THEORY OF SPEECH EMOTION ANALYSIS – THE PRODUCTION-PERSPECTIVE

Kaushal Kumar	Dr.K.P.Yadav
Research Scholar	Supervisor
Sunrise University, Alwar	Sunrise University, Alwar
Rajasthan	Rajasthan

ABSTRACT

In the growing range of interactive interfaces, the research of emotional voice is still at an early stage, not to mention a paucity of literatures on real applications. To make the mainstream use of speech emotion recognition a reality, the industry must deliver robust and high-recognition accuracy close to human-like. Although progress has been impressive, there are still several hurdles that automatic emotion recognition from speech. Here we are going to put eye on technical challenges in automatic emotion recognition through speech.

KEY WORDS: recognition, emotion, interact.

1.   INTRODUCTION:

Speech is the most basic, common and efficient form of communication method for people to interact with each other. People are comfortable with speech therefore persons would also like to interact with computers via speech, rather than using primitive interfaces such as keyboards and pointing devices. Psychological research on human emotions claimed that there are six basic expressions of emotions that are universally displayed and expressed: happiness, anger, sadness, surprise, disgust and fear. This implies that these basic emotions are probably displayed and recognized cross culturally (i.e. language-independently), and speaker-independently. Tickle also found that there was little difference for people detecting emotions of nonsense speech spoken in his native language or a foreign language. This motivates us that it is possible to do speaker-dependent emotion detection based only on acoustic information, that is to say, there is no speech recognizer and linguistic or semantic information available.

2. 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 owner lip and glottis are other active articulators
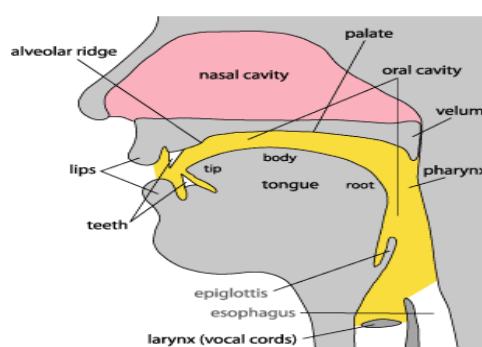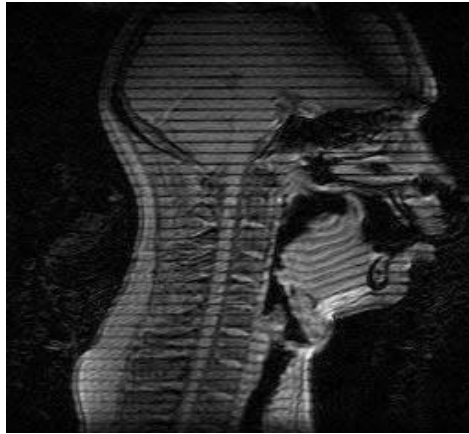

Figure 1: Speech Articulators

Figure 2: Midsagittal MRI Slice of Head

3.   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.

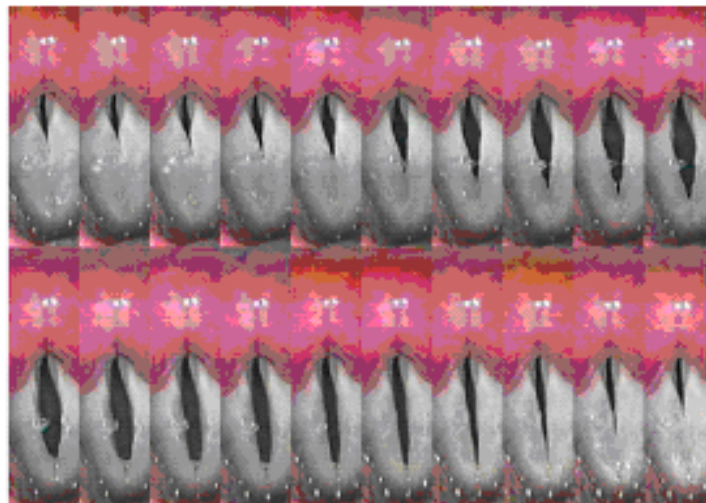Figure 3: Image of Normal Vocal Cord, courtesy of the Milton J. Dance

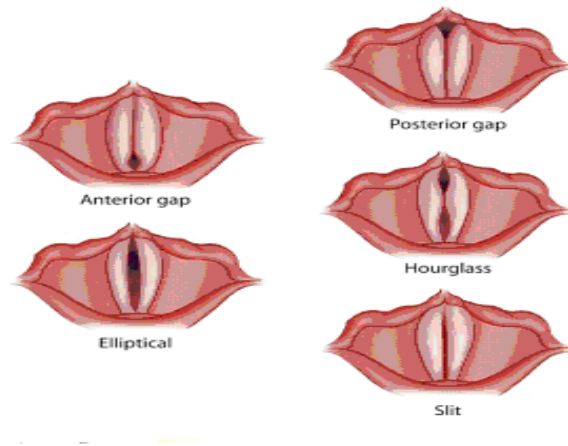Figure 4: Representative Set of Images from Stroboscopy depicting "one" vibratory cycle

Figure 5: Different Vocal Fold Closure Patterns

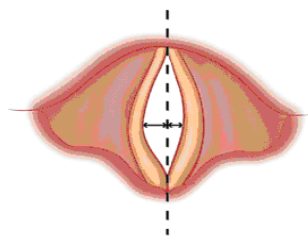Figure 6: Amplitude of Vocal Fold

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.
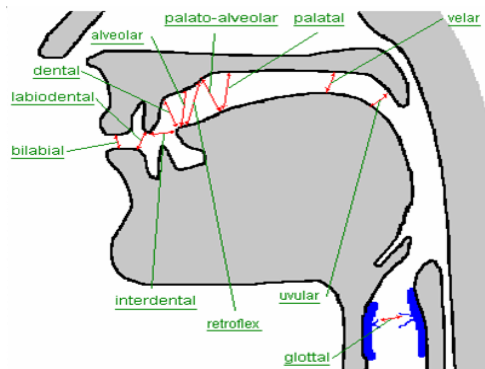

Figure 7: The Places of Articulation

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.

Emotional stimulus

↓

Physiological changes

↓

Musculatory changes

↓

Changes in Vocal tract kinematics

↓

Acoustic changes in speech

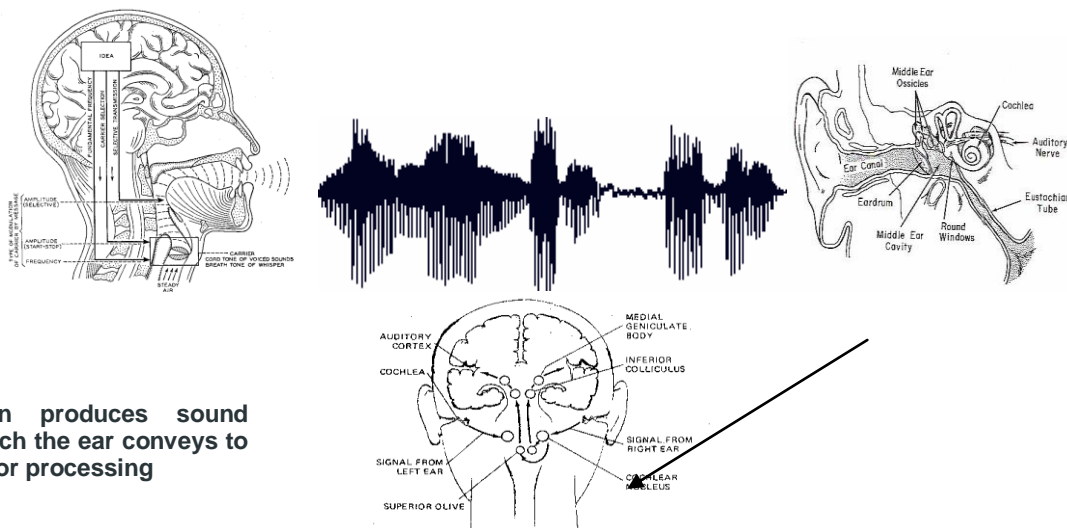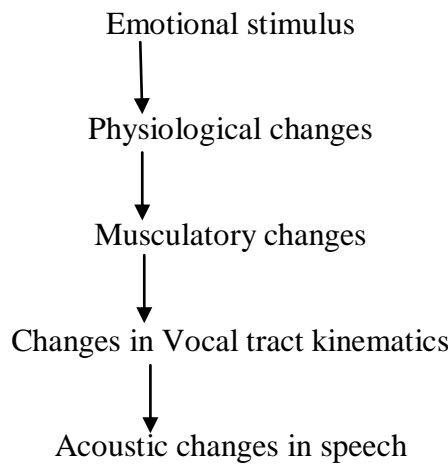**Articulation produces sound waves which the ear conveys to the brain for processing**

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.

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

1. Lazarus, R.S., 2010. "Emotion Adaptation". In Annual review of psychology, New York: Oxford Univ. Press. (pp. 1-21).
2. 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.
3. 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.
4. O'Connor, J.D. and G.F. Arnold, 1973. "Intonation of Colloquial English", 2nd ed. London, UK: Longman.
5. 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.
6. 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.
7. Otaley, K. and J.M. Jenkis, 1996. " Understanding Emotions. Oxford", UK: Blackwell.
8. Patrik N. Juslin, Klaus R. Scherer, (2008). "Speech emotion analysis", Scholarpedia, 3(10):4240.