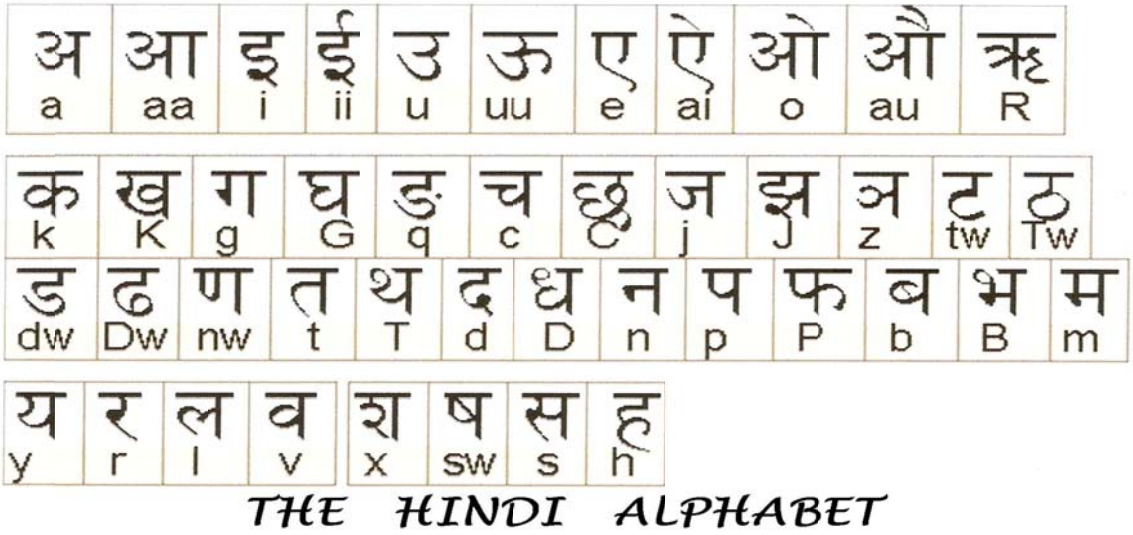
Physics 406, Spring 2015, Final semester project report

Tonal properties of Speech Unique to the Hindi Language

Max Bass*, Lawrence Hong Gang*
*University of Illinois at Urbana-Champaign

No discernable differences between native and non-native pronunciation the “nw” and “Dw” sounds could be found using the parameters of amplitude, phase, and frequency vs time of the first 6 fundamental tones in a speaker’s voice. However, other important aspects of Hindi speech and pronunciation were measured, clear marking differences between native and non-native speakers’ pronunciations. Thus, the goal of measuring unique, quantifiable aspects of speech in Hindi was achieved. Native speakers draw out pronunciation of sounds, and invoke exaggerated upward and some downward tonal inflections, compared with English speakers’ flat tones. Also, when pronouncing aspirated sounds, native speakers produce a clean, brief gap in sound, rather than trailing off in the gap.

Figure 1.



I. Intro:

The human ear receives and distinguishes sound in many ways. It gathers raw pitch, rhythm, and volume data when different cells in the cochlea in the inner ear are stimulated by different frequencies of sound. These cells each send signals to the Brain’s Auditory Cortex, in the brain’s temporal lobe, which then interprets the raw information, and helps to distinguish particular notes, rhythmic patterns, syllables, and words. In this investigation, raw vocal sound data is captured and analyzed, in order to distinguish any discernable differences between certain sounds unique to Hindi, and their analogous sounds in English.

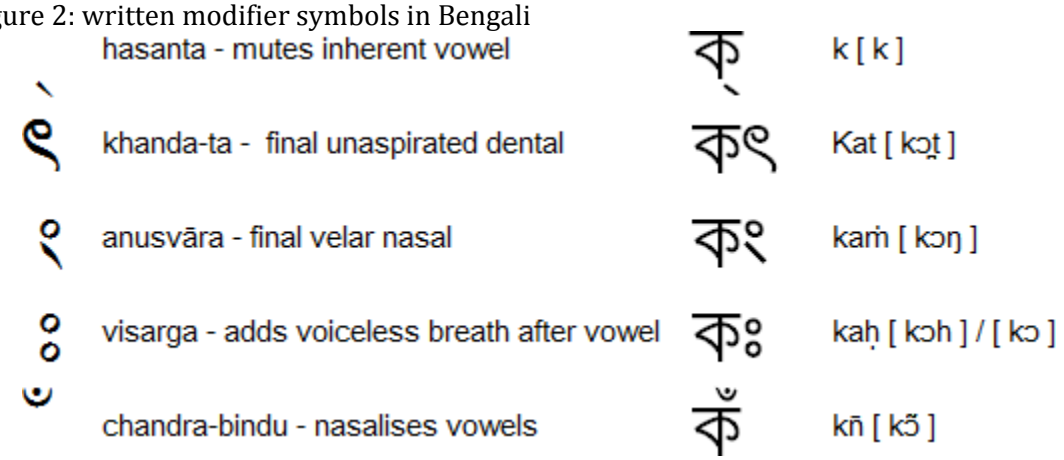
Investigation into acoustical properties of certain sounds present in Hindi, and their analogues in English could help English-speaking learners of Hindi, and other South Asian languages, to understand how to pronounce those sounds. Eventually, it also could have applications in speech recognition technology. Currently, no widely available, free, Hindi speech to text software exists.

II. Historical Background:

The language Hindi is spoken by about 42% of Indians, and along with English, is the most widely spoken language, among roughly 1600 state and regional languages across India [1]. Hindi is also mutually intelligible with Urdu, although it is written in a modified Arabic script, rather than in the Devnagri script (shown in figure 1), which comes from Sanskrit.

While Hindi and English are the national languages (note: not official languages) of India, many other Indian and South Asian languages share certain tonal properties. Bengali, Urdu, Tamil, and Telagu all contain modifiers to nasalize vowels (the Chandra-bindu symbol in figure 2, which is shared across Bengali, Hindi, Guajarati, Marathi, and other languages). Many South Asian languages contain aspirated consonant sounds. The English equivalent of this would be to put an "h" immediately after any consonant, and can be demonstrated in the letters "K", "G", "C", "J", "Tw", "Dw", "T", "D", "P", "B", and "h" shown in figure 1, above, and pronounced out loud in the audio files associated with this report.

Figure 2: written modifier symbols in Bengali

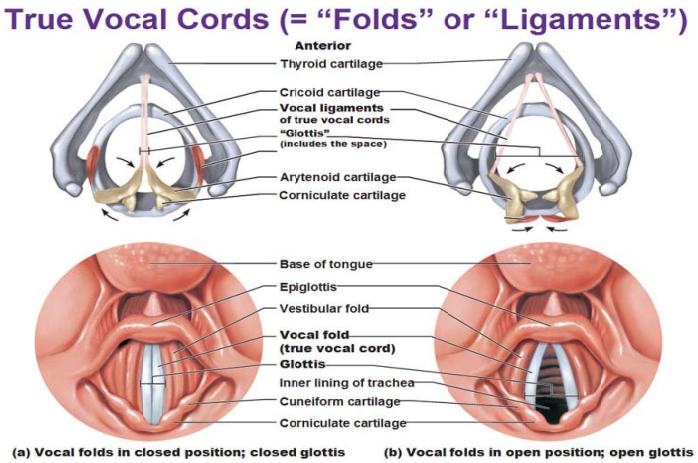


III. Technical Background:

The human voice is composed of fundamental frequencies imposed upon one another, from the various parts of the throat, and air passage vibrating. Each of these fundamental frequencies falls between 100 Hz and 5000 Hz, which is within the range of audible sound: 20Hz-20,000Hz. The human voice system interacts with the air in ways which can induce the sound wave of different volume, pitch, and vocal timbre. When producing sound through human voice system, the air flows from the lung through the larynx and reach the vocal cords. Vocal cords will remain open when not producing sound, and will come together to form a small gap to let the air from the lungs flow through when producing sound (seen in figure 3).

Sounds unique to Hindi come from a different shaping of the mouth when a sound is pronounced. Detailed description of Hindi consonant and vowel sound qualities can be found in the pronunciation table in the appendix.

Figure 3.



The pitch of sound exiting the vocal chords can be explained by the following equation:

$$y(x,t) = A\sin(kx \pm \omega t) = A\sin\left(\frac{2\pi}{\lambda}x \pm 2\pi f t\right) = A\sin k(x \pm vt)$$

When vocal cords are widely opened, the wavelengths of the fundamental tone produced is longer and for fixed sound wave speed in air v=343m/s, it has a lower frequency, thus lower pitch. More intense air flow induces higher amplitude sound waves, represented by the variable "A", and produces louder sound. More intense air flow often requires the vocal chords to open more widely, resulting in a lower pitched "shouting voice".

IV. Methods:

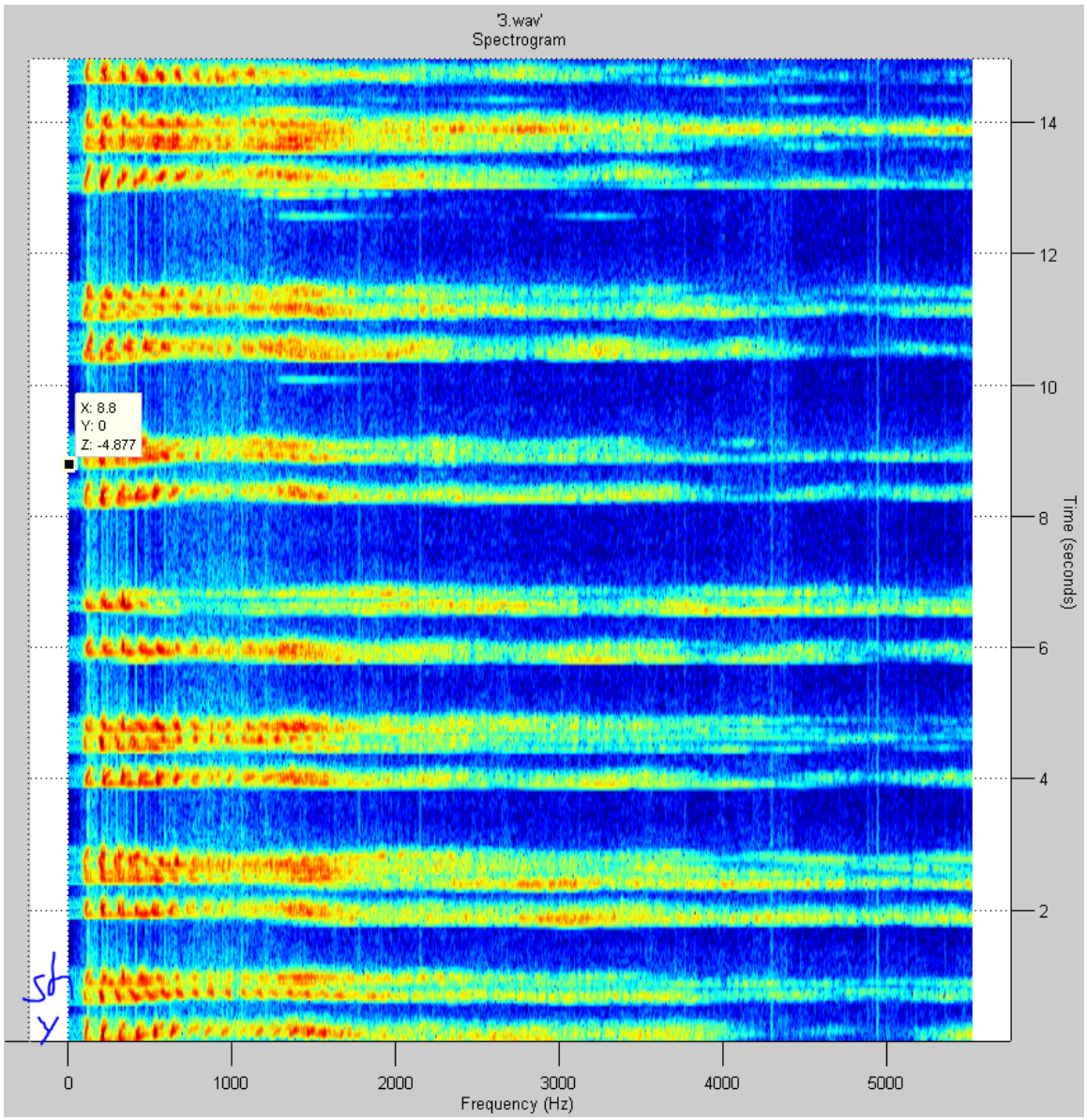
Different sounds, pronounced by different speakers are measured using the following parameters: the first 6 fundamental frequencies characteristic to each voice, measured in terms of amplitude, frequency, and phase, while different speakers are pronouncing different words and sounds.

Several Hindi speakers were recorded while pronouncing Hindi words. The speakers each pronounced at least one word for each letter in the Hindi alphabet, including some compound letters. Three speakers were recorded, including Max Bass (non-native), Pratik Naik and Ashutosh Katyal, who are each native speakers of Hindi and Marathi, and grew up in Mumbai. The recorded sound files were converted from their original mp4a format (associated with QuickTime audio recorder) to .wav files, so that the MATLAB program could analyze them.

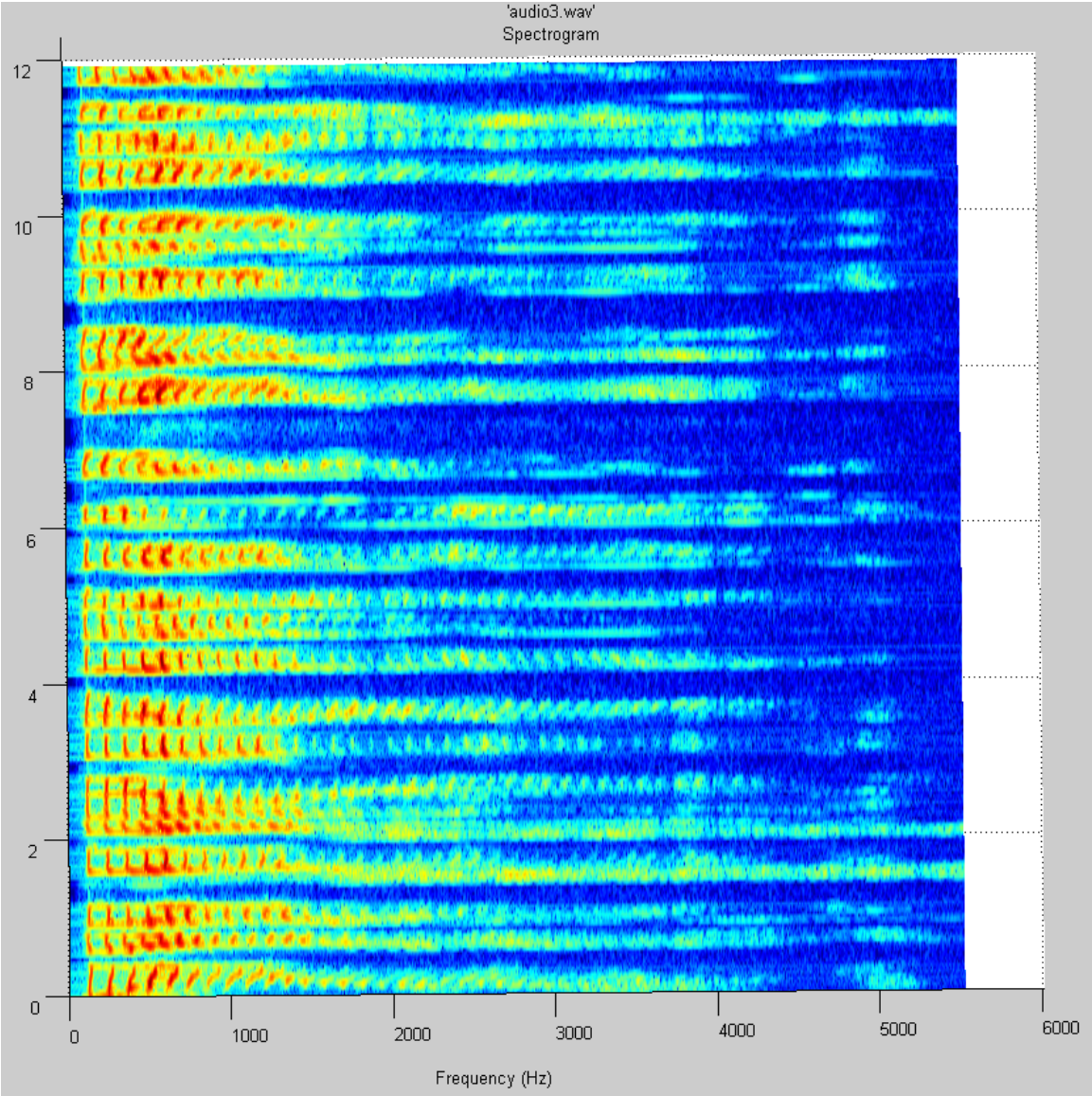
V. Results:

An immediately discernable difference in these two graphs is that the red areas, representing the high amplitude "fundamental tones" present in one's voice, exhibit different behavior in the native speakers' voices from the non-native speaker's voice. The red lines in both Pratik and Ashu's pronunciations are drawn out longer, and exhibit more a much more exaggerated change in pitch over the course of the pronunciation of any given sound or word.

Figures 4), time vs frequency vs amplitude (represented by color) plot for speaker Max, pronouncing sounds third consonant group, in order.



Figures 5), time vs frequency vs amplitude (represented by color) plot for speaker Pratik, pronouncing sounds third consonant group, in order.



The graphs in figure 7.) were all scaled as best as possible to match up individual sounds so that they could be compared. From these graphs can be discerned a noticeable difference in the pronunciation of the aspirated syllables. The aspirated syllables “kh” and “gh” appear on the second and fourth lines from the bottom. In the native speakers’ graphs, there exists a clear break across all fundamental tones, in the middle of pronunciation of an aspirated syllable. In the non native speaker’s graph, the break is less clear.

Figures 6), time vs frequency vs amplitude (represented by color) plot for speaker Ashu, pronouncing sounds third consonant group, in order.

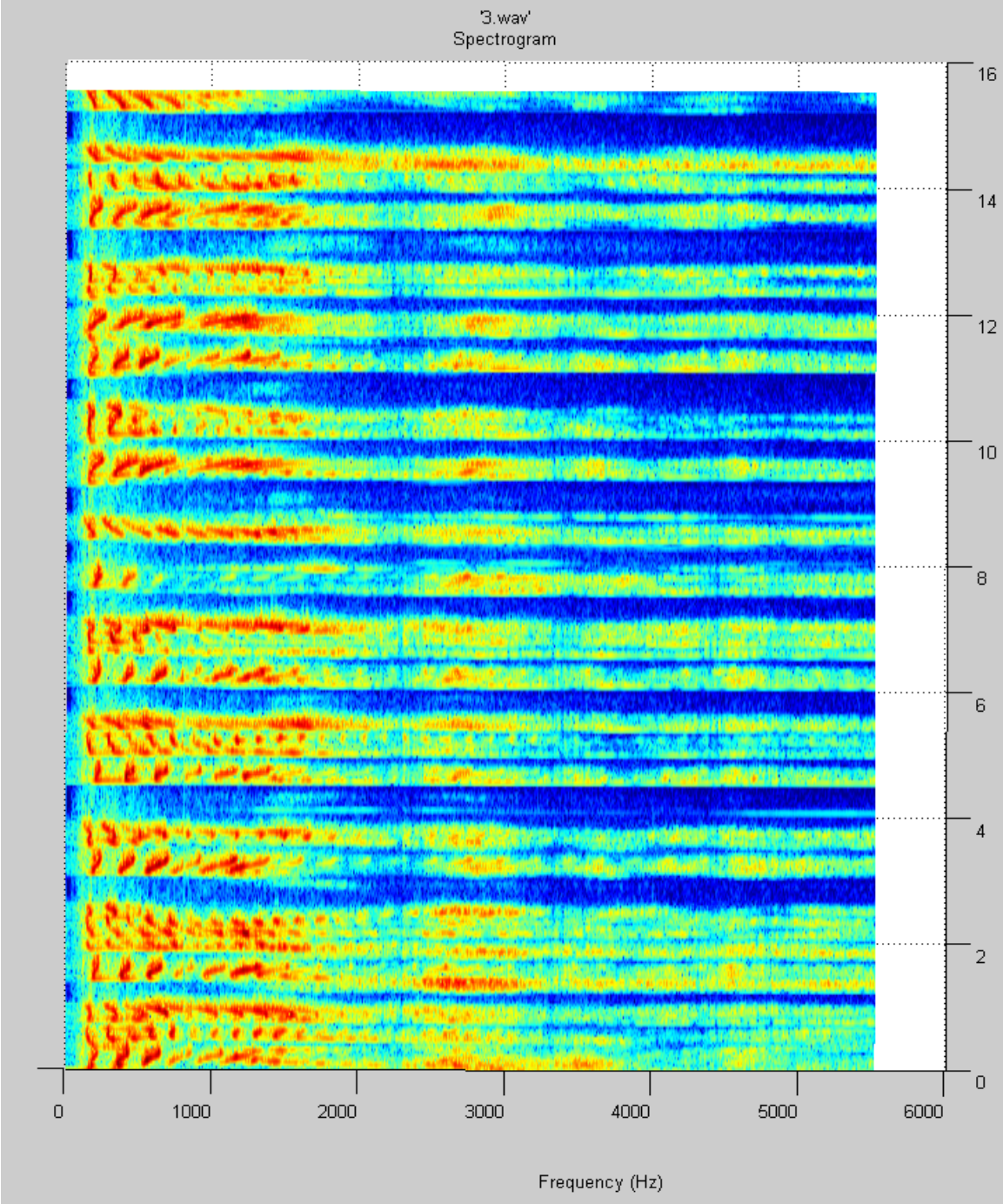
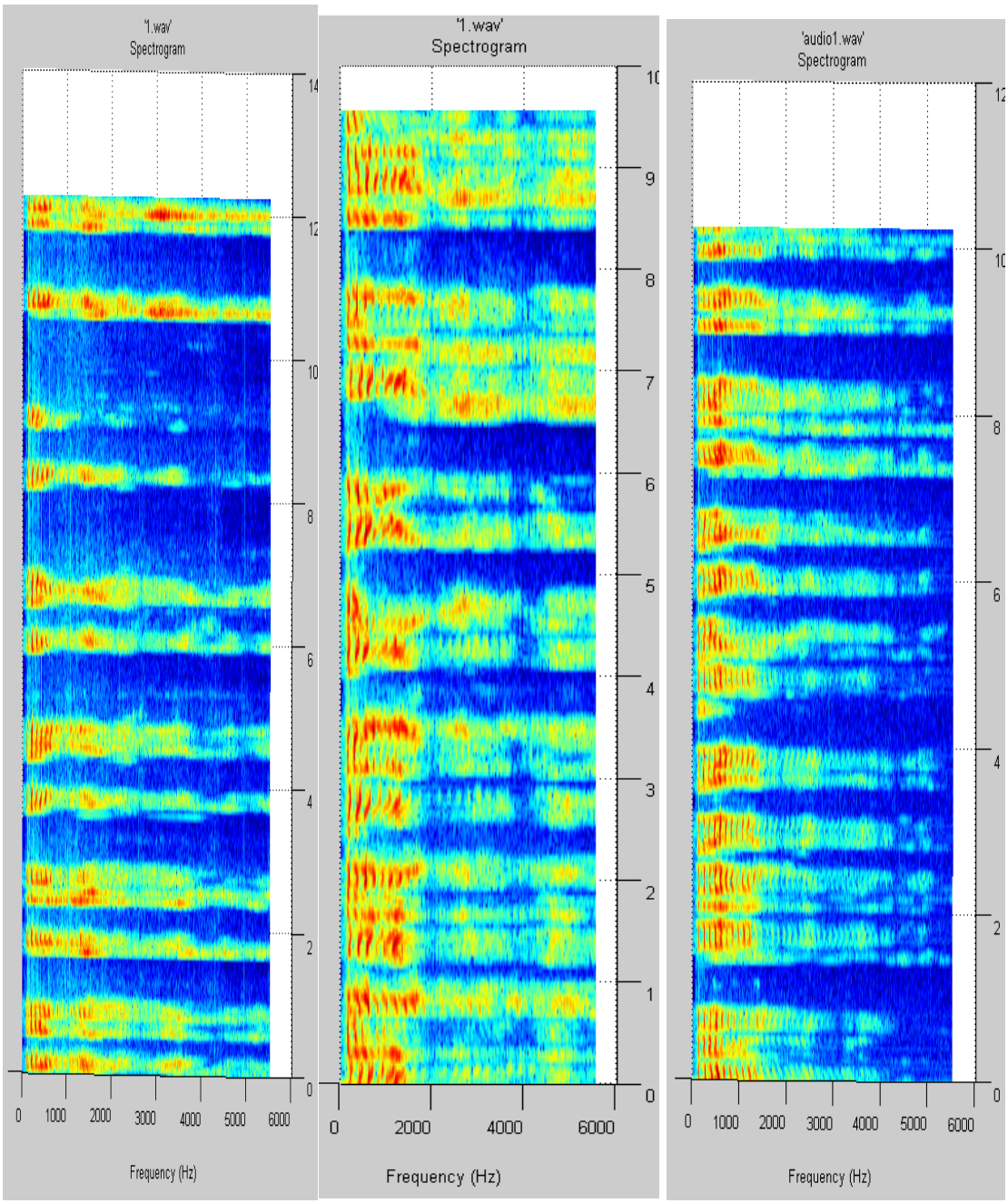


Figure 7.) Consonant group 1 being pronounced by Max, Ashu, and Pratik, respectively, measuring time vs frequency vs amplitude (represented by color)



VI. Analysis:

The greater change in pitch, and more drawn out pronunciation of any given sound by Hindi speakers demonstrates a speech pattern unique to Hindi, and Indian languages, and hence characteristic of speech in Hindi: frequent exaggerated upward and downward inflections in the voice, with the number of upward inflections dominating. While this result doesn't necessarily pertain to the pronunciation of any particular consonant unique to Hindi, it is an essential aspect of communication in Hindi. That this property of speech could be recognized, and

even quantified using a sound analysis program that looks at raw sound data indicates that it is a fundamental, and measurable difference between English and Hindi pronunciation.

That the non native speaker's pronunciation of aspirated syllables involved a less clean break in sound among all fundamental tones indicates a difference between native and non native pronunciation. Hence, it indicates a fundamental property of Hindi speech. It indicates that the non native speaker trails off in pronunciation of aspirated syllables, where he should be more cleanly and briefly "cutting the sound off".

In this project, error could manifest as statistical error, as well as experimental, procedural error. Statistical error could come from the small data sample size: one non-native speaker, and two native speakers. A more rigorous study would compare data from each group, ideally at least 20 of each. Experimental error exists in that a laptop microphone was used to gather data, whereas a higher quality microphone could have been used.

VII. Conclusions

This project's original goal was to investigate the nature of two particularly difficult sounds unique to Hindi: "nw" and "Dw", as indicated in the alphabet in figure 1. While no discernable differences between native and non-native pronunciation of those two sounds could be found using the available parameters, in the time available for this project, other important aspects of Hindi speech and pronunciation were measured, as clear differences between native and non-native speakers' pronunciations. Thus, the goal of measuring unique, quantifiable aspects of speech in Hindi was achieved. Namely, native speakers draw out pronunciation of sounds, and invoke many upward and some downward tonal inflections. Also, when pronouncing aspirated sounds, native speakers produce a clean, brief gap in sound, rather than trailing off in the gap.

Much more data is was gathered, and will be made accessible on the class shared folders. Potentially, future groups could use this data to analyze differences in voice characteristics of people from various ethnic, geographic, and linguistic backgrounds.

VIII. Acknowledgements:

The authors would like to acknowledge Ashu Katyal, and Pratik Naik for volunteering their voices. Additionally, Nicole Cox, another non-native Hindi language learner volunteered her voice, and the sound data was collected for it, but the analysis of her voice couldn't be matched against a female native Hindi speaker, so it is omitted from the analysis in this paper. Thank you to Prof. Errede for allowing us to use the wave analysis program that you made in MatLab, in order to conduct this analysis at all. Thank you to the TA, Matt Zeimann, for providing help with the wave analysis program, and letting me into the building to work on mindless data collection outside of class time.

IX. Sources:

Images and figures:

1.) Hindi Alphabet:
http://www.ac-grenoble.fr/college/le-revard.gresy/Shree,%20from%20India/Hindi%20alphabet.jpg

2.) Bengali modifier symbols:
http://www.omniglot.com/writing/bengali.htm

3.) Vocal chords image
http://antranik.org/wp-content/uploads/2011/12/true-vocal-cords-vocal-ligaments-vestibular-fold-false-vocal-fold-glottis-closed-position-open-position.jpg?56505f

Expert reference:

[1] Prof. Mithilesh Mishra of the Hindi Studies dept. at UIUC

X. Appendix:

The order of the words and sounds pronounced by each speaker is laid out in the table below. The following table makes use of all phonetic sounds of hindi (with the exception of two of the nasal sounds), and was arranged by the authors.

Phonetic pronunciation written in latin alphabet, along with english translation on the right. On the left, there is a letter, and a word containing (usually starting with) that letter, to show the sound)

Vowel Sounds	
अ - अनार आ - आम इ - इमली ई - ईख उ - उल्लू ऊ - ऊं ए - एक ऐ - ऐनक ओ औ - औरत अं - अंगूर ऋ- संस्कृत, कृपया	uh-unar (pomegranate) ah-aam (mango) i - imli (tamarind) *short "i" sound, as in "ich"* ee - eekh (reed) u - ulloo (owl) *short "u" sound, as in "cook"* oo - oon (yarn) eh - ek (one) ai - ainak (glasses) oh aw - aurat (woman) ung - ungur (grapes) rri - sanskrit, krripya (sanskrit language, please) *this is considered a vowel in Hindi*

Consonant group 1	
क - कबूतर ख - अख़बार ग - गमला घ - घडी क्ष - कक्षा, लक्ष्मी	ka- kabutar (dove/pigeon) kha - akhbar (Newspaper) *Aspirated k sound* ga - gamla (flowerpot) gha - ghadi (wrist watch) *Aspirated g sound* ksha - kaksha, lakshmi (Class, Lakshmi)
Consonant group 2	
च - चावल छ - छतरी ज - जाना झ - झंडा ज्ञ - ज्ञानी	cha - chaval (rice) chha - chatari (umbrella) *Aspirated ch sound* ja - jana (go) jha - jhada (flag) *Aspirated j sound* gya - gyani (wiseman)
Consonant group 3	
य - यात्रा श - शाकाहारी ह - हरिण ट - टमाटर ठ - ठीक ड - डमरू ढ़ - ढकना ण - आरक्षण	ya - yatra (travel) sha - shakahari (vegetarian) ha - harin (male deer (stag)) Ta - Tamatar (tomatoes) *here, the "t" sound is produced by the tongue hitting the front of the hard palate of the mouth, rather than the edge of the teeth. It is a less common sound in Hindi than the analogous "t" sound in the next consonant group, but is more similar to how the "t" is pronounced in english.* Tha - Thik (ok) *Aspirated "T" sound* Da - Damaru (drum) *This "d" sound is the same as the "d sound produced in english, as opposed to the "d sound in the next consonant group.* Dha - Dhakna (hood/cover) *This is a very difficult sound for english speakers to pronounce. In practice, it ends up sounding similar to the beginning of a "rolling r" sound.* Na - arakshaN (reservation) *This sound is also very difficult for enlish speakers to pronounce properly. It is very similar to the "n" sound in english, but the tongue is farther back on the roof of the mouth, and a slightly different part of the nasal cavity is used to produce this sound.*
Consonant group 4	
र - रस्सी ष - कष्ट त - तरबूज थ - थन	ra - rassi (rope) sha - kasht (trouble) *this is the same sound as the other sh sound, this is a redundant letter* ta - tarabuj (watermellon) *this "t" sound is produced by a burst of breath and a quick release of the tongue from the tip of the upper teeth outward. This pronunciation of "t" naturally occurs with some speakers of spanish, and other languages, although it isn't specifically called for, as is the case with Hindi*

द - दवात ध - धनुष न - नल त्र - त्रिशूल	tha - thun (udder) *Aspirated "t" sound* da - davaat (ink stand) *this "d" sound is closer to "th", as in the word "the"* dha - dhanush (longbow) *Aspirated "d" sound* na - nal (tap) *n sound common in english* tra - trishul (trident)
Consonant group 5	
ल - लम्बा स - सेब प - पतला फ - फल ब - बकरी भ - भालू म - मछली व - वन	la - lamba (long) sa - seb (apple) pa - patla (diluted/skinny) fa - fal (fruit) ba - bakri (goat) bha - bhaloo (bear) *Aspirated "b" sound* ma - machli (fish) va/wa - van (forrest) *the v and w sounds in Hindi are interchangeable*