Multifractal Analysis of Speech Imagery of IPA Vowels

Debdeep Sikdar[1], Rinku Roy[2], Koushik Bakshi[3], and Manjunatha Mahadevappa[4]

***Abstract*— In Brain Computer Interfacing (BCI), speech imagery is still at nascent stage of development. There are few studies reported considering mostly vowels or monosyllabic words. However, language specific vowels or words made it harder to standardise the whole analysis of electroencephalography (EEG) while distinguishing between them. Through this study, we have explored significance of multifractal parameters for different imagined vowels chosen from International Phonetic Alphabets (IPA). The vowels were categorised into two categories, namely, soft vowels and diphthongs. Multifractal analysis at EEG subband levels were evaluated. We have also reported significant contrasts between spatiotemporal distributions with fractal analysis for activation of different brain regions in imagining vowels.**

I. Introduction

Verbal communication is the most convenient way of interacting with one another for humans. However, due to different ailments, some individuals are unable to produce speech. Assistive technologies as alternate interpreters will be highly beneficial for them in order to communicate through speech.

Classical brain-language model as proposed by Broca, Wernicke, Lichtheim, Geschwind, and others have been very useful for the researchers to locate the brain area responsible for speech imagery as well as production. While Broca's area is responsible for language production, Wernicke's area processes words. Oro-pharyngeal-laryngeal muscle groups are coordinated by premotor cortex area for proper generation of speech. However, this classic 'Wernicke-Lichtheim-Geschwind' has been challenged by recent studies for their relevancy. Rather, new concept of whole brain theory for language production is emerging.

Several studies have been conducted to decode speech imagery with different approaches. Majority of the researchers concentrated on the brain regions responsible for speech production according to classical brain language model. DaSalla et al. proposed spatial filtering based EEG classification between imagined vowels /a/ and /u/ [1]. D'Zmura et al. have considered two syllables /ba/ and /ku/ while classifying the motor imagery with matched filter for Hilbert envelope of the EEG signals [2]. Rojas et al. have utilised Blackman-Tukey transform with SVM to achieve high accuracy in order to separate spanish vowels /a/ and /e/ [3]. Mel Frequency Cepstral Coefficients of 5 vowels were classified with different classifiers by Riaz et al [4].

In this study, we have analysed the multifractal parameters of soft vowels and diphthongs in speech imagery. These were chosen from standardised International Phonetic Alphabets (IPA) to avoid language specific dependancy. Vowel sounds solely depend on the structure of mouth aperture while speaking. Diphthongs are combination of two soft vowels. As there are hardly any meaning of isolated vowels, artefacts due to associated cognitive or comprehensive imagery are very minimal. Thus, regions extending to whole brain were considered for exploring speech imagery. To further extend our analysis, we have decomposed the EEG signals into α, β, and γ subbands to explore the effect of multifractality in frequency band specific level for different brain regions during vowel motor imagery. Statistical significance analysis revealed more insight on the contribution of different brain regions at different frequency levels towards imagining soft vowels and diphthongs.

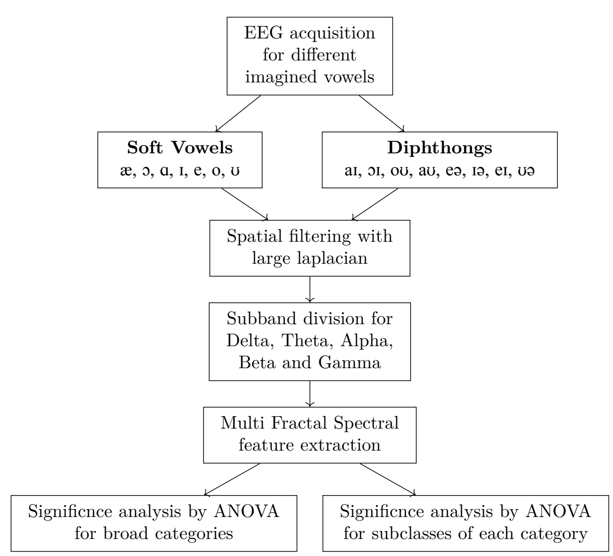


Fig. 1: Flowchart of the overall process for imagined speech.

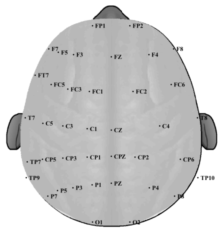
[1]Debdeep Sikdar is a PhD scholar of School of Medical Science and Technology, Indian Institute of Technology Kharagpur, India deep@iitkgp.ac.in

[2]Rinku Roy is a PhD scholar of Department of Advanced Technology Development Centre, Indian Institute of Technology Kharagpur, India rinku.roy@iitkgp.ac.in

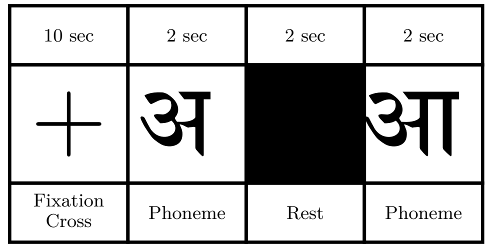
[3]Koushik Bakshi is a PhD scholar of Department of Advanced Technology Development Centre, Indian Institute of Technology Kharagpur, India koushikbakshi.bme@gmail.com

[4]Manjunatha Mahadevappa is Professor of School of Medical Science and Technology, Indian Institute of Technology Kharagpur, India mmaha2@iitkgp.ac.in

978-1-5386-3646-6/18/$31.00 ©2018 IEEE



(a) Montage



(b) Phoneme in Devnagri

Fig. 2: Experiment Protocol

TABLE I: Vowels according to IPA

(a) Soft Vowels

Type	IPA	Devnagri	As in
Open Front	/æ/	/ऐ/	c**a**t
Open-mid Back	/ɔ/	/अ/	s**aw**
Open Back	/ɑ/	/आ/	d**ar**k
Close Front	/ɪ/	/इ/	gr**i**p
Close-mid Front	/e/	/ए/	b**e**d
Close-mid Back	/o/	/ओ/	c**o**ld
Close Back	/ʊ/	/उ/	p**u**t

(b) Diphthongs

IPA	Devnagri	As in
/ɪə/	/या/	**ea**r
/eə/	/एया/	h**ai**r
/ʊə/	/वा/	t**ou**r
/eɪ/	/एइ/	w**ai**t
/aɪ/	/आई/	t**i**de
/ɔɪ/	/अय/	v**oi**d
/oʊ/	/औ/	b**ow**l
/aʊ/	/आउ/	h**ou**se

II. Materials and Methods

Multifractal Detrended Fluctuation Analysis (MFDFA) was utilised to extract the features from the EEGs. Here, we investigated changes in EEG in total 15 different vowels in imagined speech. These vowels were selected from IPA table. They were categorised into two categories, namely, soft vowels and diphthongs. Diphthongs are two soft vowels pronounced successively. These vowels have been standardised based on their usage in constructing words.

Each dataset were applied with MFDFA to get multifractal parameters. In order to distinguish between imagining soft vowels and diphthongs, significance of each parameters were evaluated for each sub-bands as well as full-band EEGs separately for each channel through ANOVA. The Flowchart of this study is given in Fig. 1.

Ten non-impaired subjects, aged between 20 and 27 years old were chosen. The subjects had no history of speech related ailments. They were asked to perform vowel imagery according to the visual cues. None of the subjects were informed about the experimental procedure in prior to minimise biasness. Written consents were taken from the subject before participating.

A. Task

The subjects were seated on an armchair in front of a table. A screen was placed on the table for visual cues. Initially there was a blank screen with a crosshair at the centre. The subjects were asked to relax and concentrate on the instruction followed. Different soft vowels and diphthongs as in Table I were shown randomly in the screen for 2 seconds. Between two successive vowels, a blank screen was shown for 2 seconds. The subjects were needed to imagine the letters shown. There were 20 trials for each case. Fig. 2b shows the timing of the stimulus. Some stressed vowels were excluded for inconsistency in isolated pronunciation. Devnagri script was used for the vowels due to their distinct inherent pronunciation capability and also the subjects were native to the scripture.

B. Signal Acquisition

EEG data were recorded at a sampling rate of 1200Hz from positions FP1, FP2, F7, F5, F3, Fz, F4, F8, FT7, FC5, FC3, FC1, FC2, FC6, T7, C5, C3, C1, Cz, C4, T8, TP9, TP7, CP5, CP3, CP1, CPz, CP2, CP6, TP10, P7, P5, P3, P1, Pz, P4, P8, O1, and O2 by scalp electrodes placed according to the International 10-20 system as shown in Fig. 2a. The reference electrode was placed to the right earlobe and the ground electrode (GND) was placed on the forehead (AFz) to form a feedback loop. The loop also drives down the common mode potential and effective impedance of ground. The EEG recording was carried out using g.Hiamp (g.tec, Graz, Austria) hardware with 39 monopolar channels. 8th order butterworth bandpass filter with 0.01 - 60Hz range and 50Hz Notch filter were added at the recording time to remove linear trends and electrical noise.

C. EEG Preprocessing

After referencing, the signals were detrended. The resulting signals were further processed for EOG and EMG artefact removal. The signals were segmented according to the stimulus timing of 2s length. For imagined dataset, α (8-12Hz), β (13-30Hz) and γ (30-60Hz) frequency subbands were extracted from the preprocessed EEG data. FIR filters were used to these sub bands. For each subbands along with $full$band EEG were processed with Multifractal Detrended Fluctuation Analysis.

III. Analysis

EEG signal has generalised Hurst's exponents due to its complex scaling attributes. This results into its multifractal behaviour. To address this multifractality, multifractal detrended fluctuation analysis (MFDFA) was proposed as an extension of standard DFA [5].

MFDFA measures local roughness from singularity exponent $\boldsymbol{\alpha}$ (thinfaced α referred to EEG band). Points with same singularity exponents are characterised by $\boldsymbol{f(\alpha)}$ as their strength of multifractality or fractal dimensions. This $\boldsymbol{\alpha}$ vs $\boldsymbol{f(\alpha)}$ plot, known as multifractality spectrum, determines the fractal nature of the original series. For multifractal series, this plot is inverted parabolic in nature with maxima at $\boldsymbol{\alpha_0}$. Smaller value of $\boldsymbol{\alpha_0}$ signifies correlated underlying process. Besides, the dynamic nature of the process with high degree of multifractality is depicted from width of the multifractal spectrum, $\boldsymbol{\Delta\alpha}$.

Furthermore, the curve is fitted with a quadratic equation with constant as $\mathbf{1}$. The linear coefficient of the equation, $\boldsymbol{B}$, is termed as asymmetry parameter. $\boldsymbol{B}$ becomes $\mathbf{0}$ for a symmetric spectrum. Positive $\boldsymbol{B}$ value indicates a left skewed spectrum for a series with extreme events dominance over low fractal exponents. On the contrary, negative $\boldsymbol{B}$ signifies finely structured series with strongly weighted low fractal exponents.

$\boldsymbol{H}$ is the Hurst exponent derived from the generalised Hurst exponent at order 2. Its range $0.5 - 1$ indicates a long-term positive autocorrelation, whereas long-term switching between high and low values in adjacent pairs of the time series at $0 - 0.5$ and at a value of $\boldsymbol{H} = 0.5$ can indicate a completely uncorrelated series.

To identify the internal nonlinear dynamics of EEG, MFDFA was performed in this study. The parameters extracted as features from EEG are $\boldsymbol{\Delta\alpha}$, $\boldsymbol{\alpha_0}$, $\boldsymbol{B}$, and $\boldsymbol{H}$. Detailed description of the methodology were explained in earlier literatures [6], [7].

IV. Results

Multifractal parameters for each subbands and full-band EEGs for each channels were computed. Table II lists the significant multifractal parameters for distinguishability between soft vowels and diphthongs for different channels at different frequency band levels based on ANOVA of the parameter values with 95% confidence interval ($p < 0.05$). It was also extended with the significant parameters for separating individual soft vowels and diphthongs. It can be seen that in *Full*band EEG, all the regions of the brain exhibited significantly distinct multifractality between imagining soft vowels and diphthongs. The lower the frequency levels, the lower are the regional distinct multifractality. $\boldsymbol{\Delta\alpha}$ being the most significant parameter among all the brain regions, it can be inferred that there are considerable difference in correlated processes while imagining soft vowels and diphthongs. In fact, it was found that diphthongs contained more correlation than soft vowels as diphthongs shared common soft vowels among themselves. Further-

TABLE II: Significant multifractal parameters ($p < 0.05$) for each subbands and band limited EEG for each channels in distinguishing between soft vowels and diphthongs

Channels	Soft Vowels & Diphthongs				Soft Vowels				Diphthongs			
	α	β	γ	*Full*	α	β	γ	*Full*	α	β	γ	*Full*
FP1			$\boldsymbol{\alpha_0, \Delta\alpha, H}$	$\boldsymbol{\alpha_0, \Delta\alpha, H}$	$\boldsymbol{H}$							
FP2		$\boldsymbol{\Delta\alpha, H}$	$\boldsymbol{\alpha_0}$			$\boldsymbol{\alpha_0}$						
F7			$\boldsymbol{\Delta\alpha, H}$	$\boldsymbol{\Delta\alpha, H}$	$\boldsymbol{H}$					$\boldsymbol{B}$		
F5		$\boldsymbol{\Delta\alpha}$	$\boldsymbol{\Delta\alpha, B}$	$\boldsymbol{\Delta\alpha, H}$	$\boldsymbol{B}$							
F3												
FZ	$\boldsymbol{B}$	$\boldsymbol{H}$	$\boldsymbol{B}$	$\boldsymbol{\Delta\alpha, H}$	$\boldsymbol{H}$					$\boldsymbol{\alpha_0}$		
F4	$\boldsymbol{\Delta\alpha}$	$\boldsymbol{\Delta\alpha, H}$		$\boldsymbol{\Delta\alpha}$	$\boldsymbol{H}$					$\boldsymbol{\alpha_0}$		
F8									$\boldsymbol{\alpha_0, \Delta\alpha}$		$\boldsymbol{\alpha_0}$	
FT7			$\boldsymbol{\Delta\alpha, H}$	$\boldsymbol{\Delta\alpha, B, H}$	$\boldsymbol{H}$	$\boldsymbol{B}$				$\boldsymbol{\alpha_0}$		
FC5	$\boldsymbol{\Delta\alpha}$		$\boldsymbol{\Delta\alpha, H}$	$\boldsymbol{\Delta\alpha, H}$					$\boldsymbol{\alpha_0, B}$			
FC3	$\boldsymbol{\Delta\alpha}$	$\boldsymbol{\alpha_0, \Delta\alpha, H}$	$\boldsymbol{\Delta\alpha, H}$	$\boldsymbol{\Delta\alpha, B, H}$								
FC1	$\boldsymbol{\Delta\alpha}$	$\boldsymbol{\Delta\alpha, H}$	$\boldsymbol{B}$	$\boldsymbol{\Delta\alpha, H}$	$\boldsymbol{\alpha_0, \Delta\alpha}$							
FC2		$\boldsymbol{H}$		$\boldsymbol{\Delta\alpha}$	$\boldsymbol{H}$							
FC6	$\boldsymbol{B}$				$\boldsymbol{B, H}$			$\boldsymbol{B}$				
T7	$\boldsymbol{B}$		$\boldsymbol{\Delta\alpha, H}$	$\boldsymbol{\Delta\alpha, H}$				$\boldsymbol{\alpha_0}$			$\boldsymbol{B}$	
C5			$\boldsymbol{\Delta\alpha, H}$	$\boldsymbol{\Delta\alpha, H}$	$\boldsymbol{H}$		$\boldsymbol{\alpha_0}$		$\boldsymbol{\Delta\alpha}$			
C3			$\boldsymbol{\Delta\alpha, H}$	$\boldsymbol{\Delta\alpha, H}$								
C1	$\boldsymbol{\Delta\alpha}$	$\boldsymbol{\Delta\alpha, H}$		$\boldsymbol{\alpha_0, \Delta\alpha, H}$	$\boldsymbol{\Delta\alpha}$							
CZ	$\boldsymbol{\Delta\alpha}$	$\boldsymbol{\Delta\alpha, H}$		$\boldsymbol{\alpha_0, \Delta\alpha, H}$	$\boldsymbol{H}$							
C4				$\boldsymbol{\Delta\alpha, H}$	$\boldsymbol{H}$							
T8	$\boldsymbol{H}$		$\boldsymbol{\Delta\alpha}$	$\boldsymbol{\Delta\alpha, H}$	$\boldsymbol{H}$							
TP9										$\boldsymbol{\alpha_0}$		
TP7			$\boldsymbol{\Delta\alpha, H}$	$\boldsymbol{\Delta\alpha, H}$	$\boldsymbol{B}$			$\boldsymbol{\alpha_0}$				
CP3	$\boldsymbol{\Delta\alpha}$	$\boldsymbol{\Delta\alpha}$	$\boldsymbol{\Delta\alpha, H}$	$\boldsymbol{\Delta\alpha, H}$	$\boldsymbol{H}$			$\boldsymbol{B}$				
CP5	$\boldsymbol{\Delta\alpha}$		$\boldsymbol{\Delta\alpha, H}$	$\boldsymbol{\Delta\alpha, H}$				$\boldsymbol{\alpha_0}$				$\boldsymbol{\alpha_0}$
CP1												
CPZ									$\boldsymbol{\alpha_0}$	$\boldsymbol{\alpha_0}$	$\boldsymbol{\alpha_0}$	
CP2		$\boldsymbol{H}$		$\boldsymbol{\Delta\alpha, H}$	$\boldsymbol{H}$				$\boldsymbol{B}$			
CP6		$\boldsymbol{\Delta\alpha, H}$			$\boldsymbol{H}$		$\boldsymbol{B}$			$\boldsymbol{\alpha_0}$		
TP10		$\boldsymbol{\alpha_0, \Delta\alpha, B, H}$		$\boldsymbol{\Delta\alpha, H}$								
P7			$\boldsymbol{\Delta\alpha, H}$	$\boldsymbol{\Delta\alpha, H}$								
P5	$\boldsymbol{\Delta\alpha}$		$\boldsymbol{\Delta\alpha, H}$	$\boldsymbol{\Delta\alpha, H}$	$\boldsymbol{\Delta\alpha}$	$\boldsymbol{B}$		$\boldsymbol{B}$	$\boldsymbol{\alpha_0}$			$\boldsymbol{\alpha_0}$
P3	$\boldsymbol{\alpha_0, B}$			$\boldsymbol{\Delta\alpha, H}$		$\boldsymbol{\Delta\alpha}$		$\boldsymbol{\alpha_0}$				
P1	$\boldsymbol{H}$	$\boldsymbol{H}$		$\boldsymbol{\Delta\alpha, H}$	$\boldsymbol{H}$		$\boldsymbol{\alpha_0}$			$\boldsymbol{B}$		
PZ	$\boldsymbol{\Delta\alpha, H}$	$\boldsymbol{\Delta\alpha, H}$		$\boldsymbol{\Delta\alpha, H}$	$\boldsymbol{H}$	$\boldsymbol{B}$	$\boldsymbol{\alpha_0}$					
P4		$\boldsymbol{H}$			$\boldsymbol{H}$					$\boldsymbol{B}$		
P8	$\boldsymbol{\Delta\alpha}$	$\boldsymbol{\Delta\alpha, B, H}$		$\boldsymbol{\Delta\alpha}$	$\boldsymbol{H}$							
O1		$\boldsymbol{H}$							$\boldsymbol{\alpha_0}$	$\boldsymbol{B}$		
O2	$\boldsymbol{B}$		$\boldsymbol{B}$	$\boldsymbol{B}$								

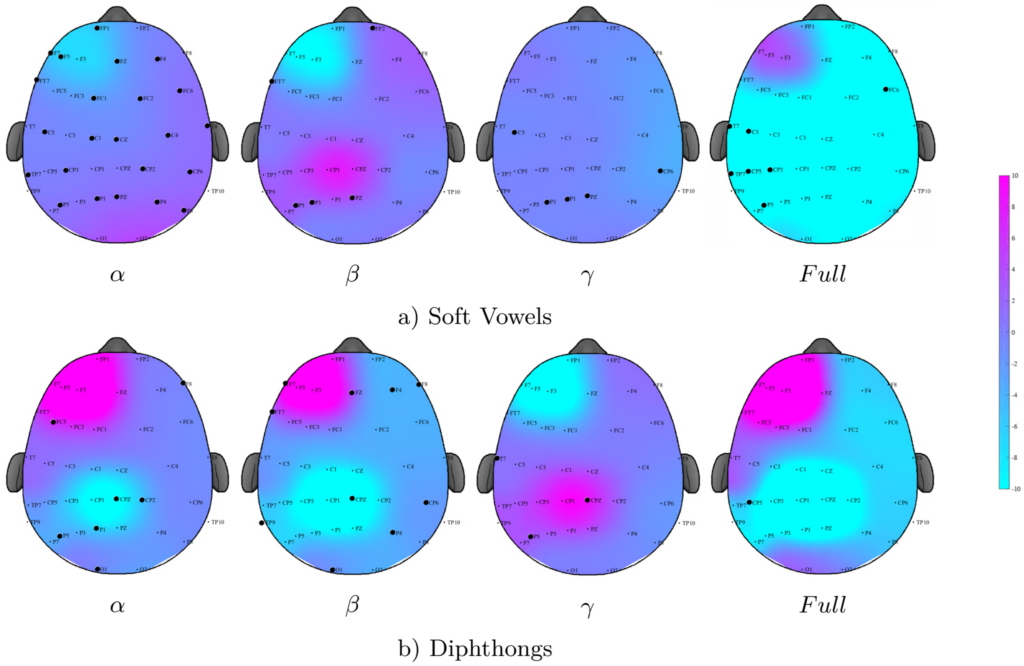


Fig. 3: EEG channel contribution on scalp

more, it can also been seen that soft vowels have the most significant parameters in α subband with $\boldsymbol{H}$, whereas, $\boldsymbol{\alpha_0}$ is the most significant in diphthongs across frequency bands.

Fig. 3 shows the brain region activation plots for the subbands and $Full$band EEGs for soft vowels and diphthongs. The channel locations with at least one significant multifractal parameter for distinguishability, as depicted from Table II, are marked with larger dots. $Full$band EEG can be be seen to have contrasting activated Broca's area with respect to subband levels. But in case of diphthongs, they are similarly activated excepting for γ subband. Moreover, in case of diphthongs, γ subband showed highly activated motor cortical region than the other bands. This has close similarity with β subband in soft vowels. It can be seen that γ subband exhibited diffused activation of brain regions for soft vowels. Although whole brain participated in speech imagery, from the spatial distribution it can also be inferred that contribution of regions related to classical brain language model is significant.

V. Conclusion

We have found although spatiotemporally different vowel imageries were indistinguishable, they are significantly distinct in non-linearity domain even in subband levels. We can further explore to understand the formation of diphthongs from multiple soft vowels. Moreover, similar analysis will also be evaluated on consonants and word formations to decode imagined speech using EEG.

References

[1] C. S. DaSalla, H. Kambara, M. Sato, and Y. Koike, "Single-trial classification of vowel speech imagery using common spatial patterns," *Neural Networks*, vol. 22, no. 9, pp. 1334–1339, 2009.

[2] M. D'Zmura, S. Deng, T. Lappas, S. Thorpe, and R. Srinivasan, "Toward EEG sensing of imagined speech," in *International Conference on Human-Computer Interaction*, Springer, 2009, pp. 40–48.

[3] D. A. Rojas, O. L. Ramos, and J. E. Saby, "Recognition of spanish vowels through imagined speech by using spectral analysis and SVM," *J. Inf. Hiding Multimed. Sig. Process*, vol. 7, no. 4, pp. 889–897, 2016.

[4] A. Riaz, S. Akhtar, S. Iftikhar, A. A. Khan, and A. Salman, "Inter comparison of classification techniques for vowel speech imagery using EEG sensors," in *Systems and Informatics (ICSAI), 2014 2nd International Conference on*, IEEE, 2014, pp. 712–717.

[5] J. W. Kantelhardt, S. A. Zschiegner, E. Koscielny-Bunde, S. Havlin, A. Bunde, and H. E. Stanley, "Multifractal detrended fluctuation analysis of nonstationary time series," *Physica A: Statistical Mechanics and its Applications*, vol. 316, no. 1, pp. 87–114, 2002.

[6] D. Sikdar, R. Roy, and M. Mahadevappa, "Epilepsy and seizure characterisation by multifractal analysis of EEG subbands," *Biomedical Signal Processing and Control*, vol. 41, pp. 264–270, 2018.

[7] D. Sikdar, R. Roy, and M. Mahadevappa, "Multifractal analysis of electroencephalogram for human speech modalities," in *Neural Engineering (NER), 2017 8th International IEEE/EMBS Conference on*, IEEE, 2017, pp. 637–640.