ABUAD Journal of Engineering Research and Development (AJERD)
ISSN (online): 2645-2685; ISSN (print): 2756-6811

Volume 4, Issue 2, 43-55

UNILAG Laughter Corpus – A Dataset for Laughter Analyses and Biometrics

Olumuyiwa Sunday ASAOLU[1], Comfort Oluwaseyi FOLORUNSO[2*], Oluwatoyin Pius POPOOLA[3], Michael Oluwadamilola OYEYEMI[4], Linda ORANYA[5], Florian CUMMINGS[6]

Department of Systems Engineering, University of Lagos, Akoka, Lagos, Nigeria

[1]oasaolu@unilag.edu.ng, [2*]cfolorunso@unilag.edu.ng, [3]opopoola@unilag.edu.ng, [4]oyeyemi.md@gmail.com, [5]oranyalinda7@gmail.com, [6]cummingsflorian16@gmail.com

Corresponding Author: cfolorunso@unilag.edu.ng, +2348080719696

Date Submitted: 09/11/2021
Date Accepted: 16/12/2021
Date Published: 31/12/2021

Abstract: *Laughter is the verbal release of air from the lungs accompanied by distinctive facial and physical gestures that show joy, pleasure, ridicule, or uneasiness. Laughter corpus exists in literature; however, these were not suitable for biometric purposes. This paper presents a laughter dataset for biometrics and laughter analyses. This dataset consists of audio recordings of 123 participants who were viewing humorous video clips. Altogether, 1763 bouts of spontaneous laughter were recorded, aside acted laughter and speech utterances. Acoustic analyses were conducted using conventional features and by modifying standard algorithms. Results obtained from different experimentations are consistent with literature. It is expected that this dataset will spur additional study in the subject, and that it will then be expanded in a gradual way across many application areas over time.*

Keywords: *Acoustic analysis, acted laughter, biometrics, laughter-bout, spontaneous laughter*

1. INTRODUCTION

Laughter is a crucial paralinguistic phenomenon that plays a significant part in human expression and conversation. The expression or manner of human communication is extremely significant in human social interactions and relationships. Laughter served as a nonverbal vocalisation that is frequently seen in continuous speech [1]. Different persons have their favourite subset of phones for laughing [2] and the eloquent expression of laughter varies by sex, individual, and circumstance [3]. Humans naturally produce and notice laughter, notwithstanding its diversity. Laughter frequently expresses a joyful mood and excitement through a mix of sound and facial emotions. Laughter under disillusionment, melancholy, or sarcasm is rare, as stated in an African proverb: 'When a situation is horrible beyond tears, it is laughing that ensues.' As a result, in addition to amusement, laughter may be utilised to identify despair in humans [3]. Communicative laughter can also provide a natural human account [4], as well as aiding human-computer connection [5].

Laughter enhances speech since it occurs frequently during conversation. With a communication objective, it can be controlled (voluntary or performed laughter) as well as spontaneous (involuntary) [3]. The automatic identification of laughter utterances in human speech can help with automatic speaker recognition, automatic speech recognition, paralinguistic tasks like speaker emotion detection, and recognising hilarious material in video clips, among other things. By distinguishing non-speech sounds, incorporating laughter detection into automated speech recognition systems can help to reduce word error rates [6]. One of the most common forms of communication in the world is vocal sound. The sound source vibrates, causing nearby air molecules to vibrate and spread. Sound is created by such continuous sound vibrations in the air or other medium. Human ears perceive, detect, and recognise speech signals based on frequency, allowing one to distinguish one's own voice from that of other speakers or sound sources.

Laughter dataset is the collection of laughter samples and related data for easy retrieval, management and update. Specific applications are in forensics, human computer interaction, animation, biometrics (person and behaviour identification) systems. This dataset is developed as an African repository, to aid research investigations on machine identification of persons from their laughter. To achieve this, laughter samples of volunteers were captured in the Control Laboratory of the Systems Engineering Department, University of Lagos, Akoka, Lagos Nigeria. Laughter-based biometrics could be used in access control if well managed and integrated with other features such as speech or voice-based systems. While culture and personal mood may affect what one finds humorous at a particular place and time, the deployment of laughing gas may be

 ISSN (online): 2645-2685; ISSN (print): 2756-6811

used to elicit laughter if absolutely necessary. More application is expected in forensics to match and track criminals rather than using this technology for routine login or authentication.

Different datasets for laughter analysis were compiled usually with a specific research focus. The research focus varies from laughter detection in meetings [7],[8] to laughter analysis [4], [9] to laughter synthesis [10],[11] with application in human computer interaction [5]. Weninger and Schuller [12] presented the audio-visual interest corpus (AVIC) which is made up of 21 participants (10 females and 11 males). The recording was done such that an experimenter and a participant were sitting facing each other on a desk. The main focus of this dataset was to identify natural interest in human conversation, this dataset however was not for laughter analysis. More so, the Audio-Visual Laughter Cycle (AVLC) dataset was built as part of the audio-visual laughter cycle (AVLaughterCycle) scheme. The goal of the dataset was to create an audio-visual laughing machine that could record a user's laughter and reply with an appropriate chuckle. It contains audio, visual and the facial motion tracking of 24 participants. Over 1000 spontaneous laughter and 27 acted laughter were recorded. This corpus is suitable for automatic laughter processing or cognitive science study [13].

In addition, the Audio-Visual Laughter Synthesis System (AVLASYN) was created for the goal of producing and synthesising audio-visual laughter. It features 251 laughter outbursts from one male participant, while viewing funny video segments that inspire laughter [14]. Though there are a number of laughter utterances generated, they were all from only one male participant and the variability characteristic of laughter may not be adequately captured. However, the Semaine dataset contains audio-visual recordings of participants communicating with avatar, which present different personalities. A total of 150 people took part in the study, with 443 instances of laughter collected from 345 video recordings [15].

An interesting and diverse corpora, is the MAHNOB Laughter dataset [16]. It's an audio-visual dataset that includes laughter, speech, acted/posed laughter, and acted/posed smiles. While watching stimulus material, a total of 22 individuals (12 men and 10 females) were recorded. The major goal was to evoke laughter. A total of 180 sessions with a total duration of 3 hours and 49 minutes were recorded. The dataset contains 563 episodes of laughter, 849 spoken utterances, 51 performed laughs, 67 speeches–laughs episodes, and 167 additional vocalisations. The dataset is suitable for research on laughter. Mckeown *et al*. [17], also, produced Belfast story-telling corpus. With a collection of varied laughter utterances, the corpus was created with the goal of capturing realistic audio and audio-visual laughter in semi-natural social interactions. Each of the six sessions recorded 21 people for over an hour, resulting in a dataset of 25 hours and 40 minutes of high-quality audio and audio-visual recordings, including both speaker and listener laughter. The corpus also includes 25 hours of motion tracking footage from Microsoft Kinect devices in various formats.

The Multimodal and Multi-Person Corpus of Laughter in Interaction (MMLI) was produced in contrast to prior research [18]. It was designed to collect multimodal laughter samples, with a focus on complete body motions and different forms of laughter. It features 16 people's acted and spontaneous laughs. 3D body location data, facial tracking, numerous audio and video channels, physiological data, and data annotation are also included. The dataset includes audio, video, and around 520 laughter episodes, among other things.

The Social Signal Processing Network (SSPNET) is a mobile dataset that includes 60 phone calls between 120 strangers [19]. It's handy for looking for linguistic occurrences. The dataset's objective, however, was to keep research on nonverbal behaviour alive rather than laughter analysis. The International Computer Science Institute of Berkeley (ICSI) corpus is made up of audio recordings made using head-worn and table-top microphones at the same time. Automatic speech recognition, noise robustness, conversation modelling, rich transcription, data recovery, and a variety of additional applications are all possible with this dataset. The dataset includes speech and laughter transcriptions as well as 72 hours of meetings by 53 distinct persons [20].

The Augmented Multi-Party Interaction (AMI) Meeting Corpus was constructed and presented, consisting of 100 hours of recordings of meetings at the University of Edinburgh (UK). Nearly 33 hours of naturally occurring laboratory condition meetings were recorded out of the 100 hours of meeting recording, consisting of 33 naturally occurring sessions. There were 25 sessions with four participants, five encounters with three conversationalists, and three meetings with five participants. The remaining 67 hours are made up of 138 role-playing sessions with a total of four participants [21].

A dataset of non-verbal conversation expression (which consist of smiles, laughter, head and eyebrow movements) was developed with the goal of developing a human-agent interaction system. It consists of 42 speakers from various nationalities speaking in 21 sessions. Nonverbal dialogue expression was the emphasis of the annotation [22].

The MULAI corpus is a multimodal dataset of 357 minutes of video, audio, and physiological data on dyadic interactions, which were frequently accompanied by laughter. This dataset was created to give researchers additional information about the expressive patterns that people exhibit when they laugh during human-human contact[23].

According to research, there is a scarcity of laughter data samples for analysis and detection [24]. The literature on laughter characteristic of people is thus augmented with our input from this region of the world. Besides, no laughter dataset has been built with a sole purpose of biometric analysis. This need has inspired the building of a novel laughter dataset aimed at investigation of laughter as a viable biometric and for laughter analysis. The summary of some existing datasets is shown in Table 1.

 ISSN (online): 2645-2685; ISSN (print): 2756-6811

Table 1: Summary of some of the existing laughter datasets

S/N	Dataset	No. of participants	No. of laughter utterances	No. of speech utterances	No. of acted /posed laugh	Aim of dataset	Reference
1.	AVLASYN	1 male	251			Audio-visual laughter synthesis	[25]
2.	AVLC	24 (9 females, 15 males)	1021 laughter episode with average of 42 episodes per person.		27	Audio-visual laughing machine	[13]
3.	AVIC	21 (10 females, 11 males).				Identification of natural interest in human conversations	[12]
4.	Mahnob	22 (10 females,12 males)	563	849	51	Laughter elicitation	[16]
5.	MMLI	16	500			Gathering multimodal laughter samples with special attention on full body movements.	[18]
6.	Belfast story telling corpus	21 (8 females,13 males)	2336			Gathering naturalistic audio and audio-visual laughter in semi-natural social interactions	[17]
7.	SSPNet	120 (63 females,57 males)				Maintain research on non-verbal behavior	[19]
8.	ICSI	53				Collect speech under a natural condition	[20]
9.	SEMAINE	150	443 instances of laughter			Elicit natural emotion	[15]

2. METHODOLOGY

Undergraduates, postgraduate students, and employees from six faculties (Engineering, Sciences, Social Sciences, Arts, Environmental Sciences, and Education) at the University of Lagos, Akoka, Nigeria, were given consent forms to sign. A total of 550 consent forms were given out while 278 forms were returned and 243 participants eventually turned up for the recording. Those who returned the form, and gave consent to use their laughter samples for research purposes were invited to the Control Laboratory of the Department of Systems Engineering for briefing/demonstration and then scheduled in batches for studio recordings. The sample of the consent form is in Appendix 1. In each session, a volunteer's laughter samples were captured with the use of a C1 condensed microphone while such viewed random video comedy clips in a very relaxed environment. Laugher bouts were extracted from participants in a controlled environment to prevent noise and interference. The recorded audio files were used to build the local laughter dataset.

The laughter recording took place across five months (February to June 2018), with each session lasting an average of 2 hours 58 minutes. Each participant was assigned a code that indicated whether they were male or female, as well as their age range, the session of the laughter recording, and the serial number they fell into throughout each session. The age groups were 17 years and upwards. Age 17 was chosen as the minimum age since it is widely documented in the literature that after reaching puberty, a teen's voice becomes entirely steady at this age [26]. To preserve data privacy, all laughter samples were kept using code-generated taxonomy that did not contain any personal information that may identify any of the real participants.

2.1 Recording Setup

Two laptops, one Behringer C1 Studio Condenser microphone, one set of Goodest GH360 stereo headphones and the PRAAT® software [27] were used for the recording. A session of the setup for the recording is shown in Fig. 1. The comedy skit was played on one laptop for the participant to watch and listen to the audio via the Goodest GH360 stereo headphones

 ISSN (online): 2645-2685; ISSN (print): 2756-6811

connected to the laptop. The second laptop had the PRAAT® software installed while the microphone was connected to it. The laughter samples were captured using the software and saved according to the codes generated.

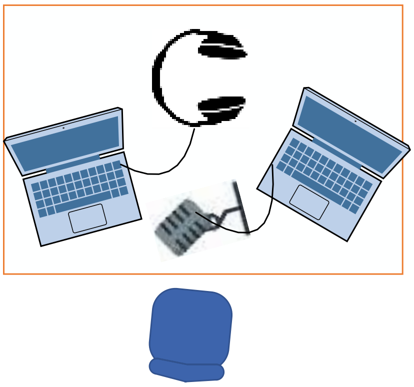

Figure 1: Recording set up

2.2 Recording of the laughter samples

The recording was done when each participant was in relaxed mood and good state of mind as they viewed one of two sets of 10-minute comedy skits from popular comedians. The stages of the dataset construction employed in this study is depicted in Fig. 2, while the recording layout is shown in Fig. 3. Some of the participants (14 females and 56 males) were given five different sentences to read and were also asked to laugh on demand. The sentences read, being somewhat common phrases amidst majority of the sample, are:

i. The boy is going to school
ii. Do you have any family in the city or near you?
iii. I am coming from the market
iv. I have a class today
v. I have test this week

The purpose of reading the sentences was to have the speech samples from the participants for speech acoustics analysis. The induced laughter was recorded using the mono sound option of PRAAT® software at a sampling rate of 44100 Hz and saved as a ‘wav' file at 16-bit Pulse Code Modulation (PCM). The recording of each individual lasted for an average of 10 minutes and then pre-processing was done in order to extract the laughter, speech, spontaneous and acted laughter generated for each participant. On the average, laughter bouts lasted 3 seconds.

ISSN (online): 2645-2685; ISSN (print): 2756-6811

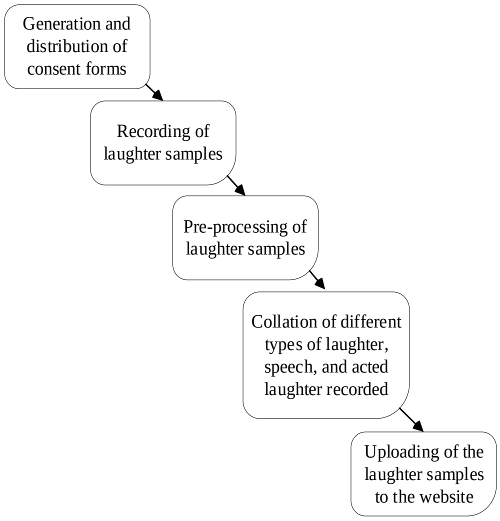

Figure 2: Flow diagram of the dataset development

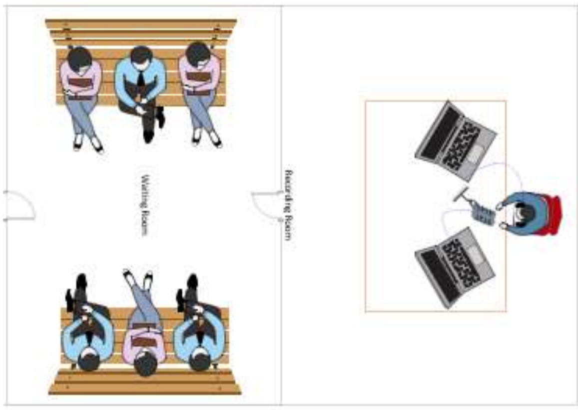

Figure 3: Recording Layout

2.3 Acoustic and Biometric analysis

The acoustic characteristics employed in literature [28][29] for laughter analysis are incorporated in this work since laughter is produced by the same process as speech. The mean intensity, pitch [fundamental frequency (F0)], and formants frequencies (F1 - F5) were all investigated. The following analyses were carried out in this study: the distinction between laughter and speech, as well as the distinction between female and male laughter. For the acoustic analysis, the same number of samples from each subject were utilised to eliminate bias. The studies included ten laughter samples from each of 42 females and 81 males (a total of 1230 laughter samples). The disparity between male and female ratios is owing to the fact that more males consented and showed up for recordings than females.

Biometric analysis was also performed on the local dataset. The VoxForge voice dataset, a well-known dataset for speaker identification tasks [30], was used for all analyses in this work to validate the proposed dataset. For training, ten speech samples from 32 females and 58 males were used, with five samples from each person used for validation. A new feature descriptor was created as a result of rigorous feature engineering to improve on the existing Mel Frequency Cepstral Coefficient (MFCC) feature, which was employed in the biometric studies (see Fig. 4). This new feature descriptor is called the Dynamic Average MFCC (DA-MFCC), which was utilised to train the Support Vector Machine (SVM) employing the Radial Basis Function (RBF) kernel and Artificial Neural Networks (ANN). Equations 1-3 provide a few of the equations that regulate the SVM. This research was carried out using a computer system with an Intel® Core TM i7-8565U CPU @ 1.8GHz (8 CPUs), 8 GB RAM, and a 64-bit Windows 10 operating system. The SVM is created with the Python programming language and the Anaconda software's Scientific Python Development Environment (SPYDER) IDE [31]. The ANN is implemented with the use of the MATLAB® program.

ISSN (online): 2645-2685; ISSN (print): 2756-6811

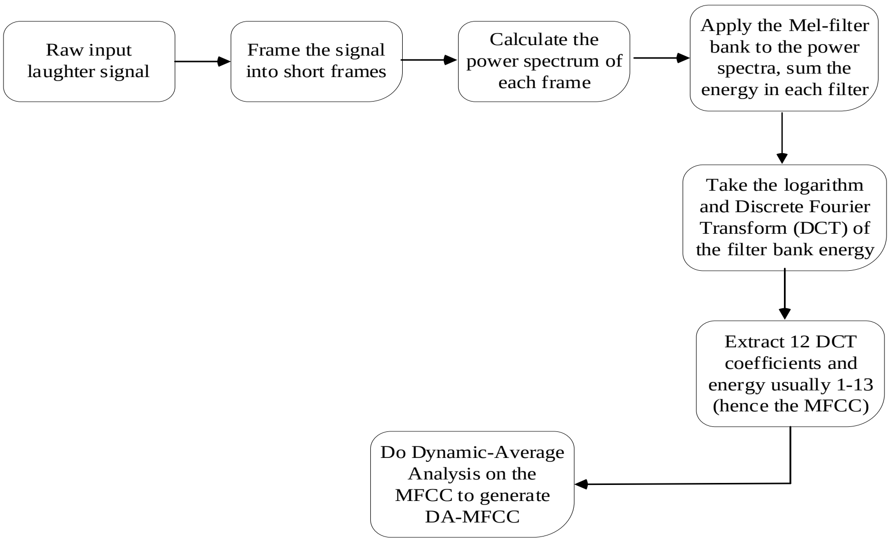

Figure 4: DA-MFCC Feature Extraction stages

The SVM determines the best hyperplane for categorising each positive and negative training sample. Eq. 1 defines the decision boundary of a linear classifier

$$w^T x + b = 0, \quad (1)$$

where 'w' is a weight vector, and 'b' is a bias term.

The operational margin of the I^{th} sample 'x_i' for a hyperplane (w, b) is determined in Eq. 2 for a given corpus and decision hyperplane.

$$\gamma_i = y_i(w^T x_i + b), \quad (2)$$

where $y_1 \in \{-1, 1\}$ and it is considered as the class label and x_i is the class input.

The operational margin of a corpus of decision boundary samples is double that of the corpus samples with minimal functional margin. On each data instance, the kernel function is used to map the original non-linear observations into a higher-dimensional space where they may be separated. Because of the following properties, the radial basis function (RBF) kernel of the SVM was used in this study: simplicity of building, strong generalisation, high tolerance to input noise, and online learning capability. The RBF kernel is described by Eq. 3, where δ is a positive radius control parameter [32].

$$K(x_i, x_j) = \exp(-||x_i - x_j||^2/2\delta^2) \quad (3)$$

This research also uses a Multilayer Perceptron (MLP) network, which is a feed forward artificial neural network. For training, the Levenberg-Marquardt library was integrated into the MATLAB platform. Figures 5 and 6 demonstrate the network design as well as the training parameters used.

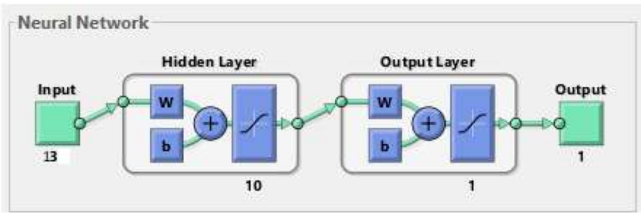

Figure 5: Network architecture implemented in this study

 ISSN (online): 2645-2685; ISSN (print): 2756-6811

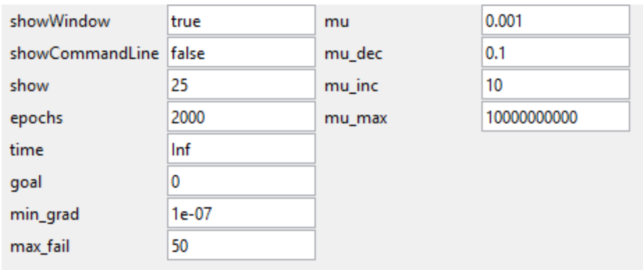

Figure 6: Network parameters used for ANN training

3. RESULTS AND DISCUSSION

In this session, the results of several acoustic analyses are presented. The differences between spontaneous laughter and speech are seen in Table 2. Table 3 also shows the disparities in spontaneous laughter of females and their male counterparts.

Table 2: Acoustic Analysis to differentiate laughter from speech

Features	Laughter	Speech	Higher Value
F0_Mean	336.69	191.86	Laughter
F1_Mean	909.76	706.93	Laughter
F2_Mean	2021.95	1947.26	Laughter
F3_Mean	3016.77	2963.31	Laughter
F4_Mean	4056.27	4045.08	Laughter
F5_Mean	4862.17	4854.07	Laughter
Intensity_Mean	50.70	54.04	Speech

Table 2 shows that the pitch or F0, as well as all the formant frequencies, are higher in laughter than in speech. Because laughter causes an increase in airflow through the vocal tract, which causes quicker vibrating of the vocal folds, the F0 and formants frequencies are higher in laughter than in speech. Speech, on the other hand, has a stronger intensity than spontaneous laughter. This is plausible since the energy of each episode of spontaneous laughter decreases with time, but the energy of speech grows as the word is generated, especially when it is stressed. This observation is in line with Bachorowski and colleagues' findings [29], [33].

Table 3: Acoustic Gender Analysis of Laughter

Features	Laughter Male	Laughter Female	Higher Value
F0_Mean	305.29	367.65	Female
F1_Mean	929.50	890.30	Male
F2_Mean	2002.82	2040.81	Female
F3_Mean	2982.98	3050.08	Female
F4_Mean	4076.93	4035.92	Male
F5_Mean	4913.99	4817.04	Male
Intensity_Mean	52.37	49.06	Male

The distinctions between male and female laughter are depicted in Table 3. The pitch (F0), second and third formant frequencies (F2 and F3) of laughter were all greater in females than in males, according to this table. This is due to the fact that female vocal folds are shorter (between 12.5 mm and 17.5 mm), whereas male vocal folds are longer (between 17 mm and 25 mm) [34]. Fundamental frequency is inversely related to vocal fold length, according to the National Centre for Voice and Speech [35]. As a result, shorter folds have a higher frequency than longer folds. The outcomes of this investigation are consistent with Bachorowski and Owren [29] observations. Males, on the other hand, have greater first, fourth, and fifth formant frequencies (F1, F4, and F5) than females. F1 differs from Bachorowski and Owren [29] findings in that it is greater in females than in males, although F4 and F5 coincide with Bachorowski *et al.*[28]. According to the F4, 60% of the laughter generated in this study consist of voiced closed mouth laughter.

Figure 7 shows the waveform, spectrogram and the transcribed spontaneous laughter of a female, while Figure 8 shows the waveform, spectrogram, intensity of an unvoiced snort-like type of laughter. The pitch, intensity and the formant frequencies of a spontaneous laughter is presented in Figs. 9 to 11. The waveform of an acted laughter of a male is presented in Fig. 12. The confusion matrix of a traditional MFCC is presented in Eq. 4 while the confusion matrix of the new feature

ISSN (online): 2645-2685; ISSN (print): 2756-6811

descriptor introduced in this study (that is the DA-MFCC) is shown in Eq. 5. The new feature developed in this study was extracted from laughter and speech from the local dataset. In order to validate the findings from this study, the same set of features were also extracted from the VoxForge standard dataset. The SVM and ANN were used as classifiers and the results is presented in Figs. 13 and 14.

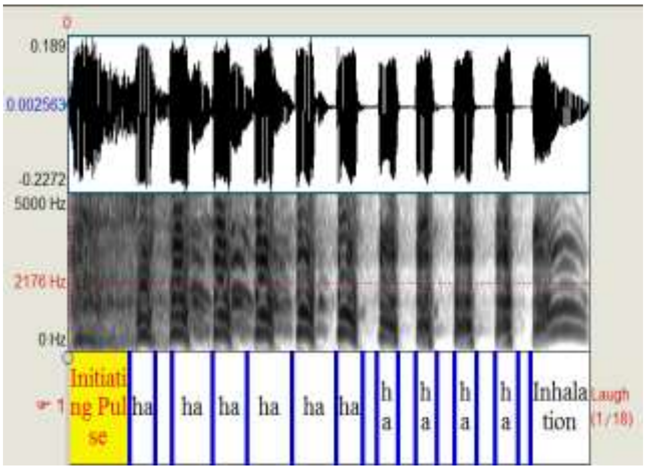

Figure 7: Waveform, spectrogram and the transcribed spontaneous laughter of a female using PRAAT software

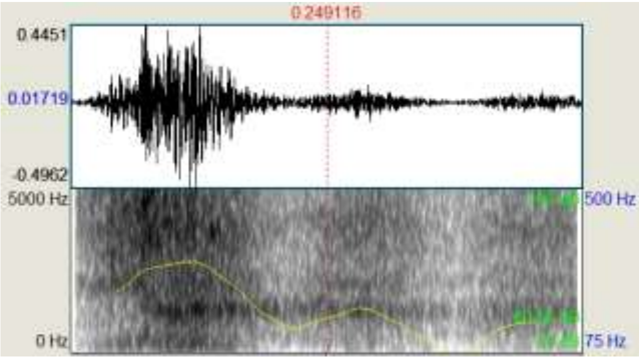

Figure 8: The waveform, spectrogram, intensity of an unvoiced snort-like type of laughter

Figure 7 shows the waveform, spectrogram and the transcription of a long duration high-pitch spontaneous laughter produced by a female. The waveform displays the value of positive and negative acoustic energy that is available at each point in the laughter signal. The spectrogram, which is displayed in the second graph of the figure, shows the energy spread over the frequency ranging from 0 Hz to 5000 Hz. The third graph shows the transcribed laughter comprising of the fricative consonant ‘h’ and vowel ‘a’. A typical laughter is made up of three phases which include the initiating pulse, a kind of outburst that precedes the laughter. The exhalation phase is the second phase of a laughter episode and it is made up of a series of laughter pulses which in this case is ten (10) pulses. The third phase is the inhalation phase which actually ends the laughter episode as shown in the figure. It can be inferred that laughter is made up of as many syllables as there is (combination of consonant-vowel). A vowel mixed with consonant make up voiced laughter. However, if a vowel is missing in any form of laughter, then such laughter is categorised as unvoiced and will not have the fundamental frequency (F0) even though it may have intensity or energy. Such type of laughter includes snort, cackle or chuckle as shown in Fig. 8.

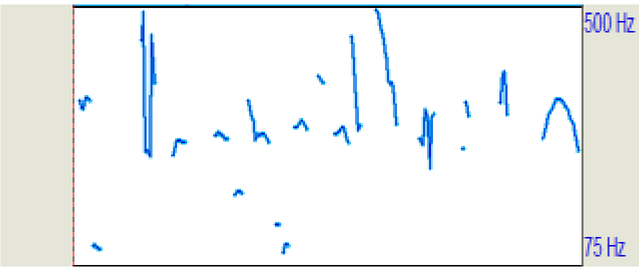

Figure 9: The pitch or fundamental frequency of a spontaneous laughter

ISSN (online): 2645-2685; ISSN (print): 2756-6811

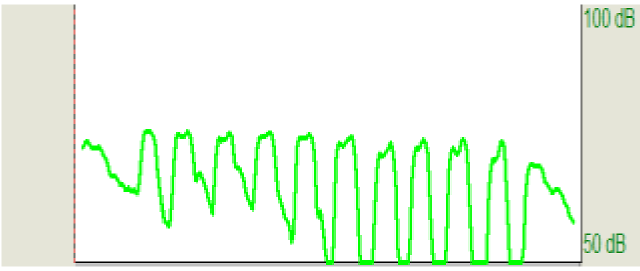

Figure 10: The intensity of a spontaneous laughter

Figure 9 shows the pitch of the periodic spontaneous laughter signal. Pitch is produced as a result of the vibration in the vocal cord during the production of such sound. All vowels sound possesses the fundamental frequency. The intensity or energy of the laughter signal is presented in Fig. 10. The intensity can also be called loudness in some cases. All type of laughter possesses this feature as long as it can be heard.

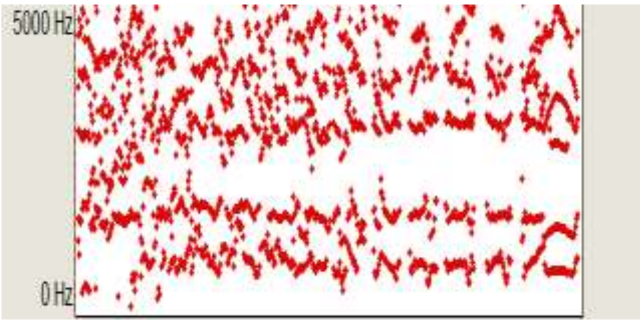

Figure 11: The formant frequencies (F1 – F5) of a spontaneous laughter


Figure 12: Waveform of an acted laughter of a male

Figure 11 shows the various formants present in the laughter signal ranging from F1 to F5. The waveform of an acted laughter of a male is presented in Fig. 12.

Confusion Matric for MFCC

$$\begin{bmatrix} \mathbf{385} & 19 & 3 & \dots & 10 & 8 & 9 \\ 5 & \mathbf{358} & 17 & \dots & 6 & 18 & 19 \\ 11 & 42 & \mathbf{155} & \dots & 37 & 6 & 20 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 1 & 0 & 0 & \dots & \mathbf{611} & 0 & 1 \\ 10 & 22 & 6 & \dots & 50 & \mathbf{358} & 26 \\ 26 & 31 & 20 & \dots & 14 & 29 & \mathbf{179} \end{bmatrix} \quad (4)$$

Confusion Matric for DA-MFCC

$$\begin{bmatrix} \mathbf{842} & 0 & 0 & \dots & 0 & 0 & 0 \\ 0 & \mathbf{811} & 0 & \dots & 0 & 0 & 0 \\ 0 & 0 & \mathbf{822} & \dots & 0 & 0 & 0 \\ \dots & \dots & \dots & \dots & 0 & 0 & 0 \\ 0 & 0 & 0 & \dots & \mathbf{816} & 0 & 0 \\ 0 & 0 & 0 & \dots & 0 & \mathbf{828} & 0 \\ 0 & 0 & 0 & \dots & 0 & 0 & \mathbf{822} \end{bmatrix} \quad (5)$$

ISSN (online): 2645-2685; ISSN (print): 2756-6811

The confusion matrix of the traditional MFCC is shown in Eq. 4, while that of the new feature descriptor (DA-MFCC) is shown in Eq. 5. The diagonal of the confusion matrix of the DA-MFCC indicated a distinct classification of everyone as compared to the traditional MFCC. Therefore, it can be concluded that the DA-MFCC of laughter signature can potentially be used in laughter-based person identification systems for better performance.

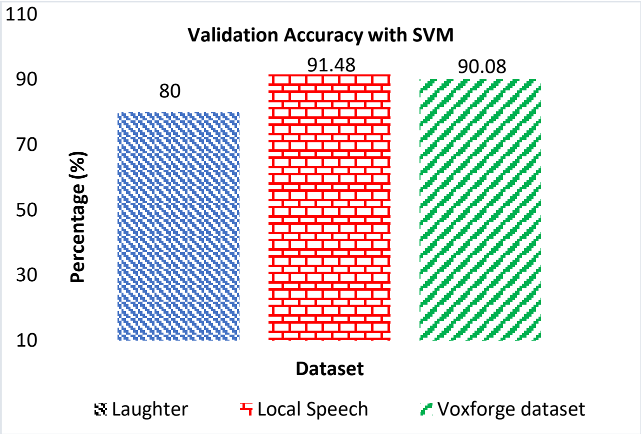

Figure 13: Performance analysis of laugher and speaker identification using SVM

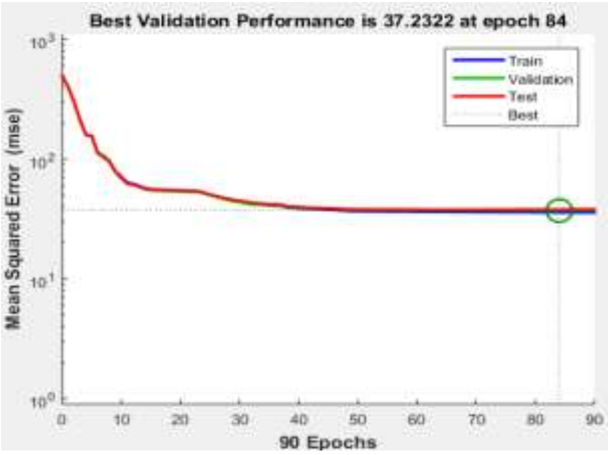

Figure 14: Performance analysis of laughter when trained with ANN

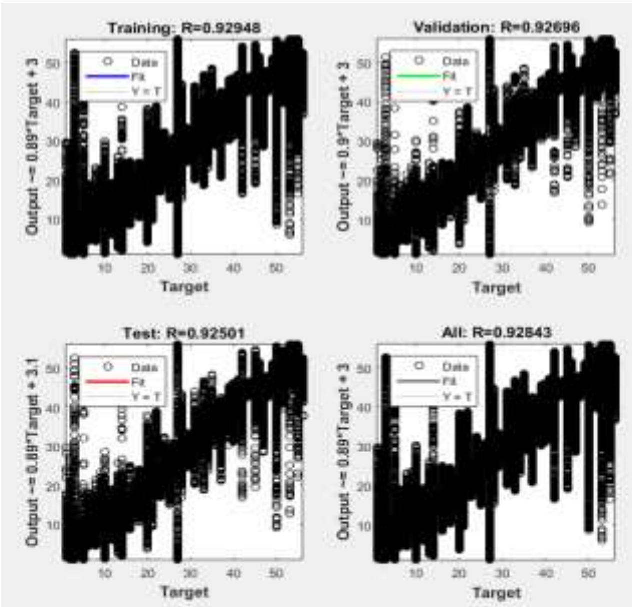

Figure 15: Regression Analysis of laughter when trained with ANN.

ISSN (online): 2645-2685; ISSN (print): 2756-6811

The performance analysis of laugher and speaker identification using the SVM classifier is shown in Fig. 13. While Figs. 14 and 15 shows the performance and regression analyses of laughter respectively using ANN classifier. From these results, it was evident that laughter can also be used as person identifier with 80% and over 90% validation accuracy with SVM and ANN respectively. The use of laughter is beneficial in the absence of speech and owing to the fact that laughter cannot be easily mimicked as compared to speech. Furthermore, laughter needs no more be considered as "noise" in traditional speaker recognition but as a component for additional analysis.

4. CONCLUSION

In this paper, a local laughter dataset is presented for biometrics and laughter analyses. This is relevant since the existing laughter datasets were unsuitable for biometrics purposes. To the best of the authors' knowledge, no previous laughter dataset has been published that might be used for biometric reasons. The dataset contains all sorts of laughter produced by university undergraduates, postgraduate students, and staff members selected at random from six faculties (Engineering, Sciences, Social Sciences, Arts, Environmental Sciences and Education). Various acoustic studies were conducted, with the majority of the results being consistent with the literature.

The resulting dataset comprises 123 individuals' tagged audio samples taken while watching amusing video clips. There were 14 sessions in all, each lasting an average of 2 hours and 58 minutes. There are 1763 episodes of laughter in this dataset. The sample also includes 1628 speech utterances, 1577 performed laughter, and 1636 speech and laughter utterances from 14 females and 56 males. This newly developed 'UNILAG Laugher Corpus' (ULC), reported in this work is available at http://laughter-db.herokuapp.com/. It is expected that this dataset will spur additional study in the subject, and that it will then be expanded in a gradual way across many application areas over time.

Some key aspects of laughter have been recognised as having the potential to be harnessed for individual identification. Laughter and its paralinguistic aspects have been identified as a biometric trait for determining a person's identity. The usefulness of laughter as a biometric, particularly in forensics, stems from its inability to be accurately mimicked due to its spontaneity when contrasted to speech.

A major strength of the study is that the use of laughter is significant in situations where speech is blurred, inaudible or unavailable. Laughter biometrics may be useful in access control with necessary provisions. A seeming weakness of the approach is that obtaining laughter samples is not as easy as collecting speech samples in humans. This study covers the development of a local laughter dataset, which security related organisations can use for purposes such as forensics. The work will spur further research interest and applications in the society.

ACKNOWLEDGMENT

The authors wish to acknowledge Raheem W.A, Ipinnimo O., Raji A.A. and Folorunso A.M. for their assistance.

REFERENCES

[1] Sudheer, K. (2010). Analysis of laugh signals for automatic detection and synthesis, *International Institute of Information Technology Hyderabad, India.*

[2] Urbain, J., & Dutoit, T. (2011). A phonetic analysis of natural laughter, for use in automatic laughter processing systems. In *International Conference on Affective Computing and Intelligent Interaction*, Springer, Berlin, Heidelberg, 397-406.

[3] Navarro, J., Fernandez Rosell, M., Castellanos, A., Del Moral, R., Lahoz-Beltra, R., & Marijuán, P. C. (2019). Plausibility of a neural network classifier-based neuroprosthesis for depression detection via laughter records. *Frontiers in neuroscience*, *13*, 267.

[4] Pietrowicz, M., Agurto, C., Casebeer, J., Hasegawa-Johnson, M., Karahalios, K., & Cecchi, G. (2019). Dimensional analysis of laughter in female conversational speech. In *ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, IEEE, 6600-6604.

[5] Ishi, C. T., Minato, T., & Ishiguro, H. (2019). Analysis and generation of laughter motions, and evaluation in an android robot. *APSIPA Transactions on Signal and Information Processing*, *8*, 1-10.

[6] Gosztolya, G., Beke, A., Neuberger, T., & Tóth, L. (2016). Laughter classification using Deep Rectifier Neural Networks with a minimal feature subset. *Archives of Acoustics*, *41*, 669-682.

[7] Kennedy, L. S. & Ellis, D. P. W. (2004). Laughter detection in meetings, in *NIST ICASSP 2004 Meeting Recognition Workshop, Montreal*, 118–121.

[8] Nergui, M., & Otake, M. (2013). Laughter detection for an assisting tool of group conversation. *methods*, *8*(9), 212-217.

[9] Jerome, U.(2014). Acoustic Laughter Processing, *TCTS Lab. Publ. Univ. Mons.PhD Thesis*.

[10] Bollepalli, B., Urbain, J., Raitio, T., Gustafson, J., & Cakmak, H. (2014). A comparative evaluation of vocoding techniques for hmm-based laughter synthesis. In *2014 IEEE international conference on acoustics, speech and signal processing (ICASSP)*, IEEE, 255-259.

[11] Urbain, J., Çakmak, H., & Dutoit, T. (2013). Automatic phonetic transcription of laughter and its application to laughter synthesis. In *2013 Humaine Association Conference on Affective Computing and Intelligent Interaction*, IEEE, 153-158.

[12] Weninger, F., & Schuller, B. (2012). Discrimination of linguistic and non-linguistic vocalizations in spontaneous speech: intra-and inter-corpus perspectives. In *Proceedings INTERSPEECH 2012, 13th Annual Conference of the International Speech Communication Association, 1-4*.
[13] Urbain, J., Niewiadomski, R., Bevacqua, E., Dutoit, T., Moinet, A., Pelachaud, C., Picart, B., Tilmanne, J. & Wagner, J. (2010). Avlaughtercycle. *Journal on Multimodal User Interfaces*, *4*(1), 47-58.
[14] Cakmak, H., Urbain, J., Tilmanne, J., & Dutoit, T. (2014). Evaluation of HMM-based visual laughter synthesis. In *2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP),* IEEE, 4578-4582.
[15] McKeown, G., Valstar, M., Cowie, R., Pantic, M., & Schroder, M. (2011). The semaine database: Annotated multimodal records of emotionally colored conversations between a person and a limited agent. *IEEE transactions on affective computing*, *3*(1), 5-17.
[16] Petridis, S., Martinez, B., & Pantic, M. (2013). The MAHNOB laughter database. *Image and Vision Computing*, *31*(2), 186-202.
[17] McKeown, G., Curran, W., Wagner, J., Lingenfelser, F., & André, E. (2015). The Belfast storytelling database: A spontaneous social interaction database with laughter focused annotation. In *2015 International Conference on Affective Computing and Intelligent Interaction (ACII),* IEEE, 166-172.
[18] Niewiadomski, R., Mancini, M., Baur, T., Varni, G., Griffin, H., & Aung, M. S. (2013). MMLI: Multimodal multiperson corpus of laughter in interaction. In *International Workshop on Human Behavior Understanding*, Springer, Cham, 184-195.
[19] Polychroniou, A., Salamin, H., & Vinciarelli, A. (2014). The SSPNet-Mobile Corpus: Social Signal Processing Over Mobile Phones. In *LREC*, 1492-1498.
[20] Knox, M. T., & Mirghafori, N. (2007). Automatic laughter detection using neural networks. In *Interspeech*, 2973-2976.
[21] Carletta, J., Ashby, S., Bourban, S., Flynn, M., Guillemot, M., Hain, T., Kadlec, J., Karaiskos, V., Kraaij, W., Kronenthal, M., Lathoud, G & Wellner, P. (2005). The AMI meeting corpus: A pre-announcement. In *International workshop on machine learning for multimodal interaction*, Springer, Berlin, Heidelberg, 28-39.
[22] El Haddad, K., Tits, N., & Dutoit, T. (2018). Annotating nonverbal conversation expressions in interaction datasets. In *Proceedings of Laughter Workshop*, 9-12.
[23] Jansen, M. P., Heylen, D. K., Truong, K. P., Englebienne, G., & Nazareth, D. S. (2018, September). The MULAI Corpus: Multimodal Recordings of Spontaneous Laughter in Dyadic Interaction. In *Proceedings of Laughter Workshop*, 58-63.
[24] Folorunso, C. O., Asaolu, O. S., & Popoola, O. P. (2019). A Review of Voice-Base Person Identification: State-of-the-Art. *Covenant Journal of Engineering Technology*, *3*(1), 38-60.
[25] Cakmak, H., Urbain, J., Dutoit, T., & Tilmanne, J. (2014, May). The AV-LASYN Database: A synchronous corpus of audio and 3D facial marker data for audio-visual laughter synthesis. In *LREC*, 3398-3403.
[26] Cleveland Clinic (2017), Voice Changes: What can they tell you as you age? Online Health Essential website. [Online]. Available: https://health.clevelandclinic.org/voice-changes-what-can-they-tell-you-as-you-age/. [Accessed: 05-May-2020].
[27] Boersma, P., & Van Heuven, V. (2001). Speak and unSpeak with PRAAT. *Glot International*, *5*(9/10), 341-347.
[28] Bachorowski, J. A., Smoski, M. J., & Owren, M. J. (2001). The acoustic features of human laughter. *The Journal of the Acoustical Society of America*, *110*(3), 1581-1597.
[29] Bachorowski, J. A., & Owren, M. J. (2001). Not all laughs are alike: Voiced but not unvoiced laughter readily elicits positive affect. *Psychological science*, *12*(3), 252-257.
[30] Verma, G. K. (2011, March). Multi-feature fusion for closed set text independent speaker identification. In *International conference on information intelligence, systems, technology and management*, Springer, Berlin, Heidelberg, 170-179.
[31] Raybaut, P. (2017). Spyder Documentation Release 3. *Available online at: pythonhosted. org*.
[32] Chamasemani, F. F., & Singh, Y. P. (2011, September). Multi-class support vector machine (SVM) classifiers--an application in hypothyroid detection and classification. In *2011 Sixth International Conference on Bio-Inspired Computing: Theories and Applications*, IEEE, 351-356.
[33] Truong, K. P., & Van Leeuwen, D. A. (2007). Automatic discrimination between laughter and speech. *Speech Communication*, *49*(2), 144-158.
[34] Science Buddies (2014). Singing Science: How High and Low Can You Go? - Scientific American. *Scientific American*. [Online]. Available: https://www.scientificamerican.com/article/singing-science-how-high-and-low-can-you-go/. [Accessed: 15-Sep-2021].
[35] National Centre for Voice and Speech (2012). Voice Classification and Life-Span Changes.[Online]. Available: http://www.ncvs.org/ncvs/tutorials/voiceprod/equation/chapter7/index.html. [Accessed: 01-Jun-2020].

 ISSN (online): 2645-2685; ISSN (print): 2756-6811

Appendix 1: Consent Form

DEPARTMENT OF SYSTEMS ENGINEERING,
FACULTY OF ENGINEERING,
UNIVERSITY OF LAGOS, AKOKA, YABA, LAGOS, NIGERIA.
CONSENT TO PARTICIPATE IN THE BUILDING OF AN INDIGENEOUS LAUGHTER DATABASE.

Introduction
- You are being invited to participate in a research study to build an indigenous Laughter database.
- This research is being conducted by the Biometrics Technology Research Group of the Departments of Systems Engineering, coordinated by Comfort Folorunso with Dr. O.S. Asaolu and Dr. O.P. Popoola as principal investigators.

Purpose of Study
- The purpose of the study is to develop a database of laughter produced by Africans (indigenous laughter).
- Eventually, this research may be published as part of a book chapter, presented as a paper, etc.
- We assure you that there is no known risk if you decide to participate in this research as there will not be any form of identification of the research participants.

Description of the Study Procedures
- If you agree to participate in this study, you will be asked to do the following things:
- You will be invited to the mass communication acoustic laboratory where the laughter samples will be recorded.
- You come in at the specified time, sit in front of a laptop and then series of jokes and comedy collections for this purpose will be displayed.
- Your laughter sample is being taken and recorded while you are having a nice time watching and reading the jokes.
- The total duration of the recording is between 5 to 10minutes, and you only need to come once for the length of the period of the recordings which will take place in the month of February 2018, at the first instance.

Risks/Discomforts of Being in this Study
- There is no risk attached to the study. Just your time is needed and that is appreciated.

Benefits of Being in the Study
- There is no direct benefit to the participants at the moment.

Confidentiality
- The records of this study will be kept strictly confidential. We will not include any information in any report we may publish that would make it possible to identify participant.

Right to Refuse or Withdraw
- The decision to participate in this study is entirely up to you. You may refuse to take part in the study *at any time* without affecting your relationship with the investigators of this study.

Right to Ask Questions and Report Concerns
- You have the right to ask questions about this research study and to have those questions answered by the research coordinator before, during or after the research. If you have any further questions about the study, at any time feel free to contact the coordinator, [*Folorunso Comfort]* at [comfortfolorunso@gmail.com] or by telephone at *08080719696*

Consent
- Your signature below indicates that you have decided to volunteer as a research participant for this study, and that you have read and understood the information provided above.

Subject's Name ………………………………………………………………………………..

Gender: Male [ ], Female [ ]; Age: 17-21 [ ]; 22-26 [ ]; 27-31 [ ]; 32-36 [ ]; 37-41 [ ]

Phone Number …………………………………………Email ……………………………………………

Subject's Signature …………………………………………….Date ………………………………

Coordinator's Signature ……………………………………….Date…………………………………