



*Journal of Advances and  
Scholarly Researches in  
Allied Education*

*Vol. VII, Issue No. XIV,  
April-2014, ISSN 2230-7540*

## **SPEECH EMOTION ANALYSIS: THE PRODUCTION-PERSPECTIVE**

AN  
INTERNATIONALLY  
INDEXED PEER  
REVIEWED &  
REFEREED JOURNAL

# Speech Emotion Analysis: The Production-Perspective

Rohit Parashar

Research Scholar, Sainath University, Ranchi, Jharkhand

**Abstract** – Automatic recognition of human emotion (anger) in speech aims at recognizing the underlying emotional state of a speaker from the speech signal. The area has received rapidly increasing research interest over the past few years. However, designing powerful spectral features for high-performance speech emotion recognition (SER) remains an open challenge. Our purpose was to investigate whether emotion (anger), as perceived by a panel of listeners, were observable in various acoustic cues of the speech signal. The cues were chosen, by examining earlier studies on the same subject and they were: the syllable rate, the minimum, maximum, median and mean of the pitch, the amplitude and the first six formants, % jitter and %shimmer. Application is first trained on person voice will provide signal if detected as angry, and a person, who is signaled that his/her manner of speaking is classified as angry, becomes aware of his mental state and could regulate his way of expressing his thoughts. It is of great advantage in this situation if the corrective comes from a machine that does not play a part in the situation and shows no emotions in itself.

**Keywords:** Recognition, Emotion, Interacts.

## INTRODUCTION

Speech emotion analysis refers to the use of various methods to analyze vocal behavior as a marker of affect, focusing on the nonverbal aspects of speech. The basic assumption is that there is a set of objectively measurable voice parameters that reflect the affective state a person is currently experiencing. This assumption appears reasonable given that most affective states involve physiological reactions, which in turn modify different aspects of the voice production process. For example, the sympathetic arousal associated with an anger state often produce changes in respiration and an increase in muscle tension, which influence the vibration of the vocal folds and vocal tract shape, affecting the acoustic characteristics of the speech, which in turn can be used by the listener to infer the respective state.

## 3. BASIC CONCEPTS OF SOUND PRODUCTION

### Speech Organs

Produce the many sounds needed for language. Organs used include the lips, teeth, tongue, alveolar ridge, hard palate, velum (soft palate), uvula and glottis.

Speech organs— or *articulators*—are of two types:

### Passive articulators:

Passive articulators remain static during the articulation of sound. Upper lips, teeth, alveolar ridge, hard palate, soft palate, uvula, and pharynx wall are passive articulators.

### Active articulators:

Move relative to these passive articulators to produce various speech sounds, in different manners. The most important active articulator is the tongue. The lower lip and glottis are other active articulators

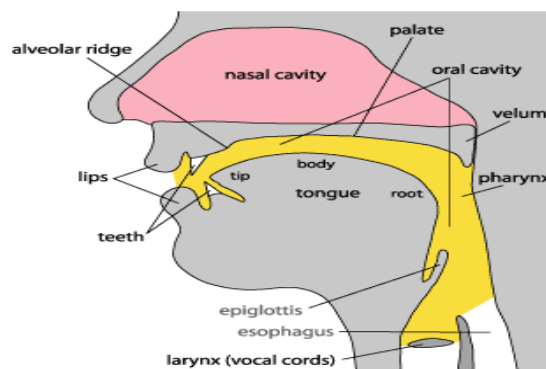


Figure 1: Speech Articulators

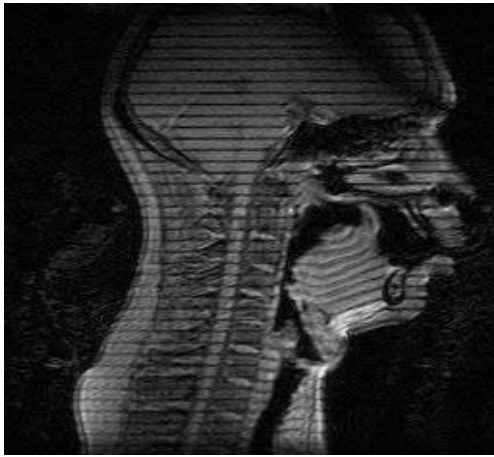


Figure 2: Midsagittal tMRI Slice of Head

### Vocal Folds

The vocal folds, also known popularly as vocal cords, are composed of twin infolding of mucous membrane stretched horizontally across the larynx. They vibrate, modulating the flow of air being expelled from the lungs during phonation. Another name for the airway at the level of the vocal cords is the glottis, and the opening between the cords is called the glottic chink. The size of the glottic chink is important in respiration and phonation.



Figure 3: Image of Normal Vocal Cord, courtesy of the Milton J. Dance

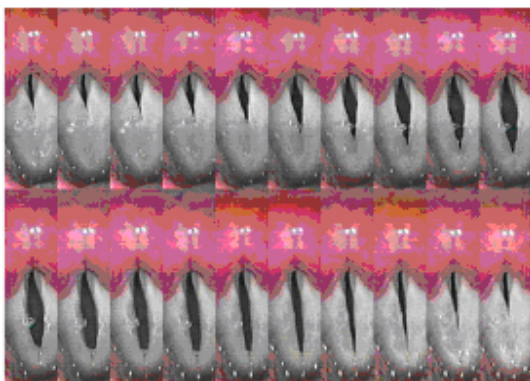


Figure 4: Representative Set of Images from Stroboscopy depicting "one" vibratory cycle

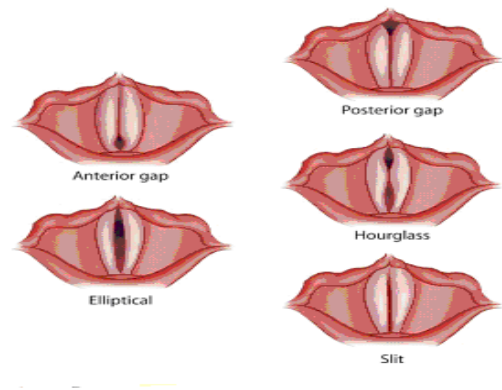


Figure 5: Different Vocal Fold Closure Patterns

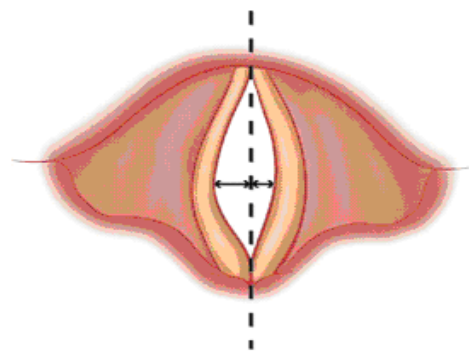


Figure 6: Amplitude of Vocal Fold

## 4. ARTICULATION

Articulation refers to the production of the speech sounds. Accurate articulation involves precise movement of the articulators including the tongue, lips, alveolar ridge, velum, and jaw coordinated with correct air flow and voicing

### Place of articulation

- Refers to the relative positions of the lips, teeth and tongue.
- There are six distinct types of classification bilabial, labiodental, interdental, alveolar, alveo-palatal, and velar.
- The six places of articulation describe the parts of the vocal tract which are responsible for the obstruction of the air flow from the lungs the degree of obstruction the airstream incurs must also be considered.

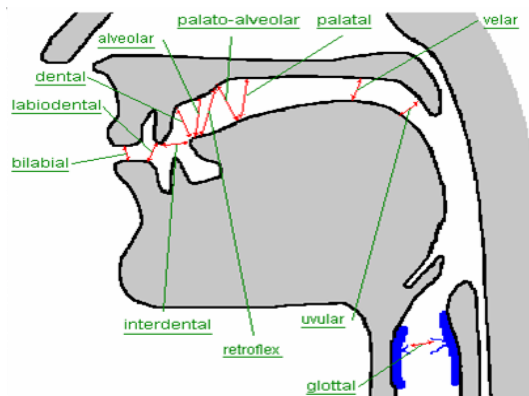


Figure 7: The Places of Articulation

## 5. PRODUCTION OF SPEECH SOUNDS

The act of speech involves three major anatomical subsystems:

1. Respiratory system including the lungs, rib cage, and diaphragm;
2. Phonatory system which includes the larynx
3. Articulatory features the lips, teeth, tongue, and jaw.

## 6. EFFECTS OF EMOTION ON HUMAN VOCAL SYSTEM

There are many situations where people perceive stress and emotion. These include heavy workload, adverse environment and social problems. Stress has impact on the body as well as on the mind of the person and this in turn affects the vocal system. There are physiological reasons for the acoustic cues changing whenever a speaker changes his or her emotional state. A Study done by Kienast at the University of Berlin examined the spectral and segmental changes caused by the articulatory behaviour of a person feeling an emotion (Kienast et al., 2000). The effects of human emotion on vocal system and variation of acoustic characteristics are analyzed.

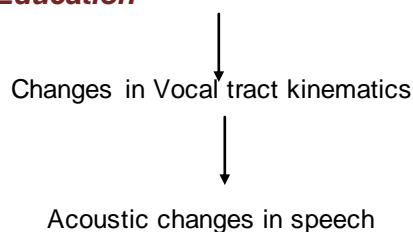
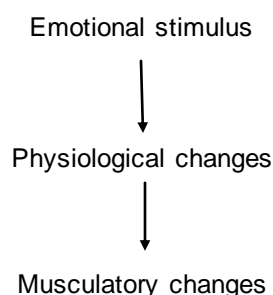


Figure 8: Effect of Emotion on Speech

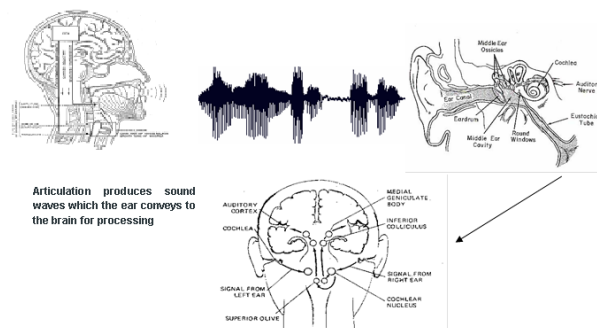


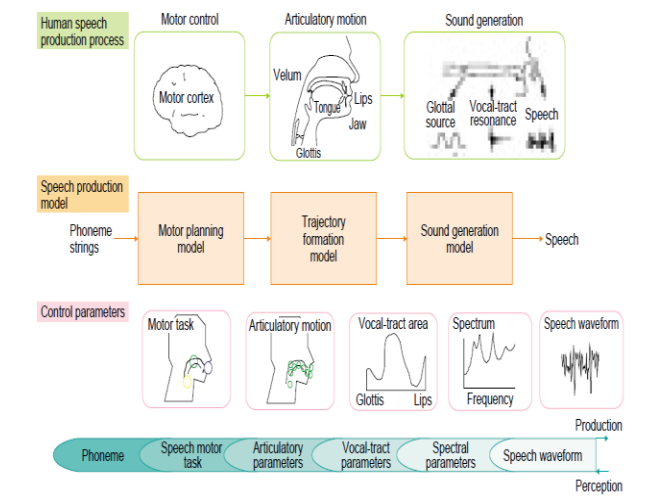
Figure 8: Speech Production and Perception

When people are undergoing under any kind of negative emotion such as anger or stress, their bodily resources are automatically changed to prepare an attack or to run away from danger. If the situation persists, considerable strain may be placed on the body and affects a person's ability to perform including producing speech. Stress is observed even in positively toned emotions. For example, Anger, Anxiety, Guilt and Sadness are regarded as stressed emotions

Positive emotions of Joy, Pride and Love are also frequently associated with stress. For example, when people are in happy mood, they may fear that the favorable conditions provoking their happiness will end. The research studies that have emphasized especially only on psychological, biological, and linguistic aspects of several emotional states can be found in. From the psychological perspective, of particular interest is the *cause-and-effect* of emotion. The activation-evaluation space provides a simple approach in understanding emotions. In a nutshell, it considers the stimulus that excites the emotion, the cognition ability of the agent to appraise the nature of the stimulus and subsequently his/her mental and physical responses to the stimuli. The mental response is in the form of emotional state.

Thus, emotion has a broad sense and a narrow sense effect. The broad sense reflects the underlying long-term emotion and the narrow sense refers to the short-term excitation of the mind that prompts people to action. In automatic recognition of emotion, a machine will not distinguish if the emotional states are due to 15 long-term or short-term effects so long as it is reflected in the speech or facial expression. From the perspective of physiology in the production of speech. As a result, heart rate and blood pressure

increase, the mouth becomes dry and there are occasional muscle tremors. On the other hand, with the arousal of the parasympathetic nervous system, as with Sadness, heart rate and blood pressure decrease and salivation increases, producing slow speech.



**Figure 9: Speech Production Process and the Model.**

The corresponding effects on speech of such physiological changes thus show up vocal system modifications and affect the quality and characteristics of the utterances. The acoustic characteristics that are altered during stressed and emotional speech production are studied in the following section.

## 7. CONCLUSIONS

In this paper, a system for anger recognition and classification is proposed. Evaluations that concentrate in identifying the effect of anger on vocal system are carried out. It is found that the characteristics of speech utterances are altered when producing stress or emotion. From this knowledge, the best acoustic features that are important for stress and emotion (Anger) detection are selected from several traditional features. The features such as pitch, amplitude, spectral distribution and speaking rate parameters function as basic acoustic parameters to characterize emotion.

## 8. REFERENCES

- ❖ Lazarus, R.S., 2010. "Emotion Adaptation". In Annual review of psychology, New York: Oxford Univ. Press. (pp. 1-21).
- ❖ McGilloway, S., R. Cowie and E. Douglas-Cowie, 1995. "Prosodic Signs of Emotion in Speech: Preliminary Results from A New Technique for Automatic Statistical Analysis", in Proc. XIIIth Int. Congr. Phonetic Sciences, Vol. 1. Stockholm, Sweden, pp. 250-253, 171.
- ❖ Murray, I.R. and J.L. Arnott, 1993. "Toward the Simulation of Emotion in Synthetic Speech: A

Review of the Literature on Human Vocal Emotion", Vol. 93, no. 2, pp. 1097-1108.

- ❖ O'Connor, J.D. and G.F. Arnold, 1973. "Intonation of Colloquial English", 2nd ed. London, UK: Longman.
- ❖ Oatley, K. and P. Johnson-Laird, 1995. "Communicative Theory of Emotions: Empirical Test, Mental Models & Implications for Social Interaction". In Goals and Affected by L. Martin and A. Tessler, Eds. Hillsdale, NJ: Erlbaum, 1995.167.
- ❖ Oster, A. and A. Risberg, 1986. "The Identification of the Mood of A Speaker by Hearing Impaired Listeners, Speech Transmission Lab". Quarterly Progress Status Report 4, Stockholm, pp. 79-90, 1986.
- ❖ Oatley, K. and J.M. Jenkins, 1996. "Understanding Emotions. Oxford", UK: Blackwell.
- ❖ Patrik N. Juslin, Klaus R. Scherer, (2008). "Speech emotion analysis", Scholarpedia, 3(10):4240.