Volume 3, Issue 12, December 2013 ISSN: 2277 128X
International Journal of Advanced Research in Computer Science and Software Engineering
Research Paper
Available online at: www.ijarcsse.com

Effect of Unvoiced Consonants on EMG Signals Generated in Zygomaticus Muscles

Himanshu Raina[1], Randhir Singh[2], Jang Bahadur Singh[3], Parveen Lehana[4*]
M.Tech, Department ECE, Shri Sai College of Engineering and Technology, Pathankot, India [1]
HOD, Dept ECE, Shri Sai College of Engineering and Technology, Pathankot, India [2]
Scholar, Dept of Physics and Electronics, University of Jammu, India [3]
Associate Professor, Dept of Physics and Electronics, University of Jammu, India [4*]

Abstract— ***Speech is one of the most complex non stationery signals and is also the easiest way of communicating between humans. Speech signals can be affected by the behaviour and emotions of speakers. EMG is one of the biomedical signals that measures electrical current generated in muscles during its contraction representing neuromuscular activities. Contraction and relaxation of the muscles are controlled by the nervous system. Hence, the EMG signal is a complicated signal and depends upon the anatomical and physiological properties of muscles. These signals from specific facial muscles are recorded for speech recognition and system automation. EMG signals are generally recorded using small surface electrodes placed near to each other. EMG activity is frequently recorded from specific muscles and plays a prominent role in the expression of elementary emotions and speech generation. The present research paper investigates the EMG patterns generated during the utterance of the unvoiced consonants. Six subjects in the age of 20-25 years were taken (three males and three females). Thirty eight vowel-consonant-vowel (VCV) syllables in Hindi were recorded along with the corresponding facial EMG signal. For each speaker, the means of log-spectral-distances (LSD) between the EMG signal of the VCVs and the reference EMG signal were computed. Analysis of the spectrograms and LSD showed that the EMD signals generated in the muscle vary with the subject and the VCV. Hence, for automatic decoding of the EMG signals, the system should be trained using both the variants.***

Keywords— ***Speech, Speech generation, EMG signals, Unvoiced consonants, Segmentation.***

I. INTRODUCTION

Affective state of an individual is revealed by the face, which is considered richest source of information for emotion expression. Affective facial expressions can be quantitatively analyzed for scientific purposes. Electromyographic (EMG) signals from specific facial muscles are recorded for speech recognition and system automation [1], [2]. EMG signals are generally recorded using small surface electrodes placed near to each other. EMG activity plays a prominent role in the expression of elementary emotions and speech generation. Fig. 1 shows some important locations for recording facial EMG [3]. Although affective facial EMG responses may show bilateral differences in individual subjects, group results generally do not show systematic differences between both sides of the face during spontaneous emotional expressions [3], [4].

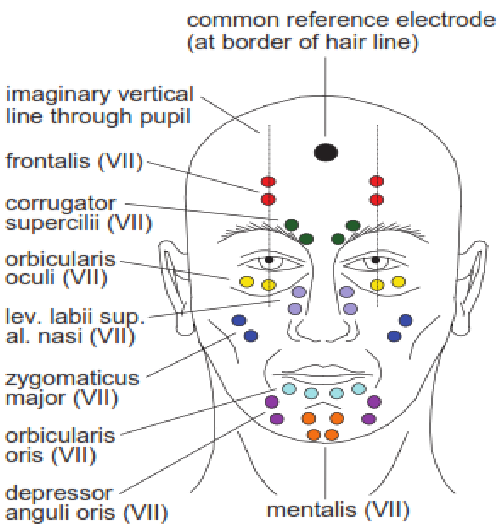

Fig. 1 Some important locations of recording spots on facial muscles

EMG signals can be recorded using three stage procedure; electrode selection, placement on the facial muscles, and signal conditioning. Every stage must be handled carefully to eliminate noise and possible measurement errors. The electrode is an important interface, which can pick up the bio-potential or the current generated in the human body. In order to record EMG signal, surface and needle electrodes are usually used. Needle electrodes can obtain good muscle selectivity and signal to noise ratio. Besides the type of the electrodes, its material also plays an important role in EMG recording. Material of electrode can greatly influence the half-cell potential leading to useless and misguiding interpretation. In EMG measurement, all recording electrodes should be made of the same material to minimize half-cell potential differences [5]. Facial EMG signal recorded during the human speech production mechanism may be implemented for speech recognition and also for designing an automated speech recognition based systems.

In this paper the Effect of unvoiced consonants on EMG signals generated over zygomaticus muscle is investigated. The basics of human speech production are presented in Section II. The methodology of the investigations is given in Section III. The results and conclusions are presented in Section IV.

II. Human Speech Production

The most complex non-stationary signal but the easiest way of communication between humans is speech [6], [7]. For many decades researchers are developing artificial models of human vocal tract [8-12]. In this section two models are presented. Fig. 2 shows the recently developed artificial human speech generation model by Anton [4] known after his name as Anton. It is an animatronics model of a human tongue and vocal tract. The main aim of Anton is to produce speech sounds using auditory feedback as producing specific speech. In this model movable tongue and jaw are actuated by servo motors. The hyoid bones are fixed in its position and the tongue is attached to the jaw. The jaw is connected to the skull by two joint capsules enclosing the mandibularcondyles.

Arai [13] designed a physical model based on electro larynx to produce vowels. This model employs acoustic theories, source filter theory, and perturbation theory. Fig. 3 shows the physical plate model of the human vocal tract. In this model every plate has a hole in the centre. When these plates are placed alongside, holes form an acoustic tube and the cross-section area changes in stepsr. When a sound source is applied to either one end of this model, vowels are produced from the other side.

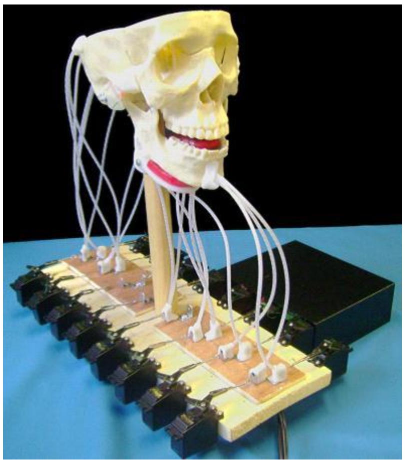

Fig. 2 The animatronics tongue and vocal tract model

Fig. 3 Physical plate model of the human vocal tract

III. Methodology

For investigating facial EMG patterns generated by uttering unvoiced consonants in VCV syllables, experiments were conducted with the six subjects (three males and three females) having age between 20-25 years. The speech signals and the corresponding EMG signals were recorded using a data acquisition system at the sampling frequency of 16 kHz and 16-bit quantization. Electrodes were placed at zygomaticus minor, zygomaticus major, and mentalispoints on the face. Block diagram of the experimental setup for recording is shown in Fig. 4. The recorded signals were segmented and labeled manually into separate files for each of the VCVs. The signals were analyzed using time-domain patterns, spectrograms, and mean log-spectral-distances. For computing LSD, the first VCV of each subject was taken as the reference.

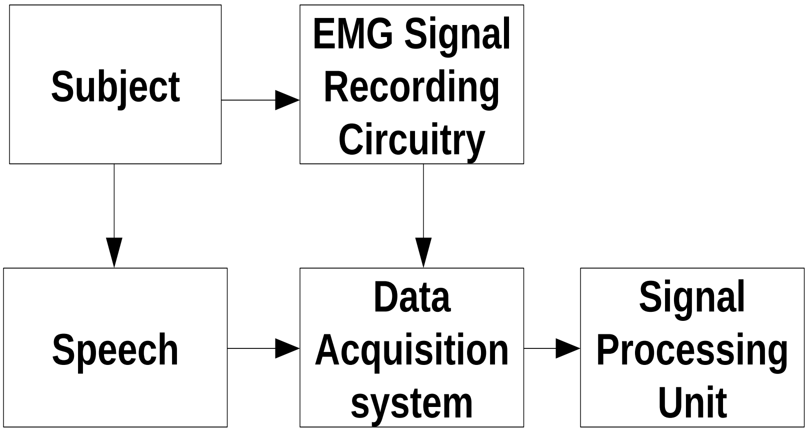

Fig. 4 Schematic block diagram for EMG signal acquisition

IV. Results and Conclusions

Time-domain signals and the corresponding spectrograms of the speech and facial EMG signals for two unvoiced consonants आटा and आसा are shown in Fig. 5 to Fig. 10 for six subjects. Here the x-axis represents the normalized time and y-axis represents the normalized frequency. The signals and the corresponding spectrograms show that the signals generated vary across the subjects and the syllables. Table I shows the means and standard deviations of unvoiced consonants आटा and आसा of all the six speakers. Analysis of the mean LSDs also suggests that the distances are a function of subject and VCV. Hence, training of the automated EMG recognition system needs both the EMG signals and the information regarding subjects

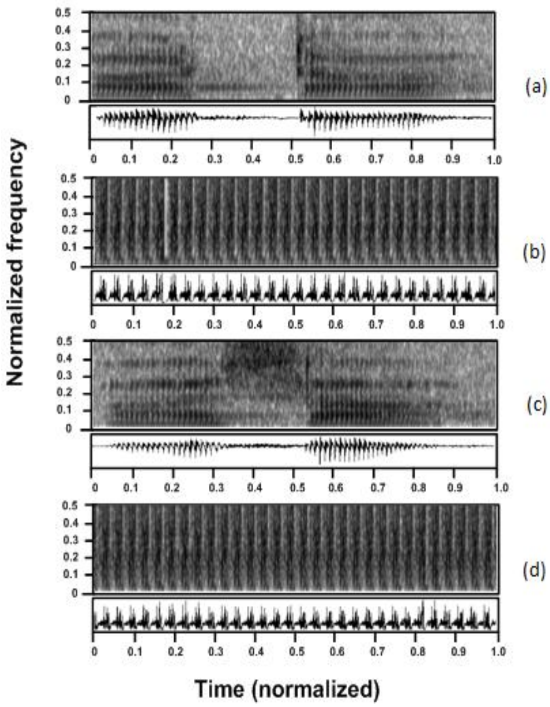

Fig. 5 Spectrogram of unvoiced consonant (a) Original speech signal (आटा) (b) Facial EMG signal (आटा) (c) Original speech signal (आसा) (d) Facial EMG signal (आसा) for Sp1

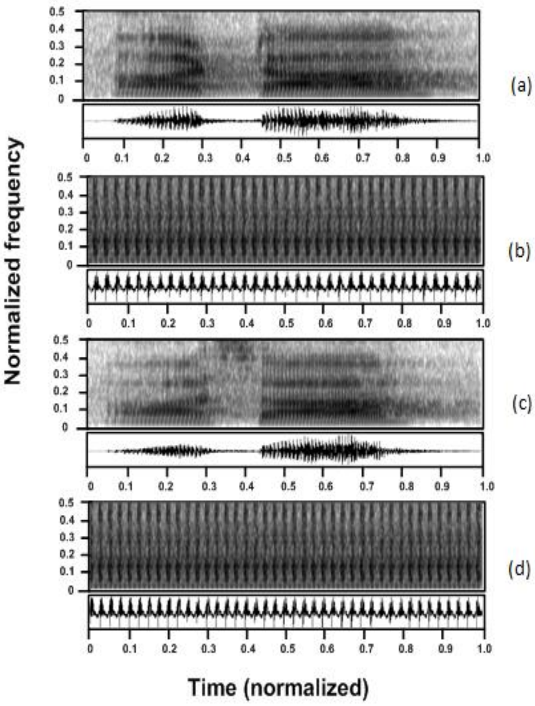

Fig. 6 Spectrogram of unvoiced consonant (a) Original speech signal (आटा) (b) Facial EMG signal (आटा) (c) Original speech signal (आसा) (d) Facial EMG signal (आसा)) for Sp2

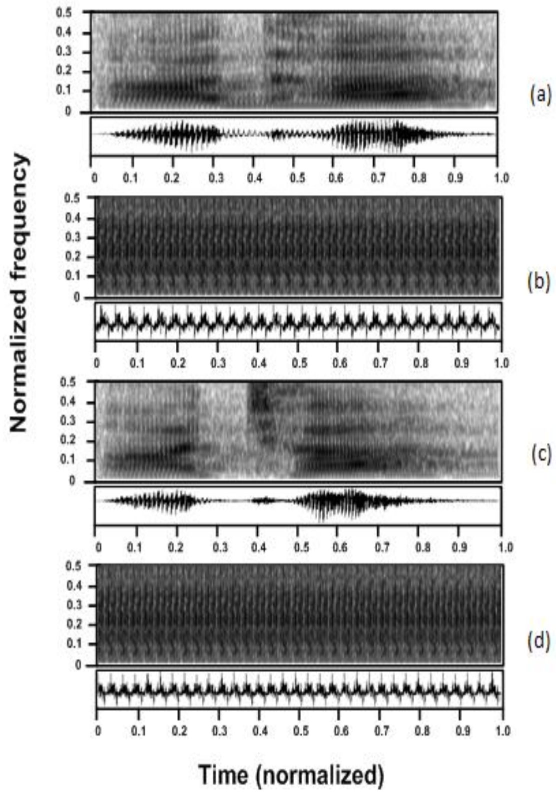

Fig. 7 Spectrogram of unvoiced consonant (a) Original speech signal (आटा) (b) Facial EMG signal (आटा) (c) Original speech signal (आसा) (d) Facial EMG signal (आसा*)* for Sp3

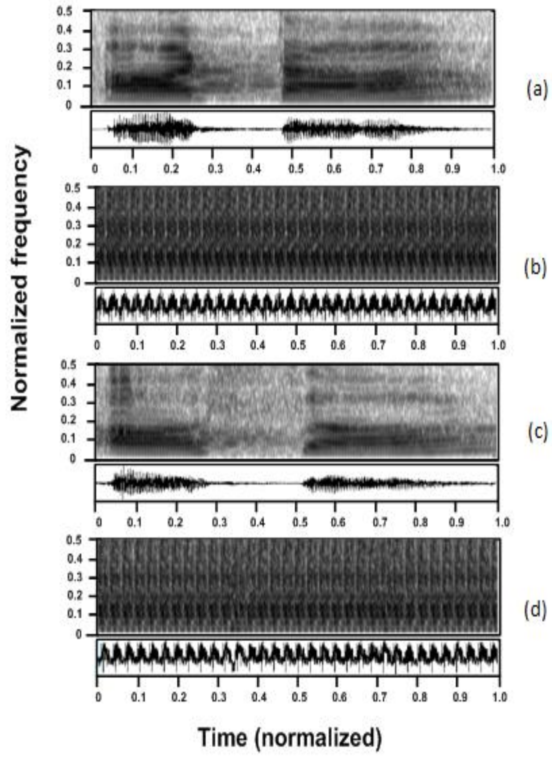

Fig. 8 Spectrogram of unvoiced consonant (a) Original speech signal (आटा) (b) Facial EMG signal (आटा) (c) Original speech signal (आसा) (d) Facial EMG signal (आसा) for Sp4

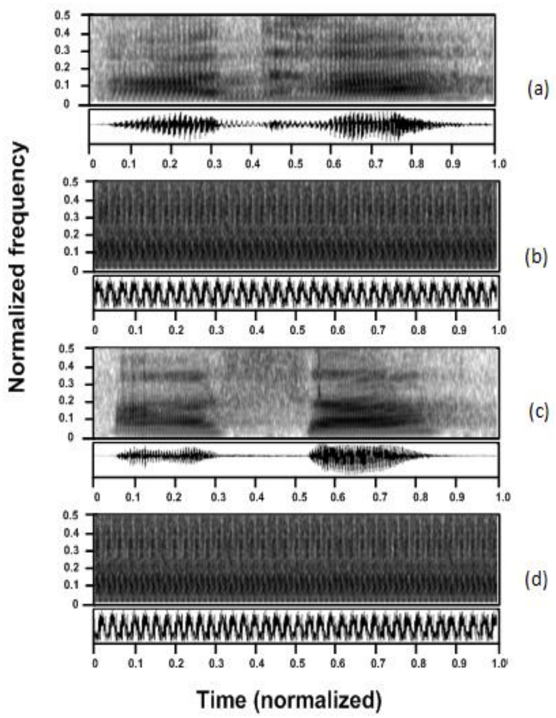

Fig. 9 Spectrogram of unvoiced consonant (a) Original speech signal (आटा) (b) Facial EMG signal (आटा) (c) Original speech signal (आसा) (d) Facial EMG signal (आसा) for Sp5

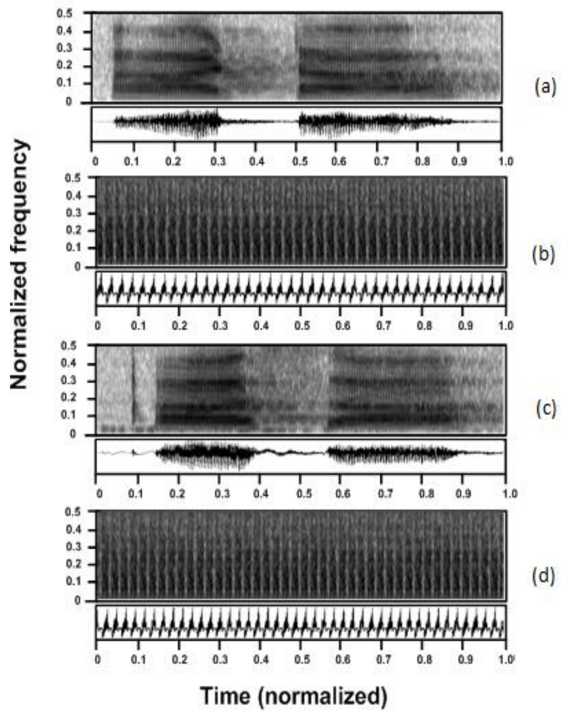

Fig. 10 Spectrogram of unvoiced consonant (a) Original speech signal (आटा) (b) Facial EMG signal (आटा) (c) Original speech signal (आसा) (d) Facial EMG signal (आसा) for Sp6

Table I. Mean and standard deviation of LSD

Speaker	*Unvoiced consonant (आटा)*		*Unvoiced consonant (आसा)*	
	Mean	*S.D.*	*Mean*	*S.D.*
Sp1	7.26	1.46	8.64	2.14
Sp2	7.37	1.53	7.52	1.31
Sp3	6.03	1.04	5.95	0.92
Sp4	6.56	0.86	7.40	0.96
Sp5	7.43	0.84	7.78	0.73
Sp6	12.56	6.04	11.75	4.54

REFERENCES

[1] D. Wied, M. V. Boxtel, A. Zaalberg, R. P. Goudena, and W. Matthys. "Facial EMG responses to dynamic emotional facial expressions in boys with disruptive behavior disorders," Journal of Psychiatric Research, vol. 40, 2006 pp. 112-121.

[2] P. Ekman and W. V. Friesen, "Facial Action Coding System (FACS)," A technique for the measurement of facial action, Consulting Psychologists Press, Palo Alto, CA, 1978.

[3] P. Ekman, J. C. Hager, and W.V. Friesen, "The symmetry of emotional and deliberate facial actions," Psychophysiology, vol18, 1981, pp. 101-106.

[4] V. B. Anton, "Facial EMG as a Tool for Inferring Affective States," Proceedings of Measuring Behavior 2010 Eindhoven, The Netherlands, 2010, pp. 24-27.

[5] J. H. Hermie, F. Bart, K. D. Catherine and R. Gunter, "Development of recommendations for SEMG sensors and sensor placement procedures," Journal of Electromyography and Kinesiology, vol10, 2000, pp. 361-374.

[6] S. Deketelaere, O. Deroo, and T. Dutoit, "Speech processing for communications what's new?" 1 Copernic Ave, Initialis scientific park, 2001, pp. 5-24.

[7] B. H. Juang, and L. R. Rabiner, "Hidden markov models for speech recognition," Techno metrics, Speech research department, AT&T bell laboratories, Murray hill, vol. 33, 1991, pp. 251-272.

[8] S. E. Masri, X. Pelorson, P. Saguet, and P. Badin, "Vocal tract acoustics using the transmission line matrix (TLM) method," in Proc. of the Fourth International Conference on Spoken Language Processing, Philadelphia, 1996, pp. 953–956.

[9] C. Lu, T. Nakai, and H. Suzuki, "Finite element simulation of sound transmission in vocal tract," J. Acoust. Soc. Jpn, pp. 1993, 2577–2585.

[10] N. Umeda, and R. Teranishi, "Phonemic feature and vocal feature: Synthesis of speech sounds, using an acoustic model of vocal tract," J. Acoust. Soc. Jpn, 1966, pp. 195-203.

[11] T. Arai, "The replication of Chiba and Kajiyama's mechanical models of the human vocal cavity," J. of the Phonetic Society of Jpn, 2001, pp. 31-38.

[12] T. Arai, E. Maeda, N. Saika, and Y. Murahara, "Physical models of the human vocal tract as tools for education in acoustics," in Proc. of the First Pan- American/Iberian Meeting on Acoustics, Cancun, 2002.

[13] T. Arai, "The replication of Chiba and Kajiyama's mechanical models of the human vocal cavity," J. of the Phonetic Society of Japan," Vol.5, 2000, pp. 31-38.

