, the fricatives [s] vs. [ʃ] were used in the *shibboleth-shibboleth* test to verify the identity of warriors, as described in Judges 12:1-7. Experts are now simply providing measurements to substantiate this well-known variability in the production of fricatives across speakers and across languages. Ladefoged and Maddieson (1996:139) said as much by noting that "The acoustic structure of fricatives seems to vary widely from individual to individual." Johnson (2012:163) adds that "It has also been noted that there may be a substantial range of inter-speaker variability in the frequencies of the spectral peaks in fricatives." The observation about interspeaker variability becomes pertinent later when measurements are displayed. They help to explain why there are huge variations between Anyi participants. This is particularly so for Speakers 1M, 3M, and 4M with regard to some correlates.

2.4 Psychoacoustic Interpretive Framework

The analytical net has been cast wide to include 12 correlates for reasons mentioned briefly in 1.2. By extracting these dozen correlates, I hope to uncover the correlates that are most robust for intelligibility and for speech synthesis in Anyi. Intelligibility is defined **acoustically, not linguistically**. In doing so, I follow in the footsteps of Klatt (1980) who provides a wide range of measurements for fricatives in American English. Measurements are deemed intelligible if the naked ear can perceive differences between them. The concept of intelligibility is important both for verbal communication and for speech synthesis. Therefore, when measurements are extracted, their robustness must be assessed. Differences between two measurements are deemed robust if they are auditorily perceptible. The theory that has done the most to propel speech intelligibility research forward is the Critical Band Theory (CBT) which was first proposed by Harvey Fletcher and later demonstrated empirically by Goerg von Bekessy. Since I have devoted several publications to CBT, only a lapidary summary is offered here.

The human auditory perceptual system has been calculated mathematically to contain 24 critical bands, which are further subdivided in 1/3 octave. This has been acknowledged universally as matching the human auditory system (Koffi 2021:55-58). Critical bands have in turn led to the discovery of Just Noticeable Difference (JND) thresholds for gauging the intelligibility of speech sounds. For nearly 100 years, these JNDs have been used in manufacturing and testing audio engineering products. The JNDs related to the 12 correlates are used to interpret the intelligibility of various correlates of /f/ and /s/. As a reminder to newcomers to speech intelligibility research, when JNDs are used, they obviate the need to appeal to statistical analyses because for a JND to be valid, it must meet a minimum of 75% of correct responses. Relevant JNDs are used subsequently in the remainder of the paper to gauge the robustness between various correlates.

3.0 Prevoicing Analysis

In 1.3, /f/ and /s/ were classified as "voiceless" fricatives. This classification is impressionistic. To verify whether these two segments are indeed voiceless, we turn to acoustic phonetics. Normally, voiceless segments are the ones for which the vocal folds do not vibrate during articulation. Halle et al. (1957:107) observed almost 70 years ago that when this definition is applied to American English, none of its segments are voiceless because, even for those so-called voiceless consonants, there are always some amounts of vocal fold vibration. This is called **prevoicing**. Koffi (2023b) provided data to show that prevoicing is widespread in the pronunciation of Anyi stops. Figure 3 shows that prevoicing also occurs in the pronunciation of fricatives. The area corresponding to vocal fold vibration is indicated by the red oval (red online). Inside of the red oval, we see a tiny bit of the blue pitch track line. This is an indication that there is a little bit of vocal fold vibration at the end of the [s] even before the onset of the vowel [o].

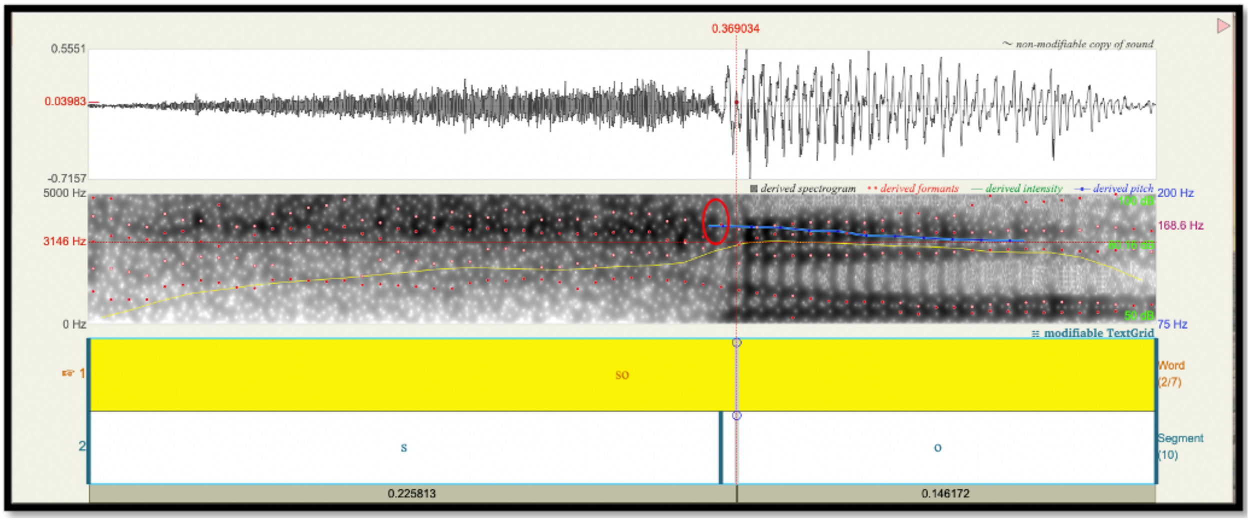

Figure 3: Illustration of Prevoicing

Even though the vocal folds vibrate a tiny bit at the end of [s], the naked ear will perceive [s] as voiceless. The 40/60 threshold explains why. This JND states that if only 40% of a segment is voiced, and 60% is unvoiced, people will still perceive it auditorily as voiced. However, if 10% or less of a segment is voiced, the naked ear perceives it as voiceless (Koffi and Lundy 2017:109-124). In Figure 3, the portion of /s/ that is prevoiced is less than 2% of the total duration of the frication noise. This explains why hearers perceive it as voiceless. So, when we say that Anyi has voiceless /f/ and /s/, what we mean is that prevoicing occurs, but in most cases, it is less that 5% of the total duration of the segment.

3.1 The Contribution of Prevoicing to Intelligibility

Prevoicing is omnipresent in Anyi, as shown in Table 4. Nine speakers produced 243 tokens of /f/. For 81 of them, Praat rendered a result of *"pitch undefined."* This means that the software did not detect any vocal fold vibration. The minimum default pitch setting in Praat is 75 Hz, which means that anything less than this threshold is taken to be voiceless. Fry (1979:68) reports that the smallest amount of vibration that the vocal folds can produce is 60 Hz. For speech analysis, the standard minimum setting for pitch is 75 Hz. So, whenever Praat says "pitch undefined," it means that the pitch detection algorithm could not find any pitch. As a matter of practice, I report 74 Hz for all such cases. It is important to note here that it is a grievous acoustic phonetics error to enter 0 Hz when Praat says "pitch undefined" because the vocal folds always vibrate, but the vibrations do not amount to 75 Hz. With this explanation, we see that in 162 out of 243 instances, the vocal folds of the participants vibrate at the rate 75 Hz or higher because of prevoicing. In other words, /f/ was prevoiced 66% of the times. The participants also prevoiced /s/ 174 out of 270 instances (64%), as reported in Table 4.

	"Voiceless" /f/	**Prevoiced /f/**	**"Voiceless" /s/**	**Prevoiced /s/**
Speaker 1M	11	16	5	22
Speaker 2M	12	15	11	16
Speaker 3M	16	11	23	4
Speaker 4M	16	11	19	8
Speaker 5M	13	14	17	10
Speaker 6M	0	27	0	27
Speaker 7M	0	27	0	27
Speaker 8M	8	19	9	18
Speaker 9M	NA	NA	1	26
Speaker 10F	5	22	11	16
Total	**81**	**162**	**96**	**174**

Table 4: Prevoicing Data

To sum it up, the participants were expected to produce 513 instances of voiceless /f/ and /s/. However, they prevoiced them in 336 cases. In other words, prevoicing occurred 65.49% of the time. We conclude that prevoicing does not only occur in Anyi, but it does so pervasively. The situation of Anyi is somewhat different from that of English described by Pirello et al. (1997:3761). They analyzed 360 utterances containing fricatives and found that 18 utterances had some voicing in the frication noise interval. Prevoicing occurred in only 5% of their data. In Anyi, on the other hand, it occurs in 65.49% of the data. Clearly, the prevoicing of fricatives is far more pervasive in Anyi than in English.

3.2 Possible Correlation between F0 and [±ATR]

In gauging intelligibility, we examine each correlate to see how the naked ear perceives it. The JND for determining intelligibility in the pitch domain is stated as follows:

Auditory Discrimination of Pitch on the F0 Frequency Bandwidth

Of two speech signals **A** and **B**, **A** is perceived as auditorily distinct from **B** if and only if there is a difference of 1 Hz or more between them.

Here, we are interested in answering two questions. First, does the vowel before which /f/ and /s/ occur cause the vocal folds to vibrate faster or slower? Second, does the pitch value in the prevoiced portion of frication noise contribute to the intelligibility of /f/ and /s/?

The measurements in Tables 5A and 5B help us answer the first question. Yet, before doing so, we must first make a short foray into the phonology of Anyi. Vowel harmony plays a very important role in the language, so much so that the vowels that occur in any given word must agree in the feature [±ATR]. The acronym stands for Advanced Tongue Root. It means that in any given word, the vowels must all be either [+ATR] or [-ATR]. The vowels [i, e, o, u] are [+ATR], while [ɪ, ɛ, ɔ, ʊ] are [-ATR]. Mismatches in [±ATR] causes the word to be ill-formed, except with central vowel [a], which can co-occur with either set.

	F0/Pitch of /f/ before Vowels								
	/fi/	/fɪ/	/fe/	/fɛ/	/fa/	/fɔ/	/fo/	/fʊ/	/fu/
Speaker 1M	102	108	141	103	75	114	179	149	180
Speaker 2M	104	105	148	105	146	135	171	74	177
Speaker 3M	106	132	103	98	74	181	86	86	141
Speaker 4M	134	74	95	110	74	82	141	87	120
Speaker 5M	154	74	99	128	74	152	126	124	164
Speaker 6M	106	154	138	121	130	134	127	137	134
Speaker 7M	102	128	145	145	142	144	143	133	116
Speaker 8M	129	120	107	89	85	78	121	154	150
AVG.	**117**	**111**	**122**	**112**	**100**	**127**	**136**	**118**	**147**
St. Dev.	19	27	22	18	33	34	29	31	24

Table 5A: F0/Pitch Measurements of [f]

	F0/Pitch of /s/ before Vowels								
	/si/	/sɪ/	/se/	/sɛ/	/sa/	/sɔ/	/so/	/sʊ/	/su/
Speaker 1M	132	137	111	99	83	91	105	119	144
Speaker 2M	93	92	89	121	83	74	97	95	132
Speaker 3M	74	98	92	74	74	100	92	74	74
Speaker 4M	113	74	85	84	74	74	74	74	133
Speaker 5M	106	74	86	90	74	98	89	74	114
Speaker 6M	142	135	126	127	124	122	135	137	134
Speaker 7M	127	128	132	137	118	123	130	133	137
Speaker 8M	112	93	104	90	100	98	121	87	136
Speaker 9M	118	93	107	91	137	123	141	143	143
AVG. /s/s	**113**	**102**	**103**	**101**	**96**	**100**	**109**	**104**	**127**
St. Dev.	20	24	17	21	24	19	23	29	21

Table 5B: F0/Pitch Measurements of [s]

Upon a closer examination of the data, we see that the vocal folds of the speakers vibrate a little faster when /f/ and /s/ occur before [+ATR] vowels than [-ATR] ones. For /f/, the average F0 before [+ATR] vowels is 130 Hz versus 117 Hz for before their [-ATR] counterparts. For /s/, F0 measurements are respectively 113 Hz and 101 Hz. We conclude, therefore, that the small amount of prevoicing in /f/ and /s/ when it precedes [±ATR] vowels contributes to intelligibility. This is so because the differences between them are higher than the JND of 1 Hz.

3.2 The Intelligibility of F0 Frication Noises

We now turn to the second question which relates to whether or not the pitch value in the prevoiced portion of frication noise contributes meaningfully to the intelligibility of /f/ and /s/. Table 5C shows that the answer is a resounding yes.

	Properties of /f/ and /s/ before Vowels									
	/i/	/ɪ/	/e/	/ɛ/	/a/	/ɔ/	/o/	/ʊ/	/u/	**AVG**
AVG. /f/s	117	111	122	112	100	127	136	118	147	121 Hz
AVG. /s/s	113	102	103	101	96	100	109	104	127	106 Hz
Difference	**4**	**9**	**19**	**11**	**4**	**27**	**27**	**114**	**20**	15 Hz

Table 5C: F0 Comparison between [f] and [s]

In the prevoiced portion of the frication noise, the vocal folds of the participants vibrate faster for /f/ (121 Hz) than for /s/ (106 Hz). Since the difference of 15 Hz is greater than the JND of 1 Hz, we conclude that the prevoicing of /f/ causes it to be distinguishable from /s/. This is what the data tells me. Unfortunately, I have no corroborating evidence from any other language because nobody has yet reported on prevoicing in voiceless fricatives, except for Pirello et al. (1997) in regard to American English. Even so, they did not concern themselves with the intelligibility of frication noise in fricatives. For this reason, we do not know if this acoustic behavior of Anyi is typical or atypical. Klatt (1980:987) to which I turn for insights in such cases has no measurements for prevoicing of /f/ and /s/ in American English.

4.0 Intensity Measurements

Ladefoged and Maddieson (1996:139) are of the view that overall intensity plays a role in the auditory intelligibility of fricatives. There is data to support this view. Stevens (2000: 1259, Table V) provides data that shows that in American English /s/ (64.9 dB) is almost twice louder than /f/ (55.7). Shadle (1985:136, 138) reports similar findings, showing that /s/ (65.82 dB) is 7.56 dB louder than /f/ (58.26 dB). In both Jongman et al. and Shadle, /s/ is two to three times louder than /f/ when considering the JND below:

Auditory Discrimination in Intensity

Of two speech signals **A** and **B**, **A** is perceived auditorily as distinct from **B** if and only if there is a difference of 3 dB or more between them.

4.1 Possible Correlation between Intensity and [±ATR]

Do /f/ and /s/ follow the same pattern in Anyi as they do in American English? Before answering this question, we must first evaluate whether or not the vowels that they precede have any impact on their loudness. The arithmetic means in Table 6A show that the intensity of /f/ remains the same irrespective of the vowel that it precedes. The intensity distances between any pairs of vowels are below the JND of 3 dB.

	Intensity of /f/ before Vowels								
	/fi/	/fɪ/	/fe/	/fɛ/	/fa/	/fɔ/	/fo/	/fʊ/	/fu/
Speaker 1M	80	80	76	76	79	76	76	79	78
Speaker 2M	73	74	70	73	67	70	71	71	75
Speaker 3M	79	78	82	82	83	73	80	84	80
Speaker 4M	63	60	73	62	61	67	69	69	69
Speaker 5M	64	65	65	66	65	66	65	63	63
Speaker 6M	76	76	76	76	76	81	75	77	74
Speaker 7M	72	69	72	69	68	69	67	68	69
Speaker 8M	71	74	73	72	71	70	70	72	74
AVG.	**72**	**72**	**73**	**72**	**71**	**71**	**71**	**72**	**72**
St. Dev.	6	6	4	6	7	4	5	6	5

Table 6A: Intensity Measurements of [f]

The previous observation also applies to /s/ in Table 6B. The acoustic distance between any two vowels is less than 3 dB. In other words, vowels do not have any impact on the loudness of /s/.

	Intensity of /s/ before Vowels								
	/si/	/sɪ/	/se/	/sɛ/	/sa/	/sɔ/	/so/	/sʊ/	/su/
Speaker 1M	78	79	78	77	76	77	77	78	77
Speaker 2M	78	79	78	78	78	76	77	78	76
Speaker 3M	79	78	82	79	80	79	79	80	77
Speaker 4M	67	65	68	67	67	64	62	63	63
Speaker 5M	74	75	74	75	74	71	69	69	69
Speaker 6M	77	78	77	78	79	77	77	77	77
Speaker 7M	65	74	68	65	68	69	70	69	70
Speaker 8M	78	78	78	77	76	77	78	77	77
Speaker 9M	77	79	80	80	78	76	75	77	76
AVG. /s/s	**74**	**76**	**75**	**75**	**75**	**74**	**73**	**74**	**73**
St. Dev.	5	4	4	5	4	4	5	5	5

Table 6B: Intensity Measurements of [s]

4.2 The Intelligibility of Intensity Frication Noises

Is the frication noise of /s/ louder than that of /f/ in Anyi, as it is in American English? The averaged measurements in the last column of Table 6C shows that, overall, /s/ (74 dB) is louder than /f/ (71 dB) in Anyi as it is in English.

	Intensity Comparison of between /f/ and /s/ before Vowels									
	/i/	/ɪ/	/e/	/ɛ/	/a/	/ɔ/	/o/	/ʊ/	/u/	**AVG**
AVG. /f/s	72	72	73	72	71	71	71	72	72	71 dB
AVG. /s/s	74	76	75	75	75	74	73	74	73	74 dB
Difference	2	**4**	2	**3**	**4**	**3**	2	2	1	3 dB

Table 6C: Intensity Comparison between [f] and [s]

Is it a universal tendency in world languages that the frication noise of /s/ would be louder than that of /f/? It is hard to tell since there is barely any data on intensity in general, let alone the intensity of /f/ and /s/. Even so, Maddieson (1984:49-50) offers a tantalizing explanation for why this may be a universal tendency. He opines that "It might be considered plausible that the more frequent sounds in the inventories of languages are those which have the greatest acoustic energy." A quick perusal of the lexical entries of Anyi shows that /f/ has 224 entries, while /s/ has 292. By Maddieson's logic, /s/ is slightly louder than /f/ in Anyi because the former has 68 more entries than the latter. This logic based on frequency of distribution should also work for English. Let's see if it does. The *Oxford Advanced Learner's Dictionary* (2000) has 77 pages (448-525) for /f/, but 194 pages (1126-1320) for /s/. The same distributional logic explains why /s/ is considerably louder in English than /f/. The page count shows that /s/ occurs more than twice as frequently as /f/ does. This also may explain why /s/ is almost three times louder than /f/.

Leaving frequency of distribution aside, let's turn to Johnson (2012:154) who appeals to aerodynamics to explain why /s/ is louder than /f/:

> The amplitude of turbulent noise is determined by the velocity of air molecules as they pass through a channel. …The faster the air molecules move, the louder the sound. … The narrower the channel, the louder the turbulent noise. … In addition to being produced when a jet of air escapes from a narrow channel, turbulent noise is also produced when a jet of

air stream hits a downstream obstacle. The presence of an obstacle results in increased amplitude of turbulent noise. It can be argued that almost all fricative noises involve turbulence produced by airflow hitting an obstacle.

Aerodynamics offers a plausible explanation for why /s/ is louder than /f/. Yet, all things considered, the difference of 3 dB between the frication noise of /f/ (71 dB) and /s/ (74 dB) in Anyi is not very robust. It barely clears the JND of 3 dB, while it does so robustly in English. The aerodynamic explanation works but in a very limited fashion in Anyi. When we look deeper into the data, we see that /s/ is auditorily louder than /f/ only before /ɪ, ɛ, a, ɔ/ but not before /i, e, o, ʊ, u/. In other words, intensity is robust for speech intelligibility in only 5 out of 9 vowels. So, the overall relative functional load (RFL) of intensity is only 55.55%, which is only slightly better than chance. We conclude, that intensity is only marginally robust in Anyi, but it is very robust in English.

5.0 Duration Measurements

The following JND is applied when gauging the intelligibility of the duration correlate:

Auditory Discrimination in Duration

Of two speech signals **A** and **B** lasting less than 200 msec, **A** is perceived as auditorily distinct from **B** if and only if there is a difference of 10 msec or more between them.[2]

Historically, the duration of the frication noise for intelligibility has been a mixed bag. It is auditorily robust in some languages but not in others. Gordon et al. (2003) report that in Gaelic, plain /f/ lasts 74 msec, while plain /s/ lasts 130.4 msec. The difference of 56.4 msec shows that duration is a very robust cue in Gaelic. Duration is also robust in Toda, another language analyzed by Gordon et al. (2003), with 111.2 msec for /f/ vs. 198.3 msec for /s/. The difference of 87.1 msec is eight times the JND. Duration is barely robust in other languages, including American English and Russian. Jongman et al.'s (2000:1260) data shows that only 12 msec separate /f/ (166 msec) from /s/ (178 msec). Kotechev's (2017:323) measurements show that duration is barely significant in Russian. There are languages in which duration is not robust at all. Such is the case of Chickasaw in which, according to Gordon et al. (2003), /f/ lasts 115.5 msec, while /s/ is 123.6 msec long. The difference of 8.1 msec is below the threshold for intelligibility.

5.1 Possible Correlation between Duration and [±ATR]

Does the duration of the frication noise of /f/ and /s/ contribute to intelligibility in Anyi? This question is answered in two steps. First, we examine the data to see if the duration of fricatives varies in accordance with the vowels that they precede. Secondly, we compare the overall duration of fricatives regardless of vowel context. Tables 7A and 7B help to answer the first question. We will turn later to Table 7C in answering the second question.

[2] The JND for duration changes for segments that last longer than 200 msec. However, since the mean duration of fricatives in Table 7C do not exceed 200 msec, we stay with this JND. I should note in passing that the duration of /f/ and /s/ in Anyi are similar to those of English discussed by Pirello et al. (1997:3759, Table I).

	Duration of /f/ before Vowels								
	/fi/	/fɪ/	/fe/	/fɛ/	/fa/	/fɔ/	/fo/	/fʊ/	/fu/
Speaker 1M	255	235	241	249	260	207	228	203	298
Speaker 2M	176	143	176	134	187	180	184	162	170
Speaker 3M	231	195	190	186	220	151	197	178	133
Speaker 4M	269	236	216	254	277	276	239	205	136
Speaker 5M	110	86	149	154	86	142	131	140	135
Speaker 6M	158	119	147	123	122	122	137	140	126
Speaker 7M	46	64	136	71	98	98	100	70	86
Speaker 8M	237	229	205	170	181	158	195	178	242
AVG.	**185**	**163**	**182**	**167**	**178**	**166**	**176**	**159**	**165**
St. Dev.	77	69	37	62	72	55	49	43	69

Table 7A: Duration Measurements of [f]

	Duration of /s/ before Vowels								
	/si/	/sɪ/	/se/	/sɛ/	/sa/	/sɔ/	/so/	/sʊ/	/su/
Speaker 1M	265	209	207	260	169	235	225	224	223
Speaker 2M	149	197	183	187	172	213	215	164	158
Speaker 3M	204	169	131	153	171	188	209	209	176
Speaker 4M	277	167	270	249	258	230	247	124	182
Speaker 5M	199	199	184	221	160	194	211	187	206
Speaker 6M	192	203	183	190	155	157	222	173	189
Speaker 7M	249	176	157	132	182	216	164	170	169
Speaker 8M	265	230	207	282	169	212	238	267	234
Speaker 9M	169	159	206	186	174	180	186	185	199
AVG.	**218**	**189**	**192**	**206**	**178**	**202**	**213**	**189**	**192**
St. Dev.	46	23	38	50	30	25	25	40	25

Table 7B: Duration Measurements of [s]

The most important insight is that the duration of the frication noise of /f/ varies according to the feature [±ATR] but that of /s/ does not. The frication noise of /f/ is longer when it precedes [+ATR] (177 msec) than before [-ATR] (163 msec) vowels. The difference of 14 msec is greater than the JND of 10 msec. Jongman et al. (2000:1260) report a similar behavior for /f/ and /s/ in American English when they precede tense and lax vowels. However, in Anyi /s/ does not vary much whether it precedes [+ATR] vowels (204 msec) or [-ATR] vowels (197 msec). The difference of 7 msec is below the JND. With regard to /s/, duration before [±ATR] vowels applies only for two pairs, [si] (218 msec) vs. [sɪ] (189 msec) and [so] (213 msec) vs. [sɔ] (202 msec), where differences are respectively 29 msec and 11 msec. For the pairs [se] (192 msec) and [sɛ] (206 msec), the opposite obtains, namely the [-ATR] is longer than the [+ATR] vowel. As for [u] (192 msec) vs. [ʊ] (189 msec), the difference of 3 msec is imperceptible to the naked ear. The take-away is that the frication noises of /f/ and /s/ behave differently.

5.2 The Intelligibility of Duration Frication Noises

Now, we turn to the duration of the frication noises of /f/ and /s/. The measurements in Table 7C indicate affirmatively that duration is a robust cue because there is a difference of 26 msec between the two fricatives. The average duration of /f/ is 171 msec, versus 197 msec for /s/.

	Duration Comparison between /f/ and /s/ before Vowels									
	/i/	/ɪ/	/e/	/ɛ/	/a/	/ɔ/	/o/	/ʊ/	/u/	**AVG**
AVG. /f/s	185	163	182	167	178	166	176	159	165	171 msec
AVG. /s/s	218	189	192	206	178	202	213	189	192	197 msec
Difference	**33**	**26**	**10**	**39**	0	**36**	**37**	**30**	**27**	26 msec

Table 7C: Duration Comparison between [f] and [s]

It cannot be ignored that the duration of the frication noises of /f/ and /s/ is identical when they precede /a/.[3] As has been noted all along, /a/ is a central vowel. Therefore, it is not a surprise that /f/ and /s/ do not vary when they occur before /a/. The lack of distinction can potentially cause words such as /sá/ (to draw) and /fá/ (to hold) to be confused. Fortunately, other cues such as F0, F2, and CoG help to differentiate between them.

6.0 CoG Measurements

CoG has been deemed a robust cue for discriminating between fricatives. Ladefoged and Maddieson (1996:139) singled it out as a correlate of great potential. Subsequent studies have confirmed its usefulness. To explain to my students what CoG is, I use the metaphor of the eye of a hurricane. It is where the maximal power resides. To get an accurate reading about the strength of a hurricane, the National Oceanic and Atmospheric Administration in the USA sends probes deep into its eye. In other words, the CoG is the place of the maximum concentration of power when fricatives are produced. This analogy is fitting because fricatives are turbulent noises inside the oral cavity when high velocity air molecules collide with a place of articulation. For /f/, the constriction area is between the lower lips and the upper teeth. This is why it is described phonetically as a labiodental fricative in many languages. The exact constriction area for /s/ is unclear in many languages. It can be apical, alveolar, or even postalveolar (Ladefoged and Maddieson 1996:151, Figure 5.9 and 164, Table 5.7). It is only by extracting CoG measurements that one can discern how /s/ is really produced.

CoG measurements lead to two separate JND thresholds depending on how they register on the frequency bandwidths. For segments between 4000-4999 Hz, the JND is calculated on the F4 bandwidth. For segments between 5000-5999 Hz, the JND is based on F5. For those between 6000-6999 Hz, the JND is based on F6. When two measurements span across two bandwidths, the segment with the higher bandwidth is used for the calculations. In most languages, CoG measurements are between 4000-5999 Hz.[4] So, we rely on F4 and F5 for the JNDs listed below:

Auditory Discrimination on the F4 Frequency Bandwidth

Of two speech signals **A** and **B**, **A** is perceived as auditorily distinct from **B** if and only if there is a difference of 600 Hz or more between them.

Auditory Discrimination on the F5 Frequency Bandwidth

Of two speech signals **A** and **B**, **A** is perceived as auditorily distinct from **B** if and only if there is a difference of 800 Hz or more between them.

[3] Pirello et al. (1997:3759) found that place of articulation plays a role in the durational characteristics of fricatives in English. This holds true for Anyi, except when /f/ and /s/ occur before /a/.

[4] The JND of critical bands are found in Pope, J. (1998:1347).

As noted previously, several studies have found that CoG is a robust correlate for discriminating between fricatives. For American English, Jongman et al. (2000:1257, Table I) found CoG to be discriminating between /f/ (5108 Hz) and /s/ (6133 Hz) because the difference of 1025 Hz between is greater than the JND of 1000 Hz. Gordon et al. (2002) presents various CoG measurements from seven languages and wrote on page 167 that "Gravity center frequencies differentiate many of the fricatives in the examined languages." Even though this statement is true for many of the fricatives they studied, we are here interested only in the languages that have both /f/ and /s/. In Chickasaw, the average CoG of /f/ is 4562 Hz, while that of /s/ is 5163 Hz. Since the difference of 601 Hz is below the JND of 800 Hz, we conclude that CoG is not a robust cue. In Gaelic, the CoG of the plain /f/ is 4415 Hz, while that of /s/ is 4884 Hz. Since the difference of 469 Hz also falls short of the JND of 600 Hz, we also conclude CoG is not a robust cue in Gaelic. In Toda, the CoG does not discriminate between /f/ (4268 Hz) and the plain [s] (4529 Hz) because the difference between them is only 261 Hz instead of the JND of 600 Hz. For the languages reviewed here, CoG is discriminatory only in English.

6.1 Interspeaker Variability in CoG

Before discussing the CoG measurements in Anyi, a quick detour is made to remind the reader of the fact that Ladefoged and Maddieson (1996:139, 173) found a great deal of interspeaker variability in the production of fricatives. This is particularly true in the pronunciation of Speakers 1M and 3M who have extremely low CoG for /f/. The measurements reported here are not erroneous. Some speakers naturally have low CoGs. I learned this the hard way one year when I nearly failed a student in my acoustic phonetics course for unusually low CoG values. I presumed that she entered measurements willy-nilly in order to avoid going through the tedious steps of extracting CoG measurements in Praat. She protested that she did not fabricate the numbers. I redid the measurements and, much to my embarrassment, I found out that she was right. Subsequently, I had her record her mother and younger sister and they too had very low CoG values. When CoG measurements are averaged across speakers, the fact that some speakers produced very low CoG values is hardly ever noted. However, this is something to anticipate.

6.2 Possible Correlation CoG and [±ATR]

How robust is the CoG of the frication noises of /f/ and /s/ in Anyi? Does it vary in accordance with the feature [±ATR] of the vowels that they precede? The first question is answered in 6.3. The second is answered in this section.

	CoG of /f/ before Vowels								
	/fi/	/fɪ/	/fe/	/fɛ/	/fa/	/fɔ/	/fo/	/fʊ/	/fu/
Speaker 1M	1467	1793	3016	4155	1467	238	2233	241	1847
Speaker 2M	4518	4736	4513	3895	4564	6228	5961	4624	5744
Speaker 3M	1337	2069	291	830	280	4371	1178	168	2113
Speaker 4M	5746	6870	2118	6388	7057	3728	2679	3125	5114
Speaker 5M	6227	6558	7023	5941	7025	6756	7301	6062	6287
Speaker 6M	4092	5321	4911	5064	4044	4186	3080	4476	4789
Speaker 7M	4267	4014	4546	4354	8097	5759	5128	4923	5539
Speaker 8M	9860	9014	9765	8898	8109	6215	9215	8614	6589
AVG.	**4689**	**5046**	**4522**	**4940**	**5080**	**4685**	**4596**	**4029**	**4752**
St. Dev.	2736	2450	2923	2321	3009	2111	2778	2843	1807

Table 8A: CoG Measurements of [f]

	CoG of /s/ before Vowels								
	/si/	/sɪ/	/se/	/sɛ/	/sa/	/sɔ/	/so/	/sʊ/	/su/
Speaker 1M	4405	6936	6857	7405	6989	6644	7032	6743	6591
Speaker 2M	4600	5545	8011	7019	5770	6963	6860	6964	6817
Speaker 3M	3957	2009	1080	2759	1741	4857	5402	4252	5431
Speaker 4M	7261	7010	7272	7436	7311	6544	6144	6623	4386
Speaker 5M	7008	7201	7196	7096	7107	6373	6540	6467	6044
Speaker 6M	6367	6448	5806	6179	4990	4740	5010	4529	4948
Speaker 7M	7624	7693	5940	6436	6582	4980	4705	4086	3694
Speaker 8M	8039	6891	7318	7924	7297	7187	7016	6218	6760
Speaker 9M	6388	1669	1167	974	2873	5002	5051	4712	5483
AVG.	**6183**	**5711**	**5627**	**5914**	**5628**	**5921**	**5973**	**5621**	**5572**
St. Dev.	1502	2273	2643	2395	2052	1003	939	1193	1093

Table 8B: CoG Measurements of [s]

The measurements in Tables 8A show that [±ATR] has no impact on the frication noises of /f/ because for any two vowel pairs, CoG measurements fall short of the JND of 600 Hz, except for [u] (4752 Hz) [ʊ] (4029 Hz) where the difference is 723 Hz. However, this a fluke due to the extremely low CoG values produced by Speakers 1M and 3M for the vowel [ʊ]. For /s/ where they did not produce low CoG values for [ʊ], we see that CoG is not a robust cue. We conclude, therefore, that the frication noises of /f/ and /s/ do not vary in accordance with [±ATR].

6.3 The Intelligibility of CoG Frication Noises

The measurements in Table 8C help answer the question regarding the robustness of the CoG values for frication noises of /f/ and /s/. The difference of 1090 Hz between /f/ (4704 Hz) and /s/ (5794 Hz) shows that CoG is robust for intelligibility. Anyi belongs to a large number of languages in which CoG is robust for intelligibility (see Gordon et al. 2002:167).

The measurements also give us some indications about the articulatory characteristics of /f/ and /s/. Evers (1998:350) explains that a higher CoG is indicative of front cavity and a smaller constriction area. The fact that /s/ has a greater CoG than /f/ means that the area of constriction for /s/ is smaller than that of /f/. This suggests that /s/ in Anyi can be described as an apical dental fricative whereas /f/ is a labiodental fricative. The surface area where constriction occurs in the production of /f/ is larger than the area where the tip of the tongue touches the upper teeth when /s/ is produced. Given this, the impressionistic description of /s/ in Table 1 should be amended from "alveolar" to "dental."

	CoG Comparison of /s/ and /f/ before Vowels									
	/i/	/ɪ/	/e/	/ɛ/	/a/	/ɔ/	/o/	/ʊ/	/u/	**AVG**
AVG. /f/s	4689	5046	4522	4940	5080	4685	4596	4029	4752	4704 Hz
AVG. /s/s	6183	5711	5627	5914	5628	5921	5973	5621	5572	5794 Hz
Difference	**1494**	**665**	**1105**	**974**	**548**	**1236**	**1377**	**1644**	**820**	1095 Hz

Table 10C: CoG Comparison between [f] and [s]

Finally, we note in passing that when /f/ and /s/ precede [ɪ] and [a], the CoG values fall below the JND. This suggests that /fí/ (to throw up) and /sí/ (to know) on the one hand, and /sá/ (to draw) and /fá/ (to hold), on the other could be confused if no other acoustic cues were available to help in the segmental discrimination process. However, F0 and intensity help in this regard.

7.0 Formant Extraction Methodology

Extracting the formants of voiceless consonants is generally challenging. If data is taken from the segment itself, the results become incorrect. For example, we see in Figure 4 that if F1 data is extracted from the frication noise itself, its value is 1283 Hz, which is unreasonable. To avoid such errors, Klatt (1980:973) recommends extracting data from 5-15 msec after the voiceless segment has ended. However, I opted to extract measurements 20 msec after the frication noise because the speech digitalization community believes that signals are stable by 20 msec installments. So, all formant measurements in this paper are taken 20 msec after the frication noise has ended, as shown in Figure 4. In so doing, I replicate the methodology used in Stevens (1992:2982).

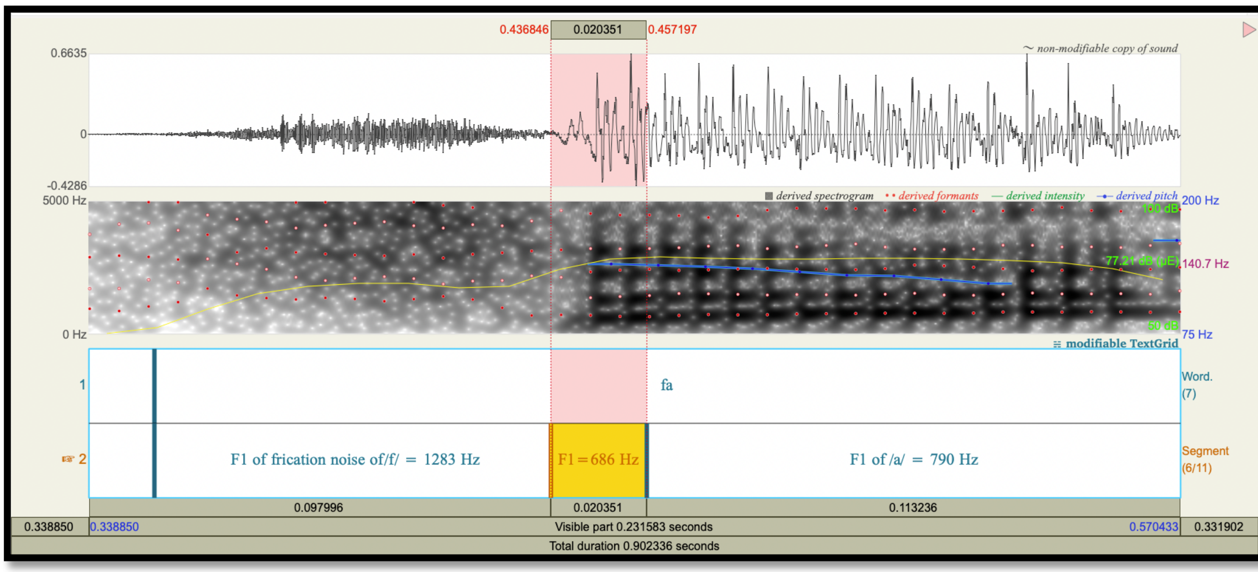

Figure 4: Formant Extraction Methodology

7.1 F1 Measurements

It is hard to find F1 measurements for fricatives because many studies do not aim at providing data for speech synthesis. The only measurements available for comparison are Kochetov (2017:345, Table A7), Jassem (1962:15-17), and Klatt (1980:987, Table III), and Jassem (1962:15-17). The first two are based on natural language data whereas the latter is based on synthesized speech. Regardless, the question to be answered is whether or not mouth aperture is robust for discriminating between the frication noises of /f/ and /s/. To help answer this question, we resort to the JND below:

Linguistic Portfolios – ISSN 2472-5102 –Volume 13, 2024 | 114

Auditory Discrimination on the F1 Frequency Bandwidth

Of two speech signals **A** and **B**, **A** is perceived as auditorily distinct from **B** if and only if there is a difference of 60 Hz or more between them.[5]

We will attempt to answer this question from two perspectives. The first is whether or not the F1 values of the segments /f/ and /s/ change considerably in regard to the feature [±ATR] of the vowels that follow them. The second is whether or not one can differentiate between /f/ and /s/ just by hearing them, irrespective of the vowels that follow them. The first question is answered in 7.2 and the second in 7.3.

7.2 Possible Correlation between F1 and [±ATR]

The correlation between F1 and the [±ATR] is very murky. The average F1 measurements for /f/ is 554 Hz in Anyi. Jassem (1962:15-17) reports that the F1 values for Swedish and Polish are respectively (480 Hz) and (540 Hz). These measurements are similar to the one in Anyi. In articulatory terms, this F1 corresponds to mouth aperture when the vowels [e] or [ɛ] are produced. Mouth aperture is considerably different when /f/ precedes [i] (1375 Hz) vs. [ɪ] (1282 Hz) and the pair [o] (1282 Hz) vs. [ɔ] (1164 Hz). In these instances, the F1 differences of 93 Hz and 118 Hz show that the feature [±ATR] is robust for intelligibility since the differences are greater than the JND of 60 Hz. However, the same is not true for when /f/ precedes the pairs [e] (1323 Hz) vs. [ɛ] (1267 Hz) and [u] (1201 Hz) vs. [ʊ] (1142 Hz), with respective differences of 56 Hz and 59 Hz. Since these values are below the JND of 60 Hz, it cannot be claimed that the feature [±ATR] is robust for intelligibility.

	F1 of /f/ before Vowels								
	/fi/	/fɪ/	/fe/	/fɛ/	fa/	/fɔ/	/fo/	/fʊ/	/fu/
Speaker 1M	423	501	602	529	702	575	521	569	490
Speaker 2M	597	590	620	658	713	585	447	407	440
Speaker 3M	411	492	603	662	699	654	614	531	508
Speaker 4M	446	477	517	643	767	669	577	501	493
Speaker 5M	417	453	514	515	613	633	520	508	416
Speaker 6M	433	468	596	606	715	642	573	496	450
Speaker 7M	534	524	544	625	670	626	509	535	426
Speaker 8M	501	534	599	656	709	685	519	493	438
AVG.	**470**	**504**	**574**	**611**	**698**	**633**	**535**	**505**	**457**
St. Dev.	67	43	42	58	43	38	51	47	34

Table 11A: F1 Measurements of [f]

When /s/ occurs before [+ATR] vowels, its F1 values are greater for [i] (1274 Hz) vs. [ɪ] (1174 Hz) and [u] (1339 Hz) vs. [ʊ] (1258 Hz). The differences are respectively 100 Hz and 81 Hz. But this is not the case for when it precedes [e] (1251 Hz) vs. [ɛ] (1204 Hz) and [o] (1240 Hz) and [ɔ] (1233 Hz). The differences of 47 Hz and 7 Hz are below the JND.

[5] Rabiner and Juang (1993:152) list slightly different set of JNDs. The JND of F1 is 62 Hz, that of F2 is 158 Hz, for F3, the JND is 355 Hz, while the JND of F4 is 480 Hz. It is important to keep in mind that the differences between these JNDs and those used in this paper do not amount to much on the 1/3 frequency bandwidth. It is universally accepted that the 1/3 frequency bandwidth replicates as faithfully as mathematically possible how the naked ear processes frequency data.

	F1 of /s/ before Vowels								
	/si/	/sɪ/	/se/	/sɛ/	/sa/	/sɔ/	/so/	/sʊ/	/su/
Speaker 1M	403	449	454	494	565	562	526	553	464
Speaker 2M	430	483	486	507	674	589	520	487	406
Speaker 3M	431	488	509	572	656	546	557	455	367
Speaker 4M	443	481	479	561	655	603	513	467	420
Speaker 5M	413	473	563	584	611	534	593	493	480
Speaker 6M	395	454	538	560	652	520	531	482	455
Speaker 7M	445	458	546	555	651	558	503	463	403
Speaker 8M	403	408	428	495	665	562	502	476	413
Speaker 9M	453	467	555	571	680	604	505	465	449
AVG.	**424**	**462**	**506**	**544**	**645**	**564**	**527**	**482**	**428**
St. Dev.	21	24	47	35	35	29	29	29	35

Table 11B: F1 Measurements of [s]

The correlation between F1 and [±ATR] is robust only when /f/ and /s/ occur before the high vowels [i] vs. [ɪ]. We see that /f/ and /s/ are produced with a smaller mouth opening before [-ATR] vowels than with [+ATR] vowels. This generalization concerns only 2 of 9 vowels, that is, 23%. This means that in 77%, the feature [±ATR] does not apply. So, it is hard to draw a firm conclusion with regard to the robustness of F1 when /f/ and /s/ precede vowels.

7.3 The Intelligibility of F1 Frication Noises

The second question, namely whether or not, by hearing the frication noises of /f/ and /s/ with one's naked ears, the hearer can differentiate between them. The answer to this question is both yes and no. The data in Table 11 suggests that auditory discrimination is possible if both voiceless fricatives occur before /e, ɛ, ɔ/, but not possible before the remaining six vowels. So, auditory neutralization occurs in 3 out of 9 vowels (34%). However, for the remaining six vowels (66%), auditory discrimination is impossible. The measurements suggest that no matter how intently somebody listens with their naked ears, they cannot perceive any difference between the syllable /fo/ and /so/ because only 8 Hz separates them. Nobody can perceive frequency distances that are ≤ 20 Hz.

	F1 Properties of /f/ and /s/ before Vowels									
	/i/	/ɪ/	/e/	/ɛ/	/a/	/ɔ/	/o/	/ʊ/	/u/	**AVG**
AVG. /f/s	470	504	574	611	698	633	535	505	457	554 Hz
AVG. /s/s	424	462	506	544	645	564	527	482	428	509 Hz
Difference	**46**	**42**	**68**	**67**	**53**	**69**	**8**	**23**	**29**	45 Hz

Table 11C: F1 Comparison between [f] and [s]

Leaving vowel contexts aside, the arithmetic means show that the naked ear cannot perceive any difference between the frication noises of /f/ and /s/ on the F1 frequency bandwidth. The 45 Hz that separates them is below the JND of 60 Hz. Anyi is not the only language in this situation. Kochetov (2017:345, Table A7) shows that F1 does not discriminate between /f/ and /s/ in Russian. Klatt (1980: 987, Table III) also confirms this for English because the distance between the F1 of /f/ (340 Hz) and /s/ (320 Hz) is only 20 Hz.

8.0 F2 Measurements

The articulatory correlate of F2 is [± anterior]. In phonological circles, /f/ and /s/ are classified as [+anterior], which means that either the lips or the alveolar area are involved in the production of these two segments (Fromkin et al. 2017:234). In assessing whether the frication noises of /f/ and /s/ are auditorily distinguishable, we use the JND below:

Auditory Discrimination on the F2 Frequency Bandwidth

Of two speech signals **A** and **B**, **A** is perceived as auditorily distinct from **B** if and only if there is a difference of 200 Hz or more between them.

Jongman et al. (2000:1259) studied fricatives in American English and found that males produce /f/ with an F2 value of 1509 Hz, and /s/ with a value of 1697 Hz. Since the difference of 188 Hz is below the threshold of 200 Hz, we conclude that both /f/ and /s/ are similar with respect to F2. For females, the F2 of /f/ is 1815 Hz, and that of /s/ is 1967 Hz. Here too, the difference of 152 Hz between them suggests that they are produced similarly.[6] Jassem (1962:6, 17, Tables 1a, 1b, 1c) provides measurements from a single speaker in American English, Polish, and Swedish. The F2s of /f/ are respectively 1550 Hz, 1620 Hz, and 1520 Hz. Those of /s/ are 1600 Hz, 1900 Hz, and 1800 Hz. We that F2 is not robust in American English since the difference between /f/ (1550 Hz) and /s/ (1600 Hz) is only 50 Hz. Yet, in the same study, Jassem (1962:16) found that F2 is a robust correlate cue in Polish because the difference between /f/ (1620 Hz) and /s/ (1900 Hz) is 280 Hz, which is greater than the JND. In Swedish, F2 is also a robust cue because the difference between /f/ (1520 Hz) and /s/ (1800 Hz) is 280 Hz. In other words, F2 is a robust cue in some languages but not in others. How is it in Anyi?

8.1 Possible Correlation between F2 and [±ATR]

Zsiga (2013:140) correlates F2 measurements directly with the place(s) of articulation of consonants. She notes that an F2 of 1600-2000 Hz corresponds to an alveolar pronunciation. A glance at Tables 12A and B show that articulatorily, both /f/ and /s/ are produced further to the front of the mouth than in the languages reviewed in 8.0. Johnson (2012:159) offers the following explanation for why this is so, "Some fricatives produced very far forward in the mouth, like [f], may have no vocal tract filtering at all." In other words, Anyi speakers tend to produce their /f/s forward in the mouth, as shown by the measurements in Table 12A.

	F2 of /f/ before Vowels								
	/fi/	/fɪ/	/fe/	/fɛ/	/fa/	/fɔ/	/fo/	/fʊ/	/fu/
Speaker 1M	1844	1644	1852	1756	1602	1715	1755	1708	1679
Speaker 2M	2122	2187	2130	2107	1951	1953	1927	2000	1881
Speaker 3M	2029	2065	2034	2039	1934	1963	2078	2130	1936
Speaker 4M	2038	2026	2168	2086	2032	2059	2182	1988	1990
Speaker 5M	2373	2392	2359	2224	2063	2019	2069	1991	1915
Speaker 6M	2231	2134	2136	2177	2183	2139	2180	2008	2075
Speaker 7M	2385	2164	2012	2023	2052	2094	1969	1881	2000
Speaker 8M	2103	2119	2114	2164	2001	1950	1967	2062	2082
AVG.	**2140**	**2091**	**2100**	**2072**	**1977**	**1986**	**2015**	**1971**	**1944**

[6] It is important to note that generally the formants in female speech are 20% higher than those produced by males.

St. Dev.	183	211	145	145	169	129	142	127	129

Table 12A: F2 Measurements of [f]

	F2 of /s/ before Vowels								
	/si/	/sɪ/	/se/	/sɛ/	/sa/	/sɔ/	/so/	/sʊ/	/su/
Speaker 1M	1990	1981	1917	1940	1894	2025	1958	2185	2366
Speaker 2M	2807	2800	2339	2461	2373	2684	2455	2478	2272
Speaker 3M	2067	1745	1956	1893	1986	1808	1819	2285	1874
Speaker 4M	1954	1937	1930	1887	1863	2018	2088	2165	2184
Speaker 5M	2100	1973	1991	1971	1968	2206	2301	2212	2244
Speaker 6M	2268	2248	2260	2261	2026	2612	2552	2596	2559
Speaker 7M	1853	1933	1891	1895	1880	2293	2407	2401	2517
Speaker 8M	2190	2062	1917	1918	1894	2298	2294	2457	2690
Speaker 9M	2138	2402	2467	2398	2299	2673	2700	2557	2632
AVG.	**2151**	**2120**	**2074**	**2069**	**2020**	**2290**	**2286**	**2370**	**2370**
St. Dev.	275	318	218	235	187	314	285	163	258

Table 12B: F2 Measurements of [s]

The question that we are interested in answering in this section is whether or not the frication noise of /f/ and /s/ changes in relation to the feature [±ATR] of the vowels that they precede. Pairwise comparisons between the F2 differences of [+ATR] and [-ATR] show that since the difference of 145 Hz is below the JND of 200 Hz required for auditory discrimination, F2 is not a robust correlate when /f/ and /s/ precede a vowel. In order words, the areas of maximal constriction are in the front of the mouth. Both /f/ and /s/ are [+anterior].

8.2 The Intelligibility of F2 Frication Noises

Regardless of vowel context, the arithmetic mean of /f/ is 2032 Hz, and that of /s/ is 2194 Hz. The difference of 162 Hz is below the JND of 200 Hz. This means that by listening to the frication noises of these two segments, one cannot auditorily perceive any difference between them.

	F2 Properties /f/ and /s/ before Vowels									
	/i/	/ɪ/	/e/	/ɛ/	/a/	/ɔ/	/o/	/ʊ/	/u/	**AVG**
AVG./f/s	2140	2091	2100	2072	1977	1986	2015	1971	1944	2032 Hz
AVG. /s/s	2151	2120	2074	2069	2020	2290	2286	2370	2370	2194 Hz
Difference	**11**	**29**	**26**	**3**	**43**	**304**	**71**	**399**	**426**	145 Hz

Table 12C: F2 Comparison between [f] and [s]

As noted in 8.1, Zsiga (2013:140) notes that an F2 of 1600-2000 Hz corresponds to an alveolar place of articulation. The fact that the averaged measurements for both /f/ and /s/ are higher than 2000 Hz means that they are produced further in the front of the mouth than the alveolar ridge. When the F2 of /f/ and /s/ are compared with those found in English, Polish, and Swedish, we see that the measurements in Anyi are higher than those in these three languages. Furthermore, the mean measurement of /s/ (2194 Hz) in Anyi is analogous to the mean F2 of [s^j] (2037 Hz) in Russian produced by females (Kochetov 2017:345, Table A8). These measurements indicate that in producing /f/ and /s/, the upper teeth are involved somehow. For /f/, the lower lips touch the upper teeth. For /s/, the tip of the tongue moves up to or right below the upper teeth. So, /f/ can be classified as a labiodental fricative and /s/ an apical dental fricative. The latter classification is in keeping with Ladefoged and Maddieson (1996:164, Table 5-7).

9.0 F3 Measurements

F3 correlates articulatorily with the movements of the lips. When the lips are rounded, F3 values decrease. But when they are unrounded, F3 values increase. Generally, 2500 Hz is taken as a threshold of reference. Measurements that are ≥ 2500 Hz are indicative of unrounded lips, whereas those that are ≤ 2500 Hz denote rounded lips. Stevens' (2000:401) modeling of /s/ shows that its typical F3 value is 2750 Hz. His F3 measurement for /f/ is below about 2500 Hz (p.390). Klatt (1980:987, Table III) synthesized /f/ and /s/ with F3 measurements of 2080 Hz and 2430 Hz. Natural data obtained by Jassem (1962:1962:16, Tables 1a-1c) shows that the F3 of /f/ in American English, Polish, and Swedish are respectively 2650 Hz, 2900 Hz, 2450 Hz. As for /s/, the measurements are 2620 Hz, 2920 Hz, and 2650 Hz. The JND threshold for assessing the robustness of segments on the F3 bandwidth is stated as follows,

Auditory Discrimination on the F3 Frequency Bandwidth

Of two speech signals **A** and **B**, **A** is perceived as auditorily distinct from **B** if and only if there is a difference of 400 Hz or more between them.

When this JND is used, we see that lip movements are not robust correlates for /f/ and /s/ because in the measurements listed in Tables 13A, 13B, and 13C, the difference of 216 Hz between /f/ and /s/ is below the JND.

9.1 Possible Correlation between F3 and [±ATR]

Pairwise comparisons of F3 values for [±ATR] show that F3 is not a robust correlate, that is, the values of the frication noises of /f/ and /s/ do not vary by much whether they precede [+ATR] or [-ATR] vowels. The measurements also indicate that the lips are unrounded when /f/ and /s/ are produced.

	F3 of /f/ before Vowels								
	/fi/	/fɪ/	/fe/	/fɛ/	/fa/	/fɔ/	/fo/	/fʊ/	/fu/
Speaker 1M	2793	2586	2653	2435	2538	2705	2655	2706	2669
Speaker 2M	2807	2983	2796	2928	2789	2724	2682	2799	2573
Speaker 3M	3017	3073	3025	2991	2876	2761	3028	3076	2899
Speaker 4M	2759	2781	2943	2672	2818	2791	3047	2841	2870
Speaker 5M	2982	2955	2891	2827	2786	2685	2718	2653	2741
Speaker 6M	3079	2990	3028	3038	3052	2917	2918	2772	2723
Speaker 7M	3183	3014	2882	2736	2880	2861	2853	2745	2724
Speaker 8M	2896	2925	2945	3020	2783	2763	2790	2814	2789
AVG.	**2939**	**2913**	**2895**	**2830**	**2815**	**2775**	**2836**	**2800**	**2748**
St. Dev.	151	157	124	208	143	79	151	126	105

Table 13A: F3 Measurements of [f]

	F3 of /s/ before Vowels								
	/si/	/sɪ/	/se/	/sɛ/	/sa/	/sɔ/	/so/	/sʊ/	/su/
Speaker 1M	3128	2898	3010	3027	2997	3319	3253	3251	3160
Speaker 2M	3735	3656	3283	3360	3360	3136	3068	2927	3054
Speaker 3M	2979	2706	2937	2889	3109	3386	3280	3262	3073
Speaker 4M	2740	2465	2694	2657	2801	2923	3016	3081	3027
Speaker 5M	3002	2855	2898	2930	3004	3068	3053	2915	2971
Speaker 6M	3157	3142	3208	3173	3321	3146	3055	3090	3060
Speaker 7M	2707	2707	2677	2754	2549	3060	3181	2909	2955
Speaker 8M	3128	3148	3010	3007	2997	3393	3306	3058	3088
Speaker 9M	2836	3317	3418	3282	3278	2968	2987	3022	3074
AVG.	**3045**	**2988**	**3015**	**3008**	**3046**	**3155**	**3133**	**3057**	**3051**
St. Dev.	308	364	252	233	261	174	122	133	61

Table 13B: F3 Measurements of [s]

9.2 Intelligibility F3 Frication Noises

The mean F3s of the frication noises of /f/ and /s/ are respectively 2839 Hz and 3055 Hz. With a difference of only 216 Hz, which is below the JND of 400 Hz, we conclude that F3 does not play any role in the intelligibility of these two fricatives.

	F3 Properties of /f/ and /s/ before Vowels									
	/i/	/ɪ/	/e/	/ɛ/	/a/	/ɔ/	/o/	/ʊ/	/u/	**AVG**
AVG. /f/s	2939	2913	2895	2830	2815	2775	2836	2800	2748	2839 Hz
AVG./s/s	3045	2988	3015	3008	3046	3155	3133	3057	3051	3055 Hz
Difference	106	75	120	178	231	380	297	257	303	216 Hz

Table 13C: F3 Comparison between [f] and [s]

The observation in 8.2 that /f/ is a labiodental fricative and /s/ is an apical fricative is confirmed by F3 measurements. The F3 values in Anyi are higher than those of the languages reviewed in the previous section. The Anyi values are in line with the measurements for Polish /f/ (2900 Hz) and /s/ (2920 Hz). When Anyi talkers produce /f/ and /s/, the lips are unrounded because the F3 values are clearly ≥ 2500 Hz.

10.0 F4 Measurements

Rare are the studies that provide F4 measurements. Those that do, like Jassem (1962), provide measurements that are based on only a single speaker's utterances. Even so, Stevens' (2000:407, 410) simulations predict that the F4s of /f/ and /s/ would be 3200 Hz and 3500 Hz respectively.[7] Jassem (1962:16, Tables 1a-1c) lists an F4 of 3850 Hz for American English. For Polish and Swedish, the F4s are respectively 4030 Hz and 3630 Hz. For the F4 of /s/, we have

[7] Most simulations are based on the following assumptions: 1) the vocal track length is 17 cm, 2) that women's formants are 15-20% higher than males, 3) lower formants, that is, F1, F2, and F3 vary with pronunciation, 4) higher formants, F4 and F5 do not vary by much. Furthermore, they contribute little to segmental intelligibility (Klatt 1980:979-80).

3950 Hz for American English, 4170 Hz for Polish, and 3820 Hz for Swedish. The JND for evaluating the robustness of F4 cues is stated as follows:

Auditory Discrimination on the F4 Frequency Bandwidth

Of two speech signals **A** and **B**, **A** is perceived as auditorily distinct from **B** if and only if there is a difference of 600 Hz or more between them.

There is no clear articulatory correspondence for F4. Ladefoged and Johnson (2015:222) contend that it correlates with the size of a person's head. However, there is no consensus on this opinion. When the frication noises of /f/ and /s/ in American English, Polish, and Swedish are compared, we see that F4 is not a robust cue because the differences are respectively only 100 Hz, 140 Hz, and 190 Hz. Stevens' simulation also yields a difference of 300 Hz, which means that F4 is not a robust cue, because in all these cases the differences are below 600 Hz.

10.1 Possible Correlation between F4 and [±ATR]

The frication noises of /f/ and /s/ are all far below the JND of 600 Hz. We conclude that F4 does not differentiate between /f/ and /s/.

	F4 of /f/ before Vowels								
	/fi/	/fɪ/	/fe/	/fɛ/	/fa/	/fɔ/	/fo/	/fʊ/	/fu/
Speaker 1M	3746	3579	3578	3355	3451	3780	3716	3784	3756
Speaker 2M	3795	3949	3768	3930	3489	3493	3450	3731	3356
Speaker 3M	3982	3951	4005	3964	3850	3228	3892	3959	3806
Speaker 4M	3789	3814	3787	3713	3676	3820	3848	3807	3788
Speaker 5M	3737	3833	3687	3641	3618	3611	3579	3619	3588
Speaker 6M	4356	4368	4370	4363	4460	4347	4393	4187	4206
Speaker 7M	3774	3753	3685	3620	3667	3637	3537	3414	3439
Speaker 8M	3814	3830	3852	4002	3646	3491	3594	3617	3637
AVG.	**3874**	**3884**	**3841**	**3823**	**3732**	**3675**	**3751**	**3764**	**3697**
St. Dev.	209	227	248	307	318	328	300	234	262

Table 14A: F4 Measurements of [f]

	F4 of /s/ before Vowels								
	/si/	/sɪ/	/se/	/sɛ/	/sa/	/sɔ/	/so/	/sʊ/	/su/
Speaker 1M	3896	3632	3831	3858	3843	3690	3594	3651	3713
Speaker 2M	4187	4331	4072	4110	4203	3771	3730	3704	3718
Speaker 3M	3988	3788	3923	3911	4106	3882	3735	3879	3508
Speaker 4M	3710	3557	3714	3596	3795	3713	3727	3694	3722
Speaker 5M	3760	3659	3717	3611	3625	3480	3463	3432	3528
Speaker 6M	3972	3974	4009	3967	4106	3638	3683	3636	3602
Speaker 7M	3705	3597	3697	3689	3577	3536	3437	3481	3530
Speaker 8M	3913	3930	3831	3851	3843	4381	3767	3703	3832
Speaker 9M	3775	4111	4261	3990	4042	3656	3774	3846	3937
AVG.	**3878**	**3842**	**3895**	**3842**	**3904**	**3749**	**3656**	**3669**	**3676**
St. Dev.	158	264	190	177	221	264	128	146	147

Table 14B: F4 Measurements of [f]

The F4 of /f/ and /s/ measurements in Anyi resemble more closely those in Swedish than those in American English or Polish. We note in passing that when F4 measurements are extracted from talkers, they are considerably higher than those obtained by Stevens (2000) through simulation.

10.2 Intelligibility F4 Frication Noises

Just by listening to the frication noise alone, one cannot determine whether the fricative is /f/ or /s/ because the difference of 64 Hz between them is far below the JND of 600 Hz required for auditory intelligibility. The measurements in Table 14C confirm that F4 is not a robust correlate at all.

	F4 Comparison between /f/ and /s/ before Vowels									
	/i/	/ɪ/	/e/	/ɛ/	/a/	/ɔ/	/o/	/ʊ/	/u/	**AVG**
AVG. /f/s	3874	3884	3841	3823	3732	3675	3751	3764	3697	3782 Hz
AVG. /s/s	3878	3842	3895	3842	3904	3749	3656	3669	3676	3790 Hz
Difference	**4**	**42**	**54**	**19**	**172**	**74**	**95**	**95**	**21**	64 Hz

Table 14C: F4 Comparison between [f] and [s]

11.0 Bandwidth Measurements

There are two ways of extracting bandwidth values. They can be extracted at the same time when F1, F2, F3, and F4 formants measurements are being extracted. Alternatively, bandwidth measurements (B1, B2, B3, and B4) can be extrapolated from their corresponding formant values. Klatt (1980:980) writes that "Bandwidths are difficult to deduce from the analysis of natural speech because of irregularities in the glottal source spectrum." In other words, it is better to estimate bandwidth measurements than to extract them. However, there is a problem. Different algorithms have been proposed for estimating bandwidths, but they differ considerably. The one that I rely on the most is Rabiner and Juang (1993:152) because of its simplicity. They propose that bandwidths can be estimated by taking 20-40% of the formant value of a segment. Klatt (1980:981) proposes 5% or 10% formant values, depending on the specific segment under consideration (Klatt 1980:879). Yet, on page 987, Table III, his estimation of B1, B2, and B3 of the bandwidths of /f/ and /s/ are approximately 40% of F1, 35% of F2, and 44% of F3. For the calculations below, a uniform of 20% of formant values is applied for calculating the bandwidths of /f/ and /s/.

	Bandwidths of /f/									
	/fi/	/fɪ/	/fe/	/fɛ/	/fa/	/fɔ/	/fo/	/fʊ/	/fu/	**AVG**
B1	94	100	114	122	139	126	107	101	91	**110 Hz**
B2	428	418	420	414	395	397	403	394	388	**406 Hz**
B3	587	582	579	566	563	555	567	560	549	**567 Hz**
B4	774	776	768	764	746	735	750	752	739	**756 Hz**

Table 15A: Bandwidths of [f]

	Bandwidths of /s/									
	/si/	/sɪ/	/se/	/sɛ/	/sa/	/sɔ/	/so/	/sʊ/	/su/	**AVG**
B1	84	92	101	108	130	112	105	96	85	**101 Hz**
B2	430	424	414	413	404	450	457	474	474	**437 Hz**
B3	609	597	603	601	609	631	626	611	610	**610 Hz**
B4	775	768	779	768	780	749	731	733	735	**757 Hz**

Table 15B: Bandwidths of [s]

Because no bandwidth measurements are available, I do not know how these values compare with other African languages. We know that Klatt (1980:987, Table III) synthesized /f/ with a B1 value of 200 Hz, a B2 of 120 Hz, and a B3 of 150 Hz. He synthesized /s/ with a B1value of 200 Hz, B2 of 80 Hz, and a B3 of 200 Hz.

12.0 Summary

The investigations of /f/ and /s/ have been comprehensive because there are many research questions to which answers are needed. First, we wanted to know if the frication characteristics /f/ and /s/ change in accordance with the feature [±ATR] of the vowels that they precede. The analyses were carried out across F0/pitch, CoG, Intensity, Duration F1, F2, F3, and F4. The most important findings are summarized in the two tables below. The symbol "+" means that the correlate is robust, and "–" means that it is not.

	[±ATR]	**Observations**
F0	+	Robust
Intensity	–	Not robust
Duration	+	Robust
CoG	–	Not robust
F1	+	Robust only for [i] vs. [ɪ]
F2	–	Not robust
F3	–	Not robust
F4	–	Not robust

Table 16: Summary of Vowel Context

The most important take-away is that the feature [±ATR] contributes significantly to the duration of the frication noise of /f/ and /s/. They are considerably longer when they precede [+ATR] vowels than [-ATR] ones, except for the central vowel /a/. Jongman (1998:1724) reports that "The effects of vowel context on fricative identification are somewhat puzzling." This is so because some vowels aid in the identification of some fricatives but not others. For American English, he found that "Fricative identification tends to be better in the context of [i, u] than in the context of [a]." F1 plays some role in Anyi, but this is confined only to the high vowels [i] and [ɪ]. This speaks to the puzzling effects that vowels have on the frication noises of fricatives.

Table 17 answers the question as to whether or not, various acoustic correlates contribute to the intelligibility of the frication noise of /f/ and /s/ irrespective of vowel contexts:

	Intelligibility of Frication Noise	Observations
F0	+	Robust
Intensity	+	Marginal
Duration	+	Robust
CoG	+	Robust
F1	–	Not robust
F2	–	Not robust
F3	–	Not robust
F4	–	Not robust

Table 17: Significance of Correlates of Frication Noise

Finally, numerous studies have highlighted important interspeaker and even intraspeaker variability in the production of fricatives. When studies report only arithmetic means, variability is less obvious. However, when individual speakers' data are made available, variability stands out. For example, Johnson (2012:163) observes that "It has also been noted that there may be a substantial range of inter-speaker variability in the frequencies of the spectral peaks in fricatives. These observations have led to the development of center-of-gravity techniques for the characterization of fricative spectra." The CoG measurements show that the fricatives of Speakers 1M, 3M, and 4M are significantly different from those produced by other participants. However, this observation is not limited only to CoG. It is visible across many of the correlates extracted for this paper.

ABOUT THE AUTHOR

Ettien Koffi, Ph.D. linguistics (Indiana University, Bloomington, IN) teaches at Saint Cloud State University, MN. He is the author of five books and author/co-author of several dozen articles on acoustic phonetics, phonology, language planning and policy, emergent orthographies, syntax, and translation. His acoustic phonetic research is synergetic, encompassing L2 acoustic phonetics of English (Speech Intelligibility from the perspective of the Critical Band Theory), sociophonetics of Central Minnesota English, general acoustic phonetics of Anyi (a West African language), acoustic phonetic feature extraction for application in Automatic Speech Recognition (ASR), Text-to-Speech (TTS), voice biometrics for speaker verification, and infant cry bioacoustics. Since 2012, his high impact acoustic phonetic publications have been downloaded **68,201** times (**47,355** as per Digital Commons analytics), (**20,846** as per Researchgate.net analytics), and several thousand downloads from Academia.edu, as of **February 2024**. He can be reached at enkoffi@stcloudstate.edu.

References

Boersma, Paul & Weenink, David. 2024. Praat: doing phonetics by computer [Computer Program]. Version 6.4.05, retrieved February 8, 2024 from http://www.praat.org/

Fromkin, Victoria., Robert Rodman., and Nina Hyams. 2017. *An Introduction to Language*. 11th edition. Boston, MA: Cengage Learning.

Fry, Dennis B. 1979. *The Physics of Speech*. New York: Cambridge University Press.

Gordon, Matthew., Barthmaier, P., and Sands, K. 2000. A Cross-linguistic Acoustic Study of

Voiceless Fricatives. *Journal of the International Phonetic Association* 32 (2):114-174.
Halle, Maurice, G.W. Hughes, and J.P. A. Radley. 1957. Acoustic Properties of Stop Consonants. *Journal of the Acoustic Society of America* 29 (1):107-116.
Jassem W. 1962. The Formant Patterns of Fricative Consonants. *Quarterly Progress and Status Report* 3:3, pp. 6-15.
Johnson, Keith. 2012. *Acoustic and Auditory Phonetics.* Third Edition. Malden, MA: Wiley-Blackwell.
Jongman, Allard and Ratree Wayland. 2000. Acoustic Characteristics of English Fricatives. *Journal of the Acoustical Society of America* 108 (3):1252-1263.
Jongman, Allard, Ratree Wayland, and Serena Wong. 1998. Acoustic Characteristics of English Fricatives: I. State Cues. *Working Papers of Cornell Phonetics Laboratory*, Volume 12:195-205.
Kent, Ray D. and Charles Read. 2002. *Acoustic Analysis of Speech*. Second Edition. Clifton Park, NY: Delmar Cengage Learning.
Klatt, Dennis H. 1980. Software for a Cascade/Parallel Formant Synthesizer. *Journal of the Acoustical Society of America 68 (3):971-995.*
Kochetov, Alexei. 2017. Acoustics of Russian Voiceless Sibilant Fricatives. *Journal of the International Phonetic Association* 47 (3): 321-348. doi:10.1017/S0025100317000019.
Koffi, Ettien. 2023a. Voice Biometrics Fusion for Enhanced Security and Speaker Verification: A Comprehensive Review. *Linguistic Portfolios* 12: 72-97.
Koffi, Ettien. 2023b. The Acoustic Phonetic Properties of Singly and Doubly Articulated Stops in Anyi: A Data-Driven Analysis. *Linguistic Portfolios* 12: 123-147.
Koffi, Ettien. 2021. *Relevant Acoustic Phonetics of L2 English: Focus on Intelligibility*. Boca Raton: FL: CRC Press.
Koffi, Ettien and Casey Lundy. 2017. An Acoustic Phonetic Account of the Production of Word-Final /z/s in Central Minnesota English. *Linguistic Portfolios* 6:109-124.
Ladefoged, Peter. 1968. *A Phonetic Study of West African Languages: An Auditory-Instrumental Survey.* Cambridge, UK: Cambridge University Press.
Ladefoged, Peter and Keith Johnson. 2015. *A Course in Phonetics.* Seventh Edition. Malden, MA: Cengage Learning.
Ladefoged, Peter and Ian Maddieson. 1996. *The Sounds of the World's Languages*. Malden, MA: Blackwell Publishers.
Lokousa, Hannah. 2023. Praat Script: Extra Sound Correlates. Praat, Version 6.1.38.
Maddieson, Ian. 1984. *Patterns of Sounds.* New York: Cambridge University Press.
Oxford Advanced Learner's Dictionary 2000. New York: Oxford University Press.
Pirello, Karen, Sheilla E. Blumstein, and Kathleen Kurowski. 1997. The Characteristics of Voicing in Syllable-initial Fricatives in American English. *Journal of the Acoustical Society of America* 101 (6): 3754-3765.
Pope, J. 1998. Analyzers. In *Handbook of Acoustics*, ed. by Malcolm J. Crocker, pp. 1341-1353. New York: A Wiley-Interscience Publication.
Rabiner, Lawrence and Biing-Hwang Juang. 1993. *Fundamentals of Speech Recognition*. Englewood Cliffs, New Jersey: Prentice Hall.
Shadle, Christine H. 1985. Acoustics of Fricative Consonants. *Technical Report* 506. Massachusetts Institute of Technology.
Stevens, Kenneth N. 2000. *Acoustic Phonetics.* The MIT Press: Cambridge, Massachusetts.

Stevens, Kenneth. 1992. Acoustic and Perceptual Characteristics of Voicing in Fricatives and Fricative Clusters. *Journal of the Acoustical Society of America* 91 (5):2979-3000.
Zsiga, Elizabeth C. 2013. *The Sounds of Language*. Wiley-Blackwell: Malden, MA.

Appendix of /f/ tokens produced by Speaker 9F

Kochetov's (2017:345, Table A8) study shows that variability of fricatives is obvious between males and females and within each gender group. Variability is also noticeable across various types of vowels. The measurements reported in the body of the paper are based on the pronunciation of nine Anyi males. Tables 18A and 18B display the pronunciation patterns of the lone female participant. Her data was not mixed in with males because of the greater gender difference between males and females in the production of fricatives. She produced 144 tokens (2 fricatives x 9 vowels x 8 correlates).

	F0	**F1**	**F2**	**F3**	**F4**	**Dur**	**Ints**	**CoG**
/f/ before /i/	208	583	2171	2916	3699	187	64	3337
/f/ before /ɪ/	235	513	1972	2764	3555	144	64	7300
/f/ before /e/	106	683	1887	2734	3474	173	63	5279
/f/ before /ɛ/	180	729	2124	2982	3734	184	71	2856
/f/ before /a/	141	724	1987	2779	3504	127	65	6570
/f/ before /ɔ/	145	565	1643	2393	3193	89	77	1286
/f/ before /o/	193	590	1846	2723	3475	90	66	5157
/f/ before /ʊ/	133	599	1904	2681	3515	130	64	6134
/f/ before /u/	206	574	1842	2747	3513	152	63	6276
AVG. Female	**171**	**617**	**1719**	**2746**	**3518**	**141**	**66**	**4910**

Table 18A: Female Pronunciations of /f/

	F0	**F1**	**F2**	**F3**	**F4**	**Dur**	**Ints**	**CoG**
/s/ before /i/	187	452	2189	3036	3863	164	72	7918
/s/ before /ɪ/	150	441	2223	3088	4055	175	70	9990
/s/ before /e/*[8]	74	535	2044	2929	4336	190	80	591*
/s/ before /ɛ/	144	546	2330	3116	3733	193	71	7927
/s/ before /a/	74	617	2039	2897	3791	191	68	10235
/s/ before /ɔ/	211	536	2086	3225	4068	198	74	7611
/s/ before /o/	171	549	2246	3326	4076	192	74	6996
/s/ before /ʊ/	154	537	2450	3320	2802	200	75	7280
/s/ before /u/	213	550	2556	3474	4129	212	76	7290
AVG. of /s/s	**153**	**529**	**2240**	**3156**	**3872**	**190**	**73**	**7315**

Table 18A: Female Pronunciations of /s/

Summary Measurements

	F0	**F1**	**F2**	**F3**	**F4**	**Dur**	**Ints**	**CoG**
AVG. of /f/s	171	617	1719	2746	3518	141	66	4910
AVG. of /s/s	153	529	2240	3156	3872	190	73	7315
Difference	**18**	**88**	**521**	**410**	**354**	**49**	**7**	**2405**

Table 18C: Female Comparative Pronunciations of /f/ and /s/

[8] The CoG measurement for /e/ is clearly an outlier. It was not included in the overall report for this speaker.