Linguistic Portfolios

Volume 9 Article 2

2020

A Comprehensive Review of Intensity and Its Linguistic Applications

Ettien Koffi
St. Cloud State University, enkoffi@stcloudstate.edu

Recommended Citation
Koffi, Ettien (2020) "A Comprehensive Review of Intensity and Its Linguistic Applications," *Linguistic Portfolios*: Vol. 9 , Article 2.
Available at: https://repository.stcloudstate.edu/stcloud_ling/vol9/iss1/2

A COMPREHENSIVE REVIEW OF INTENSITY AND ITS LINGUISTIC APPLICATIONS

ETTIEN KOFFI

ABSTRACT

Every speech sound that human beings produce or perceive is a composite of three independent, yet interdependent correlates, i.e., F0/pitch, intensity, and duration. Koffi (2019) provided a comprehensive review of F0 and its various correlations. In this paper, the focus is on intensity and its linguistic applications. Various terms related to intensity are first examined and explained. Thereafter, its psychoacoustic properties are highlighted and discussed from the standpoint of the Critical Band Theory (CBT) and Just Noticeable Difference (JND) thresholds. Intensity measurements from 55 speakers of Central Minnesota English (CMNE) are used to illustrate the aforementioned concepts. Later in the paper, Equation 4 is used to derive sonority indices that are firmly grounded in the physics of speech. It is argued that these indices are better suited to account for the Sonority Sequencing Principle (SSP) and the Minimal Sonority Distance Parameter (MSDP) than the arbitrary sonority indices commonly used in phonological analyses.

Keywords and phrases: Intensity, Intensity thresholds, Intensity SPL, Intensity weighting, Sonority Sequencing Principle, Minimal Sonority Distance Parameter, JND, Critical Band Theory, Sonority Index, Sound Power, Acoustic Comfort

1.0 Introduction

Linguists have, for the most part, not paid sufficient attention to intensity. It has been neglected and overlooked because it is not phonemic in any human language. In other words, producing a speech sound with various degrees of loudness does not change the meaning of words. Since linguists prioritize meaning (phonemic, morphological, syntactic, semantic) over anything else, it is not at all surprising that they would deem intensity not worth pursuing. Yet, I argue in this paper that intensity is crucial for formulating a coherent theory of syllable phonotactics. The goal of this paper is to highlight the relevance of intensity in linguistic analyses. This is indeed a tall order because familiarity with physics, an academic discipline with which most linguists are unfamiliar, is a prerequisite. It therefore goes without saying that the first installment of the paper must discuss the physical bases of intensity. The second portion must also deal with the psychoacoustics of intensity because, if it is overlooked, intensity measurements are likely to be misinterpreted. The third section of the paper uses intensity data collected from 55 speakers of Central Minnesota English (CMNE) to illustrate the aforementioned concepts. Then and only then do we discuss the notion of sonority index, which constitutes the foundation on which the Sonority Sequencing Principle (SSP) and the Minimal Sonority Distance Parameter (MSDP) are built. The final section of the paper introduces the concept of sound power and assesses its usefulness and applicability to phonological analysis.

1.1 Clarification of Terms

The applicability of intensity to linguistic analyses is obscured by statements such as these:

> Despite what you may have read elsewhere (and this is why I have included this section), intensity as shown in dB is usually not a very useful acoustic property to measure. It is seldom one of the distinguishing characteristics of a language (Ladefoged 2003:93).

or

> In most recordings made for phonetic research, absolute intensity measures given by Praat are largely meaningless (Styler 2017:24).

While there is an element of truth in these two statements, they also contain hyperboles. In order to sort out which is which, terminological clarification is needed. In the non-technical literature, amplitude and loudness are often used as true synonyms of intensity but they are not. To help my students understand the differences between these terms, I give them these analogies: *"Intensity is to amplitude what water is to ice,"* or *"Intensity is to loudness what water is to steam."* Let me explain what I mean by these two statements.

Strictly speaking, intensity is not the same as amplitude. Just like ice = water + freezing temperature, **amplitude = intensity + distance**. Intensity is the basic or indispensable element in amplitude, but distance cannot be ignored. If distance is omitted, the amplitude of a sound is uninterpretable. In many studies in which amplitude is used, sooner or later the author will mention a unit of distance either in the methodology section or somewhere else in the document. The standard distances in many linguistic analyses are **30 cm** (when intensity is measured from the mouth of the speaker speaking into a microphone) or **1 m** when a microphone is not in use (Fletcher 1953:76, 79, 83).

If loudness is used as a synonym of intensity, a vital element needed for interpretation is "psychological response." Just as one cannot have steam without having heat, one cannot have loudness without mentioning the hearer's psychological response to the stimulus. Therefore, **loudness = intensity + psychological response**. Fletcher (1953:176-7) explains the relationship between intensity and loudness as follows:

> Loudness is a psychological term used to describe the magnitude of an auditory sensation. Although we use the terms 'very loud,' 'loud,' 'moderately loud,' 'soft,' and 'very soft' … to define the magnitude, it is evident that these terms are not at all precise and depend upon experience, the auditory acuity, and the customs of the persons using them. If loudness depended only upon the intensity of the sound wave producing the loudness, then measurements of the physical intensity would definitely determine the loudness as sensed by a typical individual and therefore could be used as a precise means of defining it. However, no such simple relation exists.

2.0 Intensity in Physics

Amplitude and loudness are not the only terms used as synonyms of intensity. In the physics literature, we come across terms such as "sound pressure level" (SPL), "sound power," "absolute vs. relative intensity," "weighted vs. unweighted intensity," "sound vs. noise," "perceptual scale vs. logarithmic scale." These terms also need to be explained because, strictly speaking, they are not direct synonyms of intensity. However, all the terminological differences will not be explained right in this section, but in subsequent sections when most relevant. Instead, let's focus here on the units of measurement of intensity. The first is the **Bell**, named after Alexander Graham Bell, a towering figure in physics. The human ear can perceive an extraordinarily huge range of variation in intensity, from the tiniest pin drop to the loudest clap of thunder. This range is expressed arithmetically as follows: 10 x 10 x 10 x 10 x 10 x10 or 10^6. Appendix A in Reetz and Jongman (2009:290) contains useful conversion tables to this effect. The same can also be expressed linearly as 1,000,000th degrees of variation. Instead of expressing this range arithmetically, a logarithmic scale is used. Here is the rationale:

> Logarithms are particularly useful to audio engineers because they can correlate measurements to human hearing, and they also allow large ranges of numbers to be expressed efficiently. Logarithms are the foundation for expressing sound levels in decibels where the sound level is a ratio. In particular, a sound level in decibels is 10 times the logarithm to the base 10 of the ratio of two power-like quantities (Everest and Pohlmann 2015:21).

The logarithmic scale makes it possible to express intensity as **Equation 1**:

$$1 \text{ Bell} = \log_{10} = \frac{\text{Absolute Intensity}^{1}}{\text{Reference Intensity}}$$

To the left of $\log_{10}$, we see 1 Bell. The Bell is such a large unit that it has been subdivided into a decimal scale by units of tenths. This way, 1 Bell = 10 decibels. The decibel is abbreviated as **dB**. If the decibel is used, then Equation 1 can be modified as **Equation 2**:

$$1 \text{ dB} = 10 \log_{10} = \frac{\text{Absolute Intensity}}{\text{Reference Intensity}}$$

The only difference between Equations 1 and 2 is that the latter has 10 as a prefix of $\log_{10}$, that is, 1 dB = 10 $\log_{10}$ plus the rest of the equation.

Everest and Pohlmann (2015:22) quote experts as saying that sound intensity is difficult to measure directly. However, sound pressure is relatively easier to measure. Sound pressure is generally defined as the physical sensation that air molecules make when they hit the tympanic membrane. This pressure is reportedly equivalent to *a mosquito buzzing at a distance of 3 m from a human ear*. The term "**Sound Pressure Level**," commonly abbreviated as **SPL** is used to refer to this sensation. As a rule of thumb, auditory perception of intensity is always expressed in dB

---

[1] Those who wish to explore this equation further should read Hasen, Colin H. 2001.

SPL. However, more often than not, the suffix SPL is omitted. We follow the same tradition of not indicating SPL after intensity measurements. The notion of amplitude is firmly embedded in SPL because the distance of 3 m is factored into it. Other elements are also factored into SPL but the explanation is delayed until after the formulation of Equation 3. Suffice it to say for now that "intensity" and "SPL" are almost synonymous but not quite. The unit of measurement of SPL is the **Pascal**, abbreviated as **"Pa."** It is named after the French mathematical phenom, Blaise Pascal. The threshold of hearing is **20μPa**, where **"μ"** stands for **"micro,"** i.e., one millionth of a Pa. It is the tiniest air pressure that the human ear can detect. When this pressure is converted into a decibel unit, it is equivalent to 0 dB. This leads to **Equation 3**:

$$\text{SPL} = 2\text{x}10 \log_{10} = \frac{\text{Absolute Intensity}}{20\mu\text{Pa}^{2}}$$

The prefix "**2**" before 10 $\log_{10}$ in Equation 3 is explained later in section 5.1. Everest and Pohlmann (2015:22) give more detailed explanations about this equation, but we skip over them for the sake of brevity. The denominator 20μPa refers to **"standard conditions."** What are they? As noted previously, SPL includes the notion of distance. Other elements that are factored in the equation are **speed** (the speed at which air molecules travel, that is, 343 m per second), **temperature** (commonly taken to be 20 degree Celsius) and **barometric pressure** (at 760 mm, Speaks 2005:135-6). The abbreviation **"mm"** stands for "millimeter of mercury." These are some of the considerations that led Jones (2013:151) to make this astounding statement:

> Languages are spoken in different settings around the world, and as speech is primarily an aerodynamic activity, differences in altitude could potentially have long-term impacts on speech patterns.

He provides other details about atmospheric pressures at sea level and how different altitudes affect lung capacity, and so on and so forth. The way he describes them, an unsuspecting linguist would be misled into thinking these conditions affect speech. But according to Fletcher (1953:70), their impact on every day speech is negligible. He notes that "A change in temperature from 0 to 50^{o}C produces a change of only 0.36 dB, and a change in atmospheric pressure from 60 to 76 cm produces only a change of 1 dB."

2.1 Logarithmic Scale and Intensity Weighting

Other terms for which clarification is needed includes the differentiation between the **logarithmic scale** and the **arithmetic/linear scale**. This distinction must be taken seriously because failure to do so can lead to erroneous interpretations of intensity measurements. To start with, let's consider the following example from Breysse and Lees (2006:25-27):

> Since SPLs are based on a log scale, they cannot be added directly, i.e., 80 dB + 80 dB does not equal 160 dB. … Instead, 80 dB + 80 dB = 83 dB; 100 dB + 100 dB = 103 dB; 40

---

[2] Styler (2017:25) notes that "Praat uses two measures of intensity. Pascal tends to be very small numbers (like 0.00033082594541105064) where dB measurements are far larger yielding numbers like 59.23328336655995". The decibel unit is preferred because it is more manageable. More often than not, the numbers following decimal point are simply ignored. In this case, we simply have 59 dB.

> dB + 40 dB = 43 dB. … Adding two sound pressure levels of equal value will always result in a 3 dB increase…. Given four machines producing 100 dB, 91 dB, 90 dB, and 89 dB respectively, their total intensity output is 101 dB.[3]

This example underscores the fact that intensity cannot be interpreted linearly (i.e., arithmetically). Doing so leads to grave errors. For example, we will see in section 4.2 that the intensity of the vowel [i] produced by 33 women is 65 dB, while the intensity of [ɪ] is 67 dB. An unsuspecting linguist would erroneously claim that [ɪ] is louder than [i] simply because there is a 2 dB difference between them. However, this is not true because the human ear perceives both intensity levels identically. Again, intensity is not perceived arithmetically, but logarithmically. This leads to the next important concept in intensity analysis, i.e., **intensity weighting**. Four scales are in use and are described briefly below as follows:

1. **dB** and **dB(Z) weighting**: Here "Z" stands for "zero." This scale measures only absolute intensity values. Praat renders intensity measurements in dB.
2. **dB(A) weighting**: This is the default setting in many sound level meters (SLM) and apps. When this scale is used, it measures intensity as the human ear perceives it. The dB(A) is also sometimes written as dBA.
3. **dB(B) weighting**: This scale is no longer in common use. It was initially designed for measuring mid-range intensities between dB(A) and dB(C). Now, it is seen as redundant. Therefore, it has been discontinued.
4. **dB(C) weighting**: It is used to measure the intensity levels produced by heavy machinery, airplanes, rocket launchers, and sonic booms. Loubeau and Page (2018:23-30) have a very interesting article about this in *Acoustics Today*.

3.0 Intensity in Psychoacoustics

Styler's (2017:24) statement that "In most recordings made for phonetic research, absolute intensity measures given by Praat are largely meaningless" is true only if one is unware of psychoacoustics. But for seasoned phoneticians, the intensity measurements provided by Praat are not hard to interpret, thanks to the impressive work done by researchers at Bell Laboratories some 60 years ago. See Yost (2015) for an excellent short summary of the breakthroughs. Since it is impossible to talk about psychoacoustics without taking a look at the human ear, let's highlight key parts of Figure 1 that play an important role in the perception of speech signals.

---

[3] The website http://hearinglosshelp.com/blog/converting-decibels-to-sound-intensities/ contains helpful information in converting decibel levels into sound intensities. It offers an easy conversion table. The website http://www.sengpielaudio.com/calculatorSonephon.htm also helps convert sones and phons into dB(A) values, and vice versa. They were both retrieved on July 9, 2019.

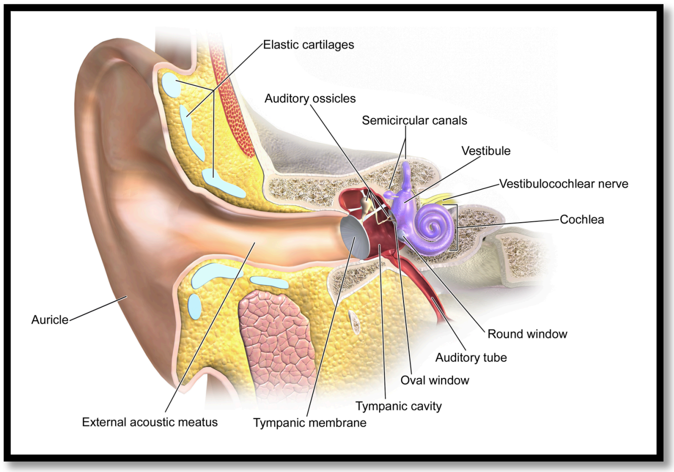

Figure 1: The Anatomy of the Ear
Blausen 0328. WikiJournal of Medicine. Wikimedia Commons

We begin with the visible part of the ear and move at brisk pace to the **middle ear**, and then to the **inner ear**. The **pinna** is the visible part of the ear. It gathers air molecules and funnels them into the ear canal. Once there, air molecules impinge on the tympanic membrane (eardrums), which acts as the boundary between the outer ear and the middle ear. The air molecules that hit the tympanic membrane set the **malleus, the incus, and the stapes/stirrup** (the three tiniest bones in the human body) inside of the middle ear into motion, which causes the signals to be amplified. The amplified signals travel through the inner ear and end up in the **cochlea** where the **basilar membrane** processes the arriving speech signals. The cochlea has two fitting nicknames: one as "an engineering marvel" and the other as "a frequency analyzer." The latter is further described in Figure 2 below.

3.1 Critical Band Theory and Relative Intensity Thresholds

A major breakthrough in how humans perceive sounds came as a result of physicist Harvey Fletcher's work. In 1940, he wrote an influential paper in which he posited mathematically that the basilar membrane can be subdivided into some 24 to 30 critical bands.[4] The length of the basilar membrane is approximately 3.5 cm. Each critical band is 1.3 mm in length and is sensitive to a specific band of frequencies. The whole length of the basilar membrane is covered by some 12,000 outer hair cells and some 3,000 inner hair cells. Lewis et al. (2016:41) write that "These two types of hair cells work together such that the auditory nerve transmits highly selective information about frequency, timing, and intensity of sounds to the brain." The transduction system is represented pictorially by Figure 2:

---

[4] The number of critical bands varies depending on whether or not the bands are overlapping. Pope (1998:1347) shows overlapping bands, while Rabiner and Juang (1993: 184) give an example of non-overlapping bands.

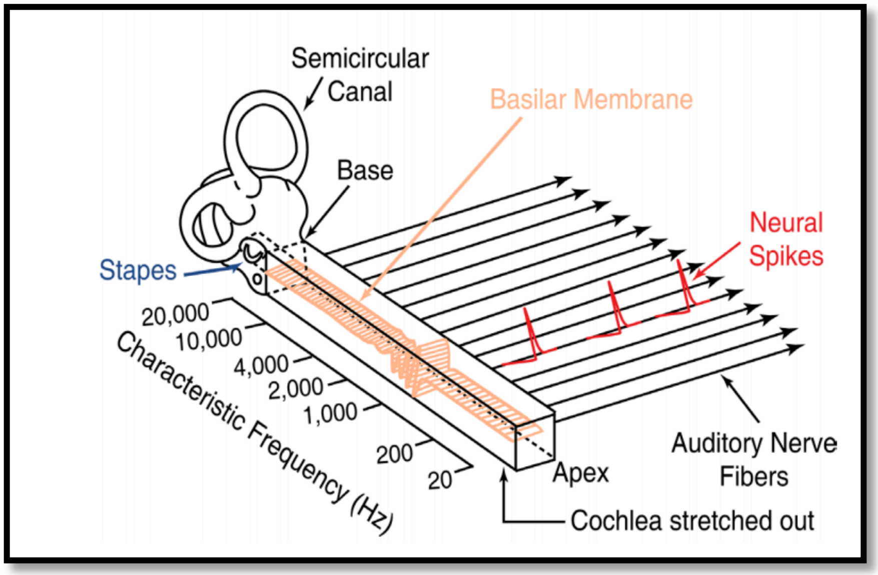

Figure 2: Audibility Range in the Frequency Domain
Sachs, M. B., Bruce, I. C., Miller, R. L., and Young, E. D. (2002). Biological basis of hearing aid design. *Annals of Biomedical Engineering* 30, 157-168. doi:10.1114/1.1458592. Reprinted by permission of © Biomedical Engineering Society.

Von Békésy, another physicist, did various astounding experiments that proved clinically that Fletcher's theory of critical bands was anchored in physiological reality. In recognition of his ingenious experimental designs and his seminal findings, he was awarded the Noble Prize in Physiology/Medicine in 1961. Fastl and Zwicker (2007:158, 233) underscore the importance of critical bands as follows:

> The critical-band concept is important for describing hearing sensations. It is used in so many models and hypotheses that a unit was defined leading to the so-called critical-band rate scale… CBT measures loudness as well as our hearing system.

Yost (2015: 49) concurs and adds that the **Critical Band Theory (CBT)** is the most influential psychoacoustic theory to date. CBT has led to incredible discoveries that have impacted a wide variety of fields, including but not limited to audio engineering, audiology, architectural acoustics, and music, to name only a few (O'Brien 2018). It has also led to the establishment of many intensity thresholds such as those described in Tables 1, 2, and 3 below. CBT is worth knowing and applying to linguistic research because, as Scharf (1961:205) puts it, it is "a basic unit of hearing." In other words, whatever the human ear perceives can be described by CBT.

3.2 A Glance at some Important Intensity Thresholds

Intensity thresholds discovered by the experts working within the CBT framework are relevant to linguistics and to auditory health. The listing in Table 1 is a summary of information collected from various sources, including Murphy (2017), Schimitta's (2016), Roy and Siebein (2019)~~-~~, Fry (1979:95), and Fletcher (1953:104). They describe intensity levels that the human ear perceives:

N0	Environmental Noise	Intensity Levels in dB(A)
1.	Rustling of leaves	30
2.	Quiet whisper	40
3.	Quiet home	50
4.	**Normal conversation**	**60**
5.	**Classroom speech**	**65-75**
6.	Flushing toilet	78
7.	City traffic (inside of a car)	85
8.	Train or truck traffic	90
9.	Motorcycle	90
10.	Jackhammer	95
11.	Subway station	100
12.	Power mower	115
13.	Rock concert	115
14.	Chainsaw	120
15.	Jet plane	140
16.	Shotgun	165
17.	Rocket launch	180
18.	Loudest sound possible	194[5]

Table 1: Intensity Levels of Environmental Sounds/Noises

Many apps and SLMs list the above-mentioned thresholds as background literature. In recent decades, governmental agencies and international regulatory bodies have issued guidelines about intensity thresholds regarded as harmful to humans, as listed in Table 2:

N0	OSHA Daily Permissible Noise Levels	Intensity Levels in dB(A)
1.	8 hours per day	90
2.	6 hours per day	92
3.	4 hours per day	95
4.	3 hours per day	97
5.	2 hours per day	100
6.	1.5 hours per day	102
7.	1 hour per day	105
8.	30 minutes per day	110
9.	About 15 to 20 minutes per day	115

Table 2: Noise Regulations by OSHA

In the US, OSHA (Occupational Safety and Health Administration) is the regulatory agency that is mandated to enforce these guidelines. However, this does not prevent other professional bodies such as the National Institute for Occupational Safety and Health (NIOSH) to propose their own guidelines. Whereas OSHA uses an exchange rate of 5 dB, NIOSH uses only a 3 dB exchange

---

[5] Lubert (2018:38) notes that rocket launches from 300 m away from the launch pad can generate intensity up to 200 dB during liftoff. She notes that the launch of NASA's Saturn V generated intensity at 204 dB.

rate. For example, NIOSH's "dangerous intensity levels" begin at 85 dB whereas for OSHA, they begin at 90 dB.

3.3 Correlations between Intensity and Frequency

Intensity and frequency are two important acoustic correlates. They are independent of each other and yet they are interdependent. This ambiguous relationship between the two has led to many erroneous statements and interpretations. It has been claimed falsely that an increase in one translates automatically into an increase in the other. However, this is not true when it comes to speech sounds that are produced or perceived. Stevens (2000:228) is on record as saying that "For normal listening levels (roughly 60 to 80 dB above the absolute threshold), the JND is independent of amplitude." In other words, there is no discernible correlation between an increase in frequency and a corresponding increase in intensity, and vice versa. This is exactly what Fletcher (1953:101) found decades earlier, as shown in Table 3.

N0	**Frequencies in Hz**	**Intensity Levels in dB**
1.	400	17
2.	600	17.2
3.	800	17.6
4.	1000	18
5.	1200	18.4
6.	1400	18.8
7.	1600	19.2
8.	1800	19.6
9.	2000	19.9
10.	2500	20.9
11.	3000	21.5
12.	3500	22.3
13.	4000	23.1

Table 3: Correlation between Frequency and Intensity

The data in Table 3 shows unambiguously that for frequencies between 400 and 4,000 Hz, the ones that are the most relevant in speech production and perception, an increase in frequency has a very negligible effect on an increase in intensity.

3.4 Psychoacoustic Classification of Acoustic Comfort

The final issue to be discussed before embarking on the linguistic applications of intensity is that of **"acoustic comfort."** Some intensity levels are deemed unpleasant. The ancient Greeks referred to them as "bad" or "evil" sounds, which translates into English as "cacophony." In contemporary acoustic phonetics, all such sounds are classified as **noise**. However, it is only recently that researchers have begun codifying when a sound is "comfortable" or "uncomfortable." Roy and Siebein (2019:25) list four thresholds:

1. **Quiet** (≤ 70 dBA): safe for hearing, great for conversation
2. **Moderate** (71-75 dBA:): safe for hearing, conducive to conversation
3. **Loud** (dBA 76-80): likely safe for hearing, difficult for conversation
4. **Very Loud** (≥ 81 dBA): Long exposure can cause hearing loss.

These thresholds have been operationalized and programmed into Gregory Scott's SoundPrint app which helps users determine the levels of acoustic comfort in certain venues before they set foot in such places.[6]

3.5 Speech Intensity Levels

The intensity level of human speech has been of interest to experts because it can be harnessed for application in many fields. It is, therefore, not surprising to find that one of the most influential publications on the subject is by NASA. It commissioned Pearsons and Horonjeff (1982) to carry out studies on speech intensity. For our purposes, the most relevant aspects of their findings are displayed in Table 4:

N0	**Speech Levels**	**Intensity Levels in dB(A)**
1.	Normal	47-65
2.	Raised	66-75
3.	Loud	76-86
4.	Shout	87-97

Table 4: Intensity Levels in Speech Levels

These four speech intensity levels line up perfectly with Fletcher's findings some 30 years earlier:

> When one talks as loudly as possible, the talker level can be raised to about 86 dB, and when talking as softly as possible, it can be lowered to 46 dB, so from a soft whisper to a loud shout there is a range of 60 dB. … About 40% of the people have talker levels within 3 dB of the average 66 dB, and no person had a talker level greater than 75 dB (Fletcher 1953:77).

Brill et al. (2018:16) report that the average intensity of teacher's speech is 67 dB(A) measured at 1 m from students. This corresponds to a "raised" speech level. Concurrent findings such as these have led experts to take 60 dB(A) as the reference speech level for two people conversing at 1 m away from each other (as noted in Table 1). Fry (1979:94) provides additional evidence for why 60 dB is a good standard reference: "The figure of 60 dB is therefore an average for the number of speakers and also an average over a stretch of conversation, taking into account the syllable variation in intensity." We will return to this threshold later in 5.1 in the formulation and justification of **Equation 4**.

3.6 Important Thresholds in Speech Perception

If 60 dB(A) is the ideal reference level for speech production, what would be the ideal intensity levels at which the human ear can perceive intensity differentials? Many psychoacoustic experiments have yielded the following results:

N0	**Perception of Increases**	**Intensity Levels in dB(A)**
1.	Imperceptible change	1
2.	Barely perceptible change	3
3.	Clearly perceptible change	5

---

[6] I'm a noise ambassador and use this app to report noise levels in restaurants, churches, stadiums, etc.

4.	Twice as loud	10
5.	Four times as loud	20

Table 5: Relative Intensity Thresholds[7]

The information in Table 5 points to two important JNDs. The first is ± 3 dB and the second is ± 5 dB. The former is used in hearing acuity research for speech intelligibility, whereas the latter is used for auditory screening for health and environmental noise pollution. Since the goal in this paper is on intensity and its linguistic applications, we concentrate only on the former. Accordingly, the JND is stated as follows:

JND of Intensity in Speech Intelligibility

> Speech segment or suprasegment A is perceived as being louder than speech segment or suprasegment B if and only if the intensity difference between them is ≥ 3dB.

This JND is widely used in audio engineering products and in sound level meter (SLM) devices. It is found in the sensitivity specification of products such as earphones and headphones. Consequently, when speech sounds are compared and contrasted on the basis of this JND, intensity measurements are not as worthless as Ladefoged (2003:93) and Styler (2017:24) would have us believe.

4.0 Intensity Measurements of Speech Segments

The preceding sections have provided the necessary background information for us to begin to apply intensity to linguistic analyses. The first issue worth considering is whether or not vowels and consonants have intrinsic intensity values. The data for vowels comes from two sources: Lehiste and Peterson (1959) and the longitudinal intensity data that I have been collecting on Central Minnesota English (CMNE) vowels for nearly 10 years. Fletcher (1953:418) is of the opinion that only the intensity of vowels is worth studying because, "As long as the vowel is heard, there is always a chance of identifying the consonant preceding or succeeding it, and consequently the threshold of a consonant so considered will be the same as that for the vowel." Even so, we will also investigate the intensity of fricatives because, according to Ladefoged and Maddieson (1996:78), intensity seems to be the only acoustic correlate that sets them apart from other consonants.

4.1 Intensity of Vowels in Lehiste and Peterson

Lehiste and Peterson (1959) investigated the intensity of vowels of American English by having a male speaker produce them in multiple experiments over four consecutive days in an anechoic room. The vowels occurred in CVC syllables within the carrier sentence *"Say the word ... again."* They measured intensity from the midpoint of vowels so as to minimize the effects of the preceding or following consonant. All in all, they measured the intensity of 1,263 vowel tokens. The averages of the 11 phonemic monophthong vowels are displayed in Table 6:

---

[7] These thresholds are endorsed by the American National Standards Institute (ANSI), NIOSH, OSHA, and by international organizations such as the World Health Organization (WHO), the International Standardization Organization (ISO), and the International Electrotechnical Commission (IEC) and many other regulatory bodies.

Vowels	**[i]**	**[ɪ]**	**[e]**	**[ɛ]**	**[æ]**	**[ɑ]**	**[ɔ]**	**[o]**	**[ʊ]**	**[u]**	**[ʌ]**
L&P (1959)	75	78	78	78	79	80	80	79	78	78	79

Table 6: Intensity Levels of Vowels

They indicated in their paper that "In listening casually to the recorded vowels, they appear to be about equally loud." The only exception is [i] whose intensity is 3 dB lower than all the other vowels.

4.2 Intensity of Vowels in CMNE

There are striking similarities between Lehiste and Peterson's (1959) findings and mine displayed in Tables 7 and 8, even though the stimuli, the recording environments, and the number of speakers are different. The data come from 55 speakers of Central Minnesota English (CMNE). They were college students who enrolled in my acoustic phonetics and sociophonetics courses from 2011 to 2019. The participants recorded themselves reading the 11 phonemic English vowels contained in Tables 7 and 8. They were instructed to read the words as naturally as possible. It is worth noting that the recordings took place early in the semester. So, the students did not know the purpose of the recordings. All the participants signed an informed consent that was approved by the Institutional Review Board (IRB) of St. Cloud State University. Each vowel was repeated three times. The 55 participants produced a total of 1, 851 vowel tokens (11 x 3 x 55). Table 7 contains female data, while Table 8 has the intensity measurements of male speakers.

Words	**heed**	**hid**	**hayed**	**head**	**had**	**hod**	**hawed**	**hoed**	**hood**	**who'd**	**hud**
Women	[i]	[ɪ]	[e]	[ɛ]	[æ]	[ɑ]	[ɔ]	[o]	[ʊ]	[u]	[ʌ]
Speaker 1F	62	65	61	63	61	62	64	64	62	63	59
Speaker 2F	68	71	71	73	74	75	74	74	75	73	80
Speaker 3F	62	62	59	64	61	64	62	62	63	63	63
Speaker 4F	64	73	58	62	68	73	74	76	69	69	72
Speaker 5F	64	64	63	64	63	63	62	65	64	66	64
Speaker 6F	73	77	74	76	73	77	76	74	75	74	77
Speaker 7F	80	84	83	85	84	84	85	85	95	84	86
Speaker 8F	67	73	70	68	65	68	66	70	72	69	67
Speaker 9F	77	78	76	72	72	75	75	77	77	78	74
Speaker 10F	65	64	64	66	67	69	70	66	66	66	68
Speaker 11F	53	41	51	53	56	58	58	53	53	53	53
Speaker 12F	58	49	50	50	50	49	53	53	53	52	51
Speaker 13F	72	70	70	69	72	74	73	70	69	70	71
Speaker 14F	52	56	54	55	54	59	55	56	58	57	55
Speaker 15F	65	66	67	71	72	75	74	72	70	67	70
Speaker 16F	75	74	75	76	76	79	81	74	75	73	79
Speaker 17F	58	60	60	63	59	59	61	58	58	59	58
Speaker 18F	69	70	69	69	69	66	71	67	69	69	70
Speaker 19F	61	59	57	58	58	58	57	59	57	59	57
Speaker 20F	58	61	61	64	63	65	62	63	64	63	60
Speaker 21F	76	79	77	78	78	79	76	79	79	79	80
Speaker 22F	76	80	76	76	73	75	73	76	78	77	77
Speaker 23F	61	64	64	62	64	64	69	63	61	62	61
Speaker 24F	75	75	74	74	74	72	73	73	75	78	73
Speaker 25F	77	85	80	83	80	81	83	81	82	76	80
Speaker 26F	62	66	67	74	67	66	62	68	65	68	69
Speaker 27F	58	60	60	63	59	59	61	58	58	59	58

Speaker 28F	68	75	72	74	79	75	74	74	75	70	73
Speaker 29F	66	71	72	71	74	73	74	72	73	70	74
Speaker 30F	41	47	47	51	49	49	50	49	49	48	47
Speaker 31F	70	73	71	71	64	69	68	70	69	71	67
Speaker 32F	71	71	71	71	67	67	67	70	69	75	69
Speaker 33F	63	72	71	72	71	68	71	73	73	71	67
CMNE Mean	**65**	**67**	**66**	**67**	**67**	**68**	**68**	**68**	**68**	**67**	**67**
St. Dev	8.4	10	8.9	8.5	8.5	8.6	8.5	8.6	9.5	8.3	9.4

Table 7: Intensity Levels of CMNE Female Vowels

Words	**heed**	**hid**	**hayed**	**head**	**had**	**hod**	**hawed**	**hoed**	**hood**	**who'd**	**hud**
Men	[i]	[ɪ]	[e]	[ɛ]	[æ]	[ɑ]	[ɔ]	[o]	[ʊ]	[u]	[ʌ]
Speaker 1M	79	82	78	82	79	82	83	82	78	78	84
Speaker 2M	58	58	55	55	54	56	54	52	55	52	55
Speaker 3M	53	61	58	58	54	54	55	58	60	55	58
Speaker 4M	66	71	69	69	67	68	68	69	69	67	70
Speaker 5M	75	75	73	76	75	74	73	76	75	74	77
Speaker 6M	67	72	71	75	74	71	73	70	70	70	70
Speaker 7M	89	90	90	90	87	89	90	90	90	89	88
Speaker 8M	86	82	84	82	85	87	87	88	88	90	80
Speaker 9M	78	85	79	79	78	77	80	81	82	77	77
Speaker 10M	77	80	78	80	79	80	80	82	80	81	80
Speaker 11M	57	63	62	62	64	64	65	62	64	62	63
Speaker 12M	65	70	67	68	67	68	67	67	69	68	70
Speaker 13M	56	60	58	59	57	58	61	61	60	58	61
Speaker 14M	73	77	76	76	74	75	74	76	77	75	75
Speaker 15M	64	66	66	66	64	66	65	64	64	63	63
Speaker 16M	67	65	62	64	63	65	64	65	63	61	61
Speaker 17M	52	57	55	51	54	52	51	51	54	50	47
Speaker 18M	60	63	59	60	60	62	65	61	62	62	58
Speaker 19M	74	74	74	76	75	77	75	76	77	77	77
Speaker 20M	57	57	55	51	53	54	55	52	58	55	51
Speaker 21M	63	65	63	62	60	63	62	63	65	63	62
Speaker 22M	51	55	50	51	50	51	51	47	52	52	50
CMNE Mean	**66**	**69**	**67**	**67**	**66**	**67**	**68**	**67**	**68**	**67**	**67**
St. Dev	10	10	10	11	11	11	11	12	10	11	11

Table 8: Intensity Levels of CMNE Male Vowels

The intensity measurements in these two tables are similar to those in Lehiste and Peterson (1959) and Fletcher (1953:77) in two important respects. First, in my study and Lehiste and Peterson's, [i] is the least sonorous vowel. Secondly, the combined mean intensity of all vowels in Tables 7 and 8 is 66 dB versus 67 dB in Fletcher.

4.3 Gender Differences and Intensity

French and Steinberg (1947:93) estimate that the intensity of speech segments produced by male speakers is ≥ 3 dB louder than that of female speakers. Presumably, this may have something to do with supralaryngeal differences in both genders. Let's see if this prediction is borne out by our measurements.

Words	heed	hid	hayed	head	had	hod	hawed	hoed	hood	who'd	hud
Men	[i]	[ɪ]	[e]	[ɛ]	[æ]	[ɑ]	[ɔ]	[o]	[ʊ]	[u]	[ʌ]
CMNE Women	65	67	66	67	67	68	68	68	68	67	67
CMNE Men	66	69	67	67	66	67	68	67	68	67	67
Difference	1	2	2	0	1	1	0	1	0	0	0

Table 9: Intensity Levels between Male and Female Vowels

The mean intensity measurement across all 11 vowels is 67.181 dB in male speech versus 67.375 dB for females. This indicates that perceptually, there is no gender-based differences in intensity among the participants in my study. However, in a similar study conducted by Lopez-Backstrom (2018:39, 138-9) on Northern Minnesota speakers, she found a perceptually salient difference of 4.37 dB between 10 male speakers (58.48 dB) and 10 female speakers (54.11 dB). It is worth noting that four of six female participants are 30 to 64 old. Since vocal fry phonation is less prominent among this age group, this would help to explain why Lopez-Backstrom found a large intensity difference between male and female but the data in Table 9 do not.[8] It seems that whether or not an intensity difference exists between male and female would depend to a large extent on the presence or absence of vocal fry. This means that French and Steinberg's estimate should be taken under advisement.

4.4 Intensity of Consonants (Fricatives)

There is a severe paucity of data on the intensity of consonants. Mentions have been made here and there that some consonants have greater intensity than others. Ladefoged and Maddieson (1996:78) note, for example, that ejectives have "a greater amplitude in the stop burst." They indicate on page 259 that clicks have very high intensity, at least 6 dB greater than other segments. Their descriptions of the labio-velar segments [k͡p] and [g͡b] on pages 332-43 suggest that these are also loud sounds. They also state on page 139 that intensity of fricatives is worth knowing because it helps to differentiate between them and other sounds. For fricatives in American English, we have intensity measurements from Jongman et al. (2000:1259), as indicated in Table 10:

Segments	**s**	**z**	**f**	**v**	**θ**	**ð**	**ʃ**	**ʒ**
Jongman (2000)	64.9	67.7	55.7	63.2	54.7	62.7	66.4	68.2

Table 10: Intensity of Fricatives

These measurements show that voiced fricatives are louder than their voiceless counterparts. In almost all instances, except for the pair [ʃ] and [ʒ], the difference is perceptually salient. Again, in light of findings such as these and others alluded to in this paper, the relevance of intensity in linguistic analysis should not be in dispute.

---

[8] There are various reasons for why this study did not uncover any intensity difference between males and females. The prevalence of vocal fry among the female college-age participants in my data could be one plausible explanation. Indeed, Koffi (2019:19) found that breathy voice due to vocal fry was prevalent among the participants in Table 7. Other possible explanations include differences in the microphones used in the recordings. Finally, Lopez-Backstrom's recordings took place in controlled environments whereas the participants in my study recorded their speech in various rooms.

4.5 The Intensity of Suprasegmentals in English

The contribution of intensity to the perception of lexical stress in English has been well-documented. Findings go as far back as Fry (1958:151). In a groundbreaking study, he ranked the acoustic correlates of lexical stress, i.e., F0, intensity, and duration, and found that intensity ranked last: F0 > Duration > Intensity. However, a replication study by Lieberman (1960:453) showed that intensity ranked second: F0 > Intensity > Duration. Kochanski et al. (2005:1046, 1052) did a major study in which intensity ranked first: Intensity > Duration > F0. What these findings seem to indicate is that there is still a lot about the linguistic applications of intensity that we do not know.

4.6 The Measurements of Intensity in Tone Languages (Anyi)

When it comes to intensity and tone, the lack of data is astounding. Consequently, even a succinct literature is not possible. For this reason, I base intensity measurements of tone languages on Anyi Morofu data I collected a while back. Anyi is an Akan language spoken in eastern Côte d'Ivoire. Lexical and grammatical tones are contrastive. The demonstration below focuses on grammatically contrastive tones in one underlying sentence produced by 10 male native speakers. Variation in prosody conveys three distinct grammatical moods:

Sentence 1: *Ɔ* *bʊka* *Kasi*
He/she *helps* *Kasi* (Indicative Mood)
Sentence 2: *Ɔ* *bʊka* *Kasi*
He/she wants to *help* *Kasi* (Intentional Mood)
Sentence 3: *Ɔ* *bʊka* *Kasi*
Let him/her *help* *Kasi* (Subjunctive Mood)[9]

Grammatical differences between these three sentences depend exclusively on prosodic modulations in F0, intensity, and duration of the subject pronoun <**ɔ**> and the vowel <**ʊ**> of the verb <bʊka>. For the purposes of this demonstration, we focus exclusively on the intensity <***ɔ***> and <***ʊ***>.

Indicative Mood	*ɔ*	*bʊ*	*ka*	*Ka*	*si*
F0	122	126	129	120	137
Intensity	**78**	**76**	75	76	74
Duration	77	54	89	67	84
Intentional Mood	*ɔ*	*bʊ*	*ka*	*Ka*	*si*
F0	152	171	129	120	137
Intensity	**78**	**74**	75	76	74
Duration	97	56	89	67	84
Subjunctive Mood	*ɔ*	*bʊ*	*ka*	*Ka*	*si*
F0	121	164	129	120	137
Intensity	**78**	**77**	75	76	74
Duration	95	65	89	67	84

Table 11: Intensity Contrast

---

[9] Technically speaking, this is a cohortative mood.

The 10 participants produced 60 measured TBUs (2 vowels x 3 sentences x 10 speakers). The intensity of the subject pronoun <ɔ> is 78 dB in the three grammatical moods. This shows that the intensity of the subject pronoun is invariable. The intensity of the TBU of <ʊ> is 76 dB in the indicative, 74 dB in the intentional mood, and 77 dB in the subjunctive. The take-away from this preliminary analysis is that intensity contributes to discriminate between the intentional mood and the subjunctive mood. In the former, the TBU of <ʊ> in <bʊka> is 74 dB, whereas it is 77 dB in the latter. Since the intensity difference is ≥ 3 dB, we conclude that Anyi speakers may rely on intensity, among other cues, to encode and decode tonal differences between these two moods, but not between the indicative (76 dB) and the subjunctive (77 dB) or the indicative (76 dB) and the intentional (74 dB).

5.0 Application of Intensity to Phonological Theory

Intuitively, people know that when some speech sounds are produced in isolation or combined with others, some are more sonorous than others. Baken and Orlikoff (2000:110-1, 265) refer to this as **vocal intensity**. Ladefoged and Johnson (2015:255) illustrate it with the following example:

> Try saying just the vowels [i, e, a, o, u]. You can probably hear that the vowel [a] has greater sonority (due, largely, to its being produced with a greater mouth opening). You can verify this fact by asking a friend to stand some distance away from you and say these vowels in a random order. You will find that it is much easier to hear the low vowels [a] than the high vowel [i, u].

In phonological circles, a speech sound is said to have its own intrinsic **"sonority index."** Bloomfield (1933:120-1) discussed it rather extensively. However, Malmberg (1963:65-6) credits Jespersen as being the first among linguists to have assigned numerical indices to specific sounds, thereby systematizing and ranking speech sounds on a sonority scale such as the one in Table 12:

N0	**Roca and Johnson (199:288)**		**Goldsmith (1990:112)**	
	Segment	**Sonority Scale**	**Segments**	**Sonority Index**
1.	Non-high vowels	6	/a/	10
2.	High vowels	5	/e, o/	9
3.	Liquids	4	/i, u/	8
4.	Nasals	3	/r/	7
5.	Fricatives	2	/l/	6
6.	Stops	1	/m, n/	5
7.			/s/	4
8.			/v, z, ð/	3
9.			/f, θ/	2
10.			/b, d, g/	1
11.			/p, t, k/	0.5

Table 12: Phonology-based Sonority Indexes

Over the years, disagreements have emerged and various indices have been proposed (See Guffey 2002 for a review of competing scales). Nevertheless, the usefulness of such a system has not diminished. On the contrary, it has been invoked to account for a universal syllable phonotactic constraints called the **Sonority Sequencing Principle (SSP)**, which Roca and Johnson (1999:266) formulate as follows:

Sonority Sequencing Principle

The sonority of a syllable must rise until it peaks, and then fall.

Sonority indices such as those in Table 12 have come under scrutiny because the correlation between segments and the numerical values assigned to them is arbitrary. Yet, Goldsmith (1990:111-2) defends the practice, arguing that,

> While there is considerable skepticism that the ultimate account of sonority is one based on an arithmetic system of this sort, there may be something right about an account that is sufficiently oriented to measuring[10] sonority differences to be able to state unambiguously that liquids are halfway between obstruents and vowels. If this is correct, then we may characterize languages with respect to how much sonority difference they demand of successive segments.

An ancillary support of the SSP is the **Minimal Sonority Distance Parameter (MSDP)**. It explains arithmetically why concatenations of some segments are permissible but others are not. Goldsmith (1990:111) states it as follows:

Minimal Sonority Distance Parameter (MSDP)

The difference in sonority of successive segments must be greater than **2**.

Unfortunately, he does not explain anywhere in his book where the "**2**" in the formulation of the MSDP comes from. We are supposed to accept it and use it without further explanation. The "**2**" in the formulation of the MSDP is also arbitrary. In spite of the arbitrariness of the SSP and the MSDP, phonologists love them because they have strong explanatory powers. Morelli (2003:358) is of the opinion that the SSP and the MSDP express strong universal tendencies even if "violations are attested across languages." I concur with him. Yet, I contend that the SSP and the MSDP can be strengthened if the sonority indices on which they are based are firmly grounded in the physics of speech, i.e., acoustic phonetics. This is what we will endeavor to demonstrate in the sections below.

5.1 Grounding Sonority Indices in Physics

Ladefoged and Johnson (2015:255) state that "The sonority of a sound can be estimated from the measurement of the acoustic intensity of a group of sounds that have been said on comparable pitches and with comparable degrees of length and stress." This is a clue that sonority indices can be calculated in a non-–arbitrary manner. The sonority indices in Table 13 are calculated according to the same logic. The data is taken from Fry (1979:126-7).

---

[10] The use of "measuring" in this quote is misleading. Nowhere does Goldsmith or any other phonologist provide proof of measurement in the assignment of these numerical values.

N0	Segments	Sonority Index	N0	Segments	Sonority Index
1.	ɔ	29	21.	ʃ	19
2.	ɑ	26	22.	ŋ	18
3.	ʌ/ə	26	23.	m	17
4.	ɛ	25	24.	tʃ	16
5.	æ	24	25.	n	15
6.	ʊ	24	26.	dʒ	13
7.	e	23	27.	ʒ	13
8.	o	23[11]	28.	z	12
9.	ɪ	22	29.	s	12
10.	u	22	30.	t	11
11.	i	22	31.	g	11
12.	ɑɪ	26; 22	32.	k	11
13.	ɑʊ	26; 24	33.	v	10
14.	ɔɪ	29; 24	34.	ð	10
15.	eɪ	23; 22	35.	b	8
16.	oʊ	22; 24	36.	d	8
17.	w	21	37.	p	7
18.	r	20	38.	f	7
19.	j	20	39.	θ	0[12]
20.	l	20	40.	h	variable[13]

Table 13: Acoustically-based Sonority Indexes

Even though Fry does not specify explicitly the formulas that yielded these indices, it is not hard to derive them from **Equation 4**:

$$\text{Sonority Index} = 2 \text{ x } 10 \log_{10} = \frac{\text{Highest Segment Intensity}}{60 \text{ dB}}$$

Here is an explanation and a demonstration of the equation. The numerator corresponds to the loudest intensity a person can produce in a sustained pronunciation of any given speech segment. In Fletcher's (1953:77) experiments at the Bell Laboratories, he found that "When one talks as loudly as possible, the talker level can be raised to about 86 dB." The denominator is kept constant at 60 dB because, as noted in section 3.2, numerous studies concur that this is the intensity level at which unmarked conversations occur. This explains why 60 dB is taken as the "reference level"

---

[11] Fry (1979:127) does not report any value for [o]. The value reported here is based on the log calculations of 33 female students whose data appear in Table 7.

[12] The value of "0" dB for [θ] does not mean absence of intensity, but rather that the intensity is so faint that it is at the threshold of hearing. The intensity of [θ] is as loud as the sound one makes when ones does not have a cold and one is breathing normally. On April 13, 2018, I conducted an experiment in a state of the art sophisticated anechoic room at the Starkey Hearing Technologies in Eden Prairie, MN. I measured the intensity of my breathing using several sound level meter apps on my smart phone. I sat my smart phone on a platform, stepped away for about 1 m, and the intensity registered as 5 dB, which translates into about 1.6 dB on the sonority index scale. Equation 4 is applied to the measurements.

[13] "Variable" means that [h] takes the intensity value of the vowel it precedes.

(Fry 1979:126). The prefix "**2**" occurs in **Equation 4** because SPL fluctuates as a factor of 2 in direct proportion to the distance between the source and the destination of the sound (Speaks 2005:135-6). If the distance between the source and the destination is reduced by half, intensity increases by 6 dB, that is, 3 dB by a factor 2. On the other hand, if the distance between the source and the destination is doubled, intensity decreases by 6 dB, also 3 dB by a factor of 2. The rest of Equation 4 is the same as Equation 3, as discussed in 2.0.

Since we know from Fletcher (1953:77) that the loudest intensity that a human can produce is 86 dB, and since, according to Fry (1979:126-7), [ɔ] has the highest sonority index, we can use Equation 4 to derive all the sonority index of [ɔ]. We divide the numerator 86 by the denominator 60 dB. It yields 1.43. Then, the product is multiplied by 10, which gives 14.33. This product is also multiplied by 2. The new product is 28.66, which is rounded up to 29 on the sonority scale. This is a reasonable assumption of how all the sonority indices in Table 13 were derived. Fry (1979:126) makes a very important statement about sonority indices that is worth quoting. He notes that "[They] remain roughly the same in loud speech, in conversational speech and in quieter speech." It is also important to underscore that sonority indices such as those in Table 13 are universal because all human beings produce speech within the same intensity ranges, that is, 86 dB for the loudest sound to 60 dB for conversational speech intensity. Furthermore, all humans are subject to the same aerodynamic speech laws. Slight changes in altitude, atmospheric conditions, and barometric pressures do not matter much, as already noted in section 2.0. When all the preceding elements are taken into account, they lead to a reformulation of the MSDP as follows:

Minimal Sonority Distance Parameter (MSDP)

The difference in sonority of successive segments must be greater than 3 dB(A).

The MSDP as reformulated above is solidly grounded in the physics of speech. The 3 dB(A) intensity response is not arbitrary. On the contrary, it is based on nearly 80 years of acoustic phonetic experiments and findings. For the sake of simplicity, the dB(A) weighting scales is dropped. Only "**3**" is used in all subsequent analyses.

5.2 Exemplification of the SSP with CV and CVC Syllables

The time has come to apply the physics-based sonority indices in Table 13 and the reformulated MSDP to language data. Since sonority indices are the same for all languages, the SSP and MSDP can provide important insights on the intensity constraints that belie syllable structures in all languages. The syllable types used in the demonstrations below come from Fletcher's (1983:94-96) study of 80,000 English words.

1. V (9.7%)
2. VC (20.3%)
3. CV (21.8%)
4. CVC (33.5%)
5. VCC (2.8%)
6. CCV (.08%)
7. CVCC (7.8%)
8. CCVC (2.8%)

9. CCVCC (0.5%)

Even though the lexical items come from English, the canonical types themselves are attested in a wide variety of languages. Clements and Keyser (1983:29) report that CV is the only type that is truly universal because it is found in all languages. How do the SSP and the MSDP apply to the CV? The word <go> [go] is used to illustrate the sonority profile of CV syllables, as seen in Figure 1:

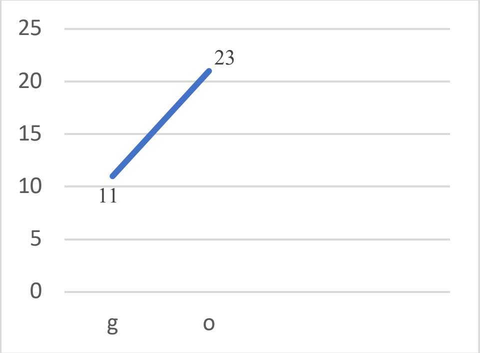

Figure 1: Sonority Profile of CV-Type

The MSDP is verified because sonority distance between [g] (11) and [o] (23) is 12, which is greater than 3.

The CVC-type is the most common in English (33.5%). The sonority profile of CVC-type syllables is exemplified by <rob> [rɑb]:

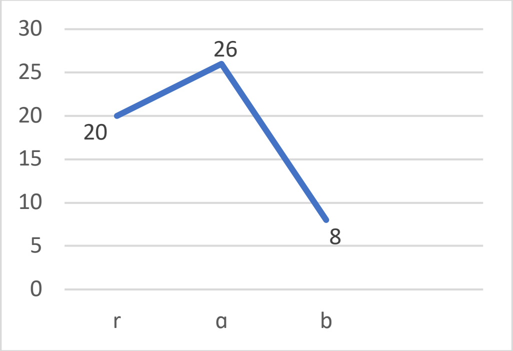

Figure 2: Sonority Profile of CVC-Type

Syllables of this type illustrate the SSP and MSDP the best because they have an onset, a nucleus, and a coda (Clements and Keyser 1983:29). They often conform to the SSP and MSDP without fail because the sonority rises from C_1 to V and then fall back on C_2. The sonority distance between C_1 and V is greater than 3, and the sonority distance between V and C_2 is also greater than 3. The word <rob> [rɑb] illustrates this perfectly. Sonority rises from [r] (20), peaks at [ɑ] (26), and then falls on [b] (8). The sonority distance between [r] (20) and [ɑ] (26) is greater than 3, and so is the sonority distance between [ɑ] (26) and [b] (8). No violation of the SSP and the MSDP has been found in any human language for CV and CVC syllables.

5.3 Exemplification of the SSP with Complex CCV and CCCVC Onsets

Complex onsets of the CCV(C) and CCCV(C) types present various challenges to the SSP and the MSDP. Both types are exemplified by Figures 3 and 4:

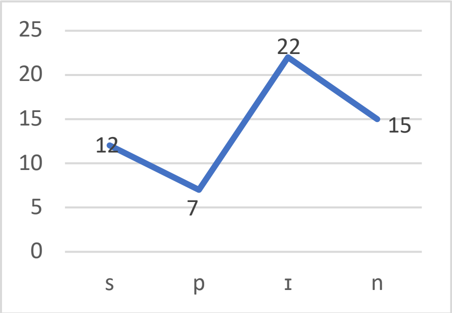

Figure 3: Sonority Profile of CCVC-Type

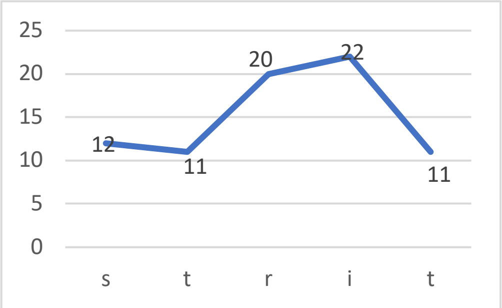

Figure 4: Sonority Profile of CCCVC-Type

Figures 3 and 4 violate the SSP because the sonority profile does not rise from the onset cluster to the nucleus. Instead, in [spɪn], the sonority dips from [s] (12) to [p] (7) before rising to [ɪ] (22). We see the same pattern in [strit]. Sonority rises from [s](12) then dips to [t] (11) before rising to [i] (22). Both onset clusters violate the SSP but only the [st] cluster in [strit] violates the MSDP because the sonority difference between [s] (12) and [t] (11) is less than 3. Whereas [sp], [st], and [sk] violate the SSP, only [st] and [sk] violate the MSDP. Clusters such as these, which are found in English and other languages, confirm the widely known fact that SSP and MSDP are "soft" universals. Yet, the phonotactic constraints that they express apply to a very large number of languages, including all African languages (Clements 2000:123-160). Welmers (1973:65) notes that complex onsets such as /mŋkp/ and /mŋgb/ are not complex in the true sense of the word because in all such cases, the /mŋ/ sequences acts as a syllable nucleus by itself.

5.4 Exemplification of the SSP with Complex CVCC(C) Codas

Some syllables have complex codas of CVCC or CVCCC-types. Some violate the SSP and the MSDP but others violate one but not the other. In the case of the words <robbed> [rɑbd] and <six> [sɪks] in Figures 5 and 6, <robbed> [rɑbd] conforms to the SSP but violates the MSDP, whereas <six> [sɪks] violates both the SSP and the MSDP. Let's explain why.

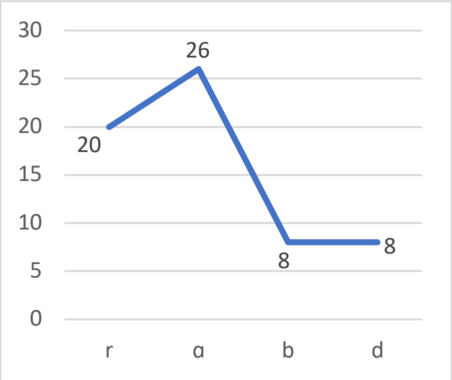

Figure 5: Sonority Profile of CVCC-Type

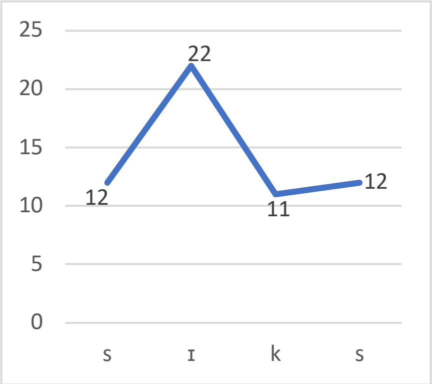

Figure 6 : Sonority Profile of CVCC-Type

In the word [rɑbd], the complex coda adheres to the SSP because sonority drops from [ɑ] (26) to [b] (8). Since [b] and [d] have the same sonority index, sonority remains flat. No violation occurs.

However, the MSDP is violated because the sonority distance between [b] (8) and d (8) is 0 dB. In [sɪks], both the SSP and the MSDP are violated. The SSP is violated because sonority rises from [k] (11) to [s] (12). Furthermore, since the sonority distance between [k] (11) an [s] (12) is only 1, the MSDP is also violated. Complex codas clusters such as [ts], [st], [ks], and [kt] violate the MSDP, while others such as [ps], [fs], [vz] violate the SSP. This may explain why languages with heavy coda clusters are, relatively speaking, uncommon. Clements' (2000:123-160) overview of the syllable structure of the 2000 or so African languages does not record any with a heavy coda cluster.

5.5 Unresolved Issues Concerning Vowel-less Syllables

The SSP and MSDP are termed "soft" universals because there is no shortage of languages that offer counterexamples. Serbian (and probably other Slavic languages) is one of them. According to Surducki (1964:177), in addition to having onset clusters and coda clusters of CC or CCC-types, it also has word-medial clusters of CC, CCC, CCCC, and CCCCC-types. There is no doubt that the SSP and the MSDP would have trouble accounting for such a complex amalgamation of consonants. However, the SSP and MSDP face their biggest challenges when they are called upon to account for vowel-less syllables such as those described by Easterday (2017:2, 16, 52, 57). The author classifies world languages into four categories with regard to syllable structure complexity: simple, moderately complex, complex, and highly complex. Languages with highly complex syllables account for 7.6% of world languages. Among them, one finds languages such as Puget Salish, a Native American language where one finds a word such as [sqwəɬ.ps] (cutthroat trout). Notice that the onset has a sequence of three consonants. Notice also that the second syllable has no vowels. Easterday (2017:2) gives also the example of Tashlhyit, an Afro-Asiatic language spoken in Morocco. In this language the word [tʃ.tk.tṣ.tṣtt] (you sprayed it) has four syllables, none of which has any vowel. Finally, we come across a language named Lendu which has the vowel-less disyllabic word [zz̩̀.zź̩] (to drink). The SSP and MSDP are completely violated by these languages and languages like those that appear in Table 2.7 on page 67 and those that appear in Table A4 on page 530. It is worth noting that Easterday's classification of highly complex syllables is made on the basis of data collected impressionistically. Nowhere in the dissertation is spectrographic data provided to support the claim of vowel-less words.

5.6 Unresolved Sonority Issues

The sonority indices in Table 13 do not cover all speech sounds. How then does one deal with segments that are unaccounted for? There are three possible solutions. The first is based on natural class. If the missing segment forms a natural class with another segment in the table, the known sonority index can be applied to the missing one. A case in point is the sounds [ɲ] and [ŋ]. The former is not represented in Table 13 but the latter is. Since both segments form a natural class, that is, [+nasal, -anterior], the same sonority applies to both. In the same vein, the index of [m] can be assigned to [ɱ]. There are numerous examples of this in Table 13: /s/ and /z/, /k/ and /g/, and /b/ and /d/. The second solution applies to the labiovelar [k͡p] and [g͡b]. These sounds do not have direct equivalents in Table 13. Yet, their constitutive segments do. In such a case, the sonority index is calculated by subtracting the sonority indices of [k] and [p] and multiplying the difference by 2.512. The justification of this analysis is found in Bauman (2016).[14] So, the sonority indices of [k͡p] and [g͡b] is 10. The third solution is for speech sounds that do not have

---

[14] The information was retrieved from https://hearinglosshelp.com/blog/converting-decibels-to-sound-intensities/ on December 1, 2019.

any representation at all in Table 13. This would include the clicks [ǀ, ǁ, ǂ, !, ʘ]. One has no choice but to recruit participants, bring them to a lab, and have them produce these clicks as loudly as they can possibly produce them, and apply Equation 4 to calculate the sonority indices of such sounds. For such an experiment, it is recommended not to have less than six participants (Ladefoged 2003:67).

5.6 Sound Power

Sonority indices are not the only acoustic cues showing that speech sounds have inherent intensity attributes. Fletcher's experiments at Bell Laboratories led him to investigate the voltage power in individual speech sounds on the assumption that vocal intensity generates power. This power is created, according to Baken and Orlikoff (2000:110), by the "interaction of subglottal pressure, the biomechanics and aerodynamics at the level of vocal folds, and the status of the vocal track." Fletcher embarked on this experiment because, in the early days of landline telephone communication, answers needed to be found to wonderments such as these:

> For the purposes of engineering telephone systems, it is desirable to know both the acoustic and electric power of the speech being transmitted. If the power becomes too small, it is masked by extraneous noise. If it becomes too large, parts of the transmitting apparatus become overloaded, that is, they fail to transmit the speech without distortion. Also, when the speech is transmitted along one pair of wires which is close to another pair, the cross talk between the two may become objectionable when the speech power in one or both of them becomes too great (Fletcher 1953:86).

Researchers set out to find answers to these questions. For this, Fletcher and his team measured the sound power of every segment in English. They recorded 500 conversations containing 80,000 words which they studied in a variety of ways. Table 14 contains the information found in Fletcher (1953:86) regarding the voltages of individual speech sounds: [15]

N0	Segments	Power in μw[16]	N0	Segments	Power in μw
1.	ɔ	680	21.	ʃ	80
2.	ɑ	600	22.	ŋ	73
3.	ʌ/ə	510	23.	m	52
4.	æ	490	24.	tʃ	42
5.	o	470	25.	n	36
6.	ʊ	460	26.	dʒ	23
7.	e	370	27.	ʒ	20
8.	ɛ	350	28.	z	16
9.	u	310	29.	s	16
10.	ɪ	260	30.	t	15
11.	i	220	31.	g	15
12.	ɑɪ	600; 260	32.	k	13
13.	ɑʊ	600; 460	33.	v	12

---

[15] The information represents data obtained from 16 participants. Their speech was recorded seating approximately 30 cm away from the recording device (p. 79, 83).

[16] 1 watt = 1 million microwatts!

14.	ɔɪ	680; 260	34.	ð	11
15.	oʊ	470; 460	35.	b	7
16.	eɪ	370; 260	36.	d	7
17.	w	210	37.	p	6
18.	r	210	38.	f	5
19.	j	210	39.	θ	1
20.	l	100	40.	h	variable

Table 14: The Phonetic Power of Segments

Fletcher (1953:82-3) was careful to note that the power of speech sounds is not static, but variable as evidenced by the following statement:

> In the course of conversation the fundamental vowel and consonant sounds are produced with varying degrees of power depending upon their position in the sentence and the emphasis desired. In spite of this variation, some of the speech sounds are always more powerful than others.

He also underscored the fact that the voltage in speech is extremely weak. To illustrate this point, he wrote on page 68 that "It would take 500 people talking continuously for a year to produce enough energy to heat a cup of tea." In other words, the voltage of speech is so minute that it is not in danger of overheating telephone lines. Fry (1979:91) echoed the same findings, noting that "It would take more than three million voices all talking at once to produce power equivalent to that which lights a 100 watt lamp." Recent findings by Gray et al. (2019:19) confirm these earlier studies. They report that "The screams from a football stadium full of people barely produce enough sound energy to boil an egg." For the purposes of engineering, the voltage produced by speech power is inconsequential. Yet, according Fry (1979:91), the voltage level is ideal for optimal auditory perception, "The ear and its associated system of nerve fibres form a receiver designed to deal with the very low levels of energy found in sound waves." Recent findings by Lewis et al. (2016:40), Brownell (2017:20-27), and Anderson et al. (2018:11, 12) show that electromotility and otoacoustic emissions make use of these "low levels of energy" to optimize speech perception. Sound power data open up a new line of research that may prove applicable to linguistics.

6.0 Summary

This paper has shown that intensity has more to offer to linguists than previously thought possible. It is true that intensity is not phonemically contrastive in any human language. Yet, it is highly relevant to phonological analyses of syllable phonotactics. The physics-based insights discussed in the latter portion of the paper have shown that the correlation between speech segments and their individual indices should no longer be arbitrary. By appealing to Equation 4, it has been demonstrated that such indices are derived from the loudest possible sustained pronunciation of a segment in isolation versus its occurrence in a normal conversation. Because human beings are subject to the same aerodynamic speech laws and because they have similar laryngeal and supralaryngeal endowments, the intensity levels that they produce in speech are identical, irrespective of the languages that they speak. Because of this, Equation 4 makes it possible to calculate sonority indices of known and unknown speech sounds. Voltage indices of speech segments have not yet been applied in phonological analyses. Yet, the potential is there.

By bringing together physics-based sonority indices and sound power indices, this paper has shown that acoustic phonetic findings about intensity are applicable to linguistic theorization and descriptions of syllable phonotactics.

ABOUT THE AUTHOR

Ettien Koffi, Ph.D. linguistics, teaches at Saint Cloud State University, MN. He is the author of four 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 perspectives 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) and Text-to-Speech (TTS), and voice biometrics for speaker verification. He can be reached at enkoffi@stcloudstate.edu.

References

Anderson, Samira, Sandra Gordon-Salant, and Tudy R. Dubno. 2018. Hearing and Aging Effects on Speech Understanding: Challenges and Solutions. *Acoustics Today* 14 (4): 11-18.

Bloomfield, Leonard. 1933. *Language*. New York: Holt, Rinehart and Winston.

Breysse, Patrick N. and Peter S.J. Lees. 2006. Noise. The John Hopkins University. Online at: The website http://hearinglosshelp.com/blog/converting-decibels-to-sound-intensities/ Retrieved on July 9, 2019.

Brill, Laura C. Kieren Smith, and Lily M. Want. 2018. Building a Sound Future for Students: Considering the Acoustics in Occupied Classrooms. *Acoustics Today* 2018 (14):14-22.

Brownell, William E. 2017. What is Electromotility? The History of its Discovery and its Relevance to Acoustics. *Acoustics Today* 13 (1):20-27.

Clements, George N. 2000. Phonology. In *African Languages: An Introduction*, ed. by Bernd Heine and Derek Nurse, pp. 123-160. New York: Cambridge University Press.

Clements, George N. and Samuel Jay Keyser. 1985. *CV Phonology: A Generative Theory of the Syllable*. Cambridge, MA: The MIT Press.

Easterday, Shelece. 2017. *Highly Complex Syllable Structure: A Typological Study of its Phonological Characteristics and Diachronic Development*. Ph.D. Dissertation. Albuquerque, New Mexico: University of New Mexico.

Everest, Alton F. and Ken C. Pohlmann. 2015. *Master Handbook of Acoustics.* 6th edition. New York: McGraw Hill Education.

Fletcher, Harvey. 1953. *Speech and Hearing in Communication*. New York: D. Van Nostrand Company, Inc.

French, N. R. and J.C. Steinberg (1947:93). Factors Governing the Intelligibility of Speech Sounds. *Journal of the Acoustical Society of America*, 19 (1): 90-119.

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

Fry, Dennis. B. 1958. Experiments in the Perception of Stress. *Language and Speech* 1 (2): 126-152.

Goldsmith, John A. 1990. *Autosegmental and Metrical Phonology*. Cambridge, MA: Basil Blackwell.

Gray, Michael D, Eleanor P. Stride, and Constantin C. Coussios. 2019. Snap, Crackle, and Pop: Theracoustic. *Acoustics Today* 15 (1): 19-27.

Guffey, Karen. 2002. *Spanish Syllable Structure*. New York: University Press of American.

Hasen, Colin H. 2001. Fundamentals of Acoustics. *Occupational Exposure to Noise: Evaluation, Prevention, and Control,* ed by Bernice Goelzer, Colin H. Hasen, and Gustav A. Sherndt. Dortmund, Germany: World Health Organization. Retrieved from: Retrieved from: http://www.who.int/occupational_health/publications/noise1.pdf on March 26, 2016.

Jones, Mark J. 2013. Phonetic Universals and Phonetic Variation. In *The Bloomsbury Companion to Phonetics*, ed. by. Mark J. Jones and Rachael-Anne Knight), pp. 140-154. New York: Bloomsbury.

Kochanski, G., E. Grabe, J. Coleman, and B. Rosner. 2005. Loudness Predicts Prominence: Fundamental Frequency Lends Little. *Journal of the Acoustical Society of America* 118 (2): 1038-1054.

Koffi, Ettien. 2019. A Comprehensive Review of F0 and its Various Correlations. *Linguistic Portfolios* 8: 2-24.

Ladefoged, Peter and Keith Johnson. 2015. *A Course in Phonetics.* Seventh Edition. Malden, MA: Cengage Learning.

Ladefoged, Peter. 2003. *Phonetic Data Analysis: An Introduction to Fieldwork and Instrumental Techniques*. Malden, MA: Blackwell Publishing.

Ladefoged, Peter. 1968. *A Phonetic Study of West African Languages: An Auditory-Instrumental Survey*. Second Edition. New York: Cambridge University Press.

Lewis, Rebecca M. Edwin W. Rubel, and Jennifer S. Stone. 2016. Regeneration of Auditory Hair Cells: A Potential Treatment for Hearing Loss on the Horizon. *Acoustics Today* 12 (2): 40-48.

Lieberman, Philip. 1960. Some Acoustic Correlates of Word Stress in American English. *Journal of the Acoustical Society of America* 12 (4): 451-454.

Lopez-Backstrom, Michel. 2018. *An Acoustic Phonetic Analysis of Northern Minnesota English Vowels*. St. Cloud: MA Thesis. St. Cloud State University, MN.

Loubeau, Alexandra and Juliet Page. 2018. Human Perception of Sonic Booms from Supersonic Aircraft. *Acoustic Today* 14 (3): 23-30.

Lubert Caroline P. 2018. From Sputnik to SpaceX: 60 Years of Rocket Launch Acoustics. *Acoustics Today* 14 (4): 38-46. Fletcher (1953:103, 104) provides noise levels in many places in New York City.

Malmberg, Bertil. 1963. *Phonetics: Physiological Phonetics, Experimental Phonetics, Evolutionary Phonetics, Phonemics*. New York: Dover Publications, Inc.

Morelli, Frida. 2003. The Relative Harmony of /s+Stop/ Onsets: Obstruents Clusters and the Sonority Sequencing Principle. In *The Syllable in Optimality Theory*, ed. by Caroline Féry and Ruben van de Vijver, pp. 356-371. New York: Cambridge University Press.

Murphy, Enda. What to do about Environmental Noise? *Acoustics Today* 13 (2):18-25.

O'Brien, Gabrielle E. The New Age of Sound: How Bell Telephone Laboratories and Leopold Stokowski Modernized Music. *Acoustics Today* 14 (2): 17-24.

Pearsons, Karl S. and Richard D. Horonjeff. 1982. Measurement of Speech Levels om the Presence of Noise. *NASA Report 3547*.

Pope, J. 1998. Analyzers. *Handbook of Acoustics*, pp. 1341-1354. Ed by. Malcom J. Crocker. New York: A Wiley-Interscience Publication, John Wiley & Sons, Inc.

Rabiner, Lawrence and Biing-Hwang Juang. 1993. *Fundamentals of Speech Recognition*. Englewood Cliffs, New Jersey: Prentice Hall.

Reetz, Henning and Allard Jongman. *Phonetics: Transcription, Production, Acoustics, and Perception*. Malden, MA: Wiley-Blackwell.

Roca, Iggy and Wyn Johnson. 1999. *A Course in Phonology*. Malden, MA: Blackwell Publishers, Inc.

Roy, Kenneth P. and Keely Siebein (2019). Satisfying Hunger, Thirst, and Acoustic Comfort in Restaurants, Diners and Bars … Is this an Oxymoron? *Acoustics Today* 2019 (15): 20-28.

Schnitta, Bonnie. 2016. Residential Quietude, the Top Luxury Requirement. *Acoustics Today* 12 (3): 49-56.

Scharf, Bertram. 1961. Complex Sounds and Critical Bands. *Psychological Bulletin* 58 (3):205-215.

Speaks, Charles E. 2005. *Introduction to Sound: Acoustics for the Hearing and Speech Sciences.* Third Edition. Clifton Park, NY: Thomson Delmar Learning.

Styler, Will. 2017. *Using Praat for Linguistic Research*. Online at: http://wstyler.ucsd.edu/praat/

Surucki, Milan I. 1964. The Distribution of Serbo-Croatian Consonants. *American Journal of Slavic and East European Languages* 8 (2):159-181.

Welmers, William E. 1973. *African Language Structure*. Los Angeles, CA: University of California Press.

Yost, William A. 2015. Psychoacoustics: A Brief Historical Overview. *Acoustics Today* 11 (3): 46-53.